MALIC ENZYMES: METABOLOMICS AND PROTEIN INTERACTIONS

MALIC ENZYMES OF SINORHIZOBIUM MELILOTI: A STUDY OF METABOLOMICS AND PROTEIN-PROTEIN INTERACTIONS

By

LAURA ANNE SMALLBONE, B.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree

Master of Science

McMaster University

MASTER OF SCIENCE (2006) McMaster University
(Biology) Hamilton, Ontario

TITLE: Malic Enzymes of *Sinorhizobium meliloti*: A Study of Metabolomics and Protein-Protein Interactions.

AUTHOR: Laura Anne Smallbone, B.Sc. (McMaster University)

SUPERVISOR: Professor T.M. Finan

NUMBER OF PAGES: xiv, 134

ABSTRACT

Malic enzymes catalyze the oxidative decarboxylation of malate to pyruvate with the simultaneous reduction of a nicotinamide cofactor. It was previously reported that the nitrogen-fixing bacterium, *Sinorhizobium meliloti,* has two malic enzymes, a diphosphopyridine-dependent malic enzyme (DME) and a triphosphopyridine-dependent malic enzyme (TME). The *dme* gene is essential for symbiotic nitrogen-fixation in alfalfa root nodules and this symbiotic requirement cannot be met through increased expression of *tme*. In order to determine if a metabolic difference exists between the *dme* and *tme* mutants which might explain the symbiotic phenotypes, we conducted an analysis of intracellular and extracellular polar metabolomes. Differences noted between the intracellular profiles of the *dme* and *tme* mutant strains hinted at osmotic stress or a disturbance in central carbon metabolism. Extracellular studies indicated that *dme* mutant cells excreted at least 10-fold greater concentrations of both malate and fumarate. When considered together, the metabolic data implies that the DME enzyme is primarily responsible for the conversion of malate to pyruvate to generate acetyl-CoA whereas the TME enzyme must serve a secondary function within the cell.

While the C-terminal 320 amino acid regions from both DME and TME are similar in sequence to phosphotransacetylase enzymes, enzyme assays with DME and TME proteins have failed to detect PTA activity. Here we report that the chimeric malic enzyme structure is conserved among various gram negative bacteria including *Agrobacterium tumefaciens*, *Escherichia coli*, *Bradyrhizobium japonicum* and *Porphyromonas gingivalis*. Moreover these chimeric proteins are also present in the

archaebacteria, *Halobacterium salinarum* and *Haloarcula marismortui*. To further our understanding of the functions of DME and TME in *S. meliloti,* we have fused protein domains from DME to an affinity tag consisting of strepII and a calmodulin binding peptide. To identify proteins interacting with this fusion, we expressed these protein fusion constructs in *S. meliloti*, prepared extracts containing the soluble proteins and passed these through tandem affinity chromatography columns. All proteins that co-eluted with the fusion proteins appeared to be interacting with antibodies specific for the DME protein and so may have been aggregates or break-down products of DME.

ACKNOWLEDGMENTS

To begin, I would like to thank Dr. Turlough M. Finan for giving me the opportunity to work in his lab for the past few years and for his guidance and support. I would also like to thank the members of my supervisory committee, Dr. Elizabeth Weretilnyk and Dr. Christian Baron for their advice and support.

Special acknowledgments to the members of the Finan lab for all of their help- I never could have done it without you guys! Allyson MacLean for being my confidant, my shoulder to cry on and most importantly, my best friend; Rahat Zaheer for being my teacher and mentor; Branka Poduska for her endless support; Andrea Sartor for being a great friend; Jane Fowler for her kind words; Shawn MacLellan for sharing his knowledge with a little undergrad; Alison Cowie for her organizing and editing; Dr. Cheng for his wisdom; Kris Knorr for trying to teach me to not worry so much; Michelle Anstey for her friendship and for being a fellow BSB fan; Katie Taylor for the horse stories and for all her help with *S. meliloti* transconjugations; Tim Soh for his friendship; Cheryl Patten, Bridget Kelly, Chris Sibley and Zechun Yuan for their help and friendship; and many undergrads that have come in and out of the lab over the past few years.

A special thanks to Dr. Brian Golding and Weilong Hao for helping me with my bioinformatics, particularly tree-making.

I would also like to express my gratitude to the members of the Weretilnyk lab: Peter Summers for teaching me about GC-MS; Chris Wang for his patience and flexibility; Dave Geuvara, Mike Begora, Amber Gleason and Jeff Dedrick for making me laugh: "Laura did it"; and of course a slew of undergrads for joining in on the fun.

I would also like to thank all the wonderful people and amazing friends I have met, especially all the gorgeous girls for their support: Abha, Melanie Sr., Maria A., Andrea M., Melanie Jr., Karen, Danya, Mai, Heidi, Natalie, Leena, Diana, Maria P., Nadia, Veronica, Ferinne, Jen and Willow. I would also like to express my gratitude to the two BGSS crews, Pat Hayward and all the wonderful ladies in the main office.

I am also grateful to my "bestest" friend Katherine Lomas, who shares half my brain, for her support from London whenever I was in crisis mode. Also a big thanks to my parents for their love and continuous support; my sister, Emily, her boyfriend, Brenden and my brother, Colin for their love and companionship; my grandparents for their love and guidance (and financial support!); all my Aunts, Uncles and cousins for their love and shared laughter.

This work was funded by NSERC and Genome Canada.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

3PG	3-phosphoglycerate
6PS	6-phosphate sugar
ADP	adenosine diphosphate
AMDIS	Automated mass spectral deconvolution and identification system
ANOVA	analysis of variance
ATP	adenosine triphosphate
CBP	calmodulin binding peptide
DME	diphosphopyridine nucleotide-dependent malic enzyme
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DTT	DL-dithiotheitol
ED	Entner-Doudoroff
EDTA	ethylenedinitrilo tetraacetic acid
EGTA	ethylenebis(oxyethylenenitrilo)) tetraacetic acid
FA	fatty acid
GASP	GC-MS Data Analysis Software Program
GC-MS	gas chromatography mass spectrometry
IPTG	isopropyl-β-D-thiogalactopyranoside
LB	Luria-Bertani
M9	minimal salts media
MSTFA	N-Methyl-N-(trimethylsilyl)trifluoroacetamide
NAD+	nicotinamide adenine dinucleotide
NADP+	nicotinamide adenine dinucleotide phosphate
NCBI	National Centre for Biotechnology Information
NIST	National Institute of Standards and Technology
NMR	nuclear magnetic resonance
OAA	oxaloacetate
OD	optical density
ONPG	o-nitrophenyl β-D galactopyranoside
PBST	phosphate-buffered saline with Tween
PCR	polymerase chain reaction
PEP	phosphoenolpyruvate
PMSF	phenylmethylsulfonyl fluoride
PNPG	p-nitrophenyl β-D-glucuronide
PTA	phosphotransacetylase
PVDF	Polyvinylidene fluoride
RF	response factor
RNA	ribonucleic acid
RRF	relative response factor
RT	retention time
RI	retention index

SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
TBE	tris borate EDTA
TBST	tris-buffered saline with Tween
TCA	tricarboxylic acid cycle
TME	triphosphopyridine nucleotide-dependent malic enzyme

CHAPTER 1

Introduction: general information and metabolite profiling

Rhizobia, root nodule formation and nitrogen fixation

Sinorhizobium meliloti is a Gram negative alphaproteobacterium of the family *Rhizobiaceae* that is found in the soil and forms a symbiotic association with leguminous plants such as alfalfa (*Medicago sativa)*, sweet clover (*Melilotus*) and *Trigonella* species. Plants infected with *S. meliloti* display characteristic nodules which harbor a differentiated form of the bacteria known as bacteroids.

The root nodule has four zones which are characterized as follows (Vasse *et al*, 1990): Zone 1 is located at the tip of the nodule, is meristematic, consists of dividing plant cells and does not contain bacteria. Zone 2 is termed the infection zone and is where new bacteroids are released into the plant cells. Zone 3 is the site of nitrogen fixation and Zone 4 is the senescent zone where bacteroids are degraded.

The process of nodulation begins as the bacteria are attracted to the rhizosphere by phenolic compounds or flavonoids that are produced by the plant (as reviewed in Long, 1989). These flavonoids then induce the expression of nodulation genes (*nod)* in *S. meliloti* which leads to a series of morphological changes to the plant root, including root hair branching and curling followed by nodule formation (as reviewed in van Rhijn and Vanderleyden, 1995). The rhizobia attach to the plant then enter via infection threads and upon release into the plant cells, they remain enveloped in a plant cytoplasmic membrane

which is also known as the peribacteroid membrane. This differentiated form of *S. meliloti* is termed a bacteroid (as reviewed in van Rhijn and Vanderleyden, 1995).

Through the expression of *nif* and *fix* genes, *S. meliloti* bacteroids are capable of converting atmospheric dinitrogen (N_2) into ammonia (NH_3) ($N_2 + 8H^+ + 8e^- + 16ATP \rightarrow 2NH_3 + H_2 + 16ADP + 16P_i$) which can be used by the host plant to make proteins (as reviewed in Roberts and Brill, 1981). This conversion requires the breaking of a triple bond in the dinitrogen molecules and is catalyzed by the nitrogenase enzyme (*nif* genes) of *S. meliloiti* (as reviewed in Fischer, 1994). This process is valuable to the agricultural industry as legumes infected with *S. meliloti* do not need supplementation with nitrogen-containing fertilizer (which is costly).

Carbon metabolism in S. meliloti

In rhizobia, glycolysis is conducted via the Entner-Doudoroff (ED) pathway (Keele *et al*, 1969) and gluconeogenesis via the reverse Embden-Meyerhoff-Parnas (EMP) pathway, both of which are shown below in Figure 1-1. The ED pathway is induced in the presence of glucose but not in the presence of succinate nor in the bacteroid (Finan *et al*, 1988; Finan *et al*, 1991). Pyruvate, a product of glycolysis, may be shunted into the tricarboxylic acid cycle (TCA) through conversion to acetyl-CoA by the pyruvate dehydrogenase enzyme. In equimolar amounts, acetyl-CoA can combine with oxaloacetate (OAA) via the citrate synthase enzyme to make citrate (Figure 1-1).

Figure 1-1: Tricarboxylic Acid Cycle (TCA) and related pathways

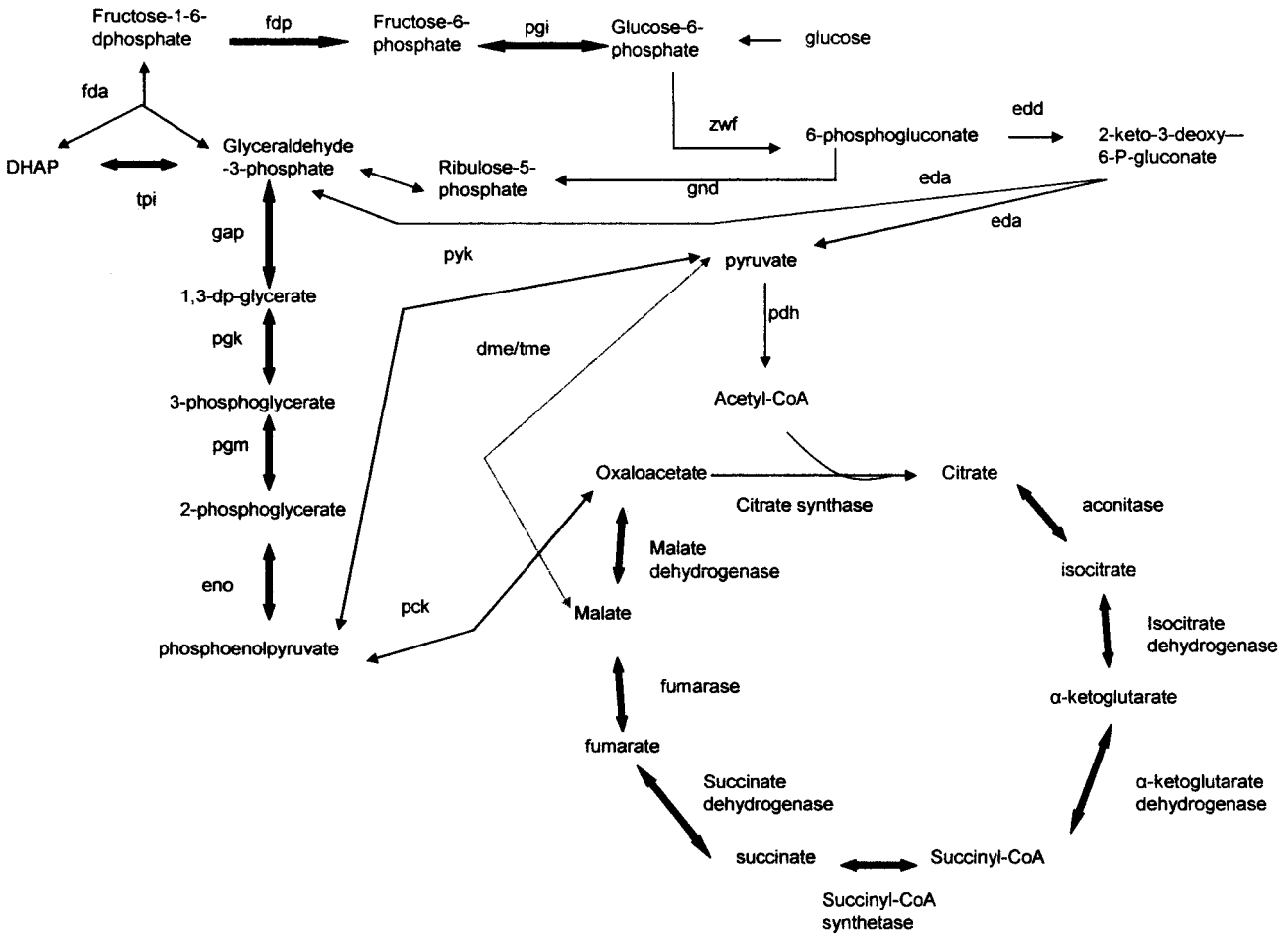


Figure 1-1: TCA cycle and related pathways. Abbreviations: DHAP (dihydroxyacetone phosphate), dme (diphosphopyridine-dependent malic enzyme), *eda2* (2-keto-3-deoxy-6-phosphogluconate aldolase), *edd* (6-phosphogluconate dehydratase), *eno* (enolase), *fda* (fructose-1-6-P_2 aldolase), *fdp* (fructose-1,6-P_2 phosphatase), *gap* (glyceraldehydes-3-phosphate dehydrogenase), *gnd* (6-phosphogluconate dehydrogenase), *pck* (phosphoenolpyruvate carboxykinase), pdh (pyruvate dehydrogenase), *pgi* (phosphoglucose isomerase), *pgk* (3-phosphoglycerate kinase), *pgm* (phosphoglycerate mutase), *pyk* (pyruvate kinase), tme (triphosphopyridine-dependent malic enzyme), *tpi* (triose phosphate isomerase), *zwf* (glucose-6-phosphate 1-dehydrogenase)

The primary carbon source of *S. meliloti* bacteroids consists of C_4-dicarboxylic acids, such as malate, succinate or fumarate (reviewed in Vance and Heichel, 1991), these organic acids are supplied by the plant to the bacteroids and are then taken up via the Dct (dicarboxylic acid transport) system (Yarosh *et al*, 1989). In order to sustain the activity of the TCA cycle when C_4-dicarboxylic acids constitute the sole carbon source, acetyl-CoA must be generated so that citrate can be synthesized from OAA. In bacteroids, malic enzymes catalyze the oxidation of malate to pyruvate which, via pyruvate dehydrogenase, may be used to synthesize acetyl-CoA.

Malic Enzymes

Malic enzymes were first isolated from pigeon liver (Ochoa *et al*, 1948) and are found in a diverse spectrum of organisms including animals, plants, bacteria and archea. In the presence of a divalent metal cation (Mg^{2+}), malic enzymes catalyze the oxidative decarboxylation of malate to pyruvate while simultaneously reducing a nicotinamide co-factor (NAD+ and/or NADP+) and producing CO_2 as a by-product (Figure 1-2). The International Union of Biochemistry classified L-malic enzymes into three groups based upon co-factor requirement and ability to decarboxylate OAA; NAD+-dependent class (EC 1.1.1.38) which can decarboxylate OAA; NAD+-dependent class (EC 1.1.1.39) which is not able to decarboxylate OAA and the strictly NADP+-dependent class, (EC 1.1.1.40) which is able to decarboxylate OAA. Note that the enzymes of classes EC 1.1.1.38 and EC 1.1.1.39 preferentially use NAD+ as a cofactor but may also use

NADP+. The EC 1.1.1.38 class is found in plants and bacteria, the EC1.1.1.39 class only in plant mitochondria and the EC1.1.1.40 class in plants, animals and bacteria (As reviewed in Wedding, 1989).

Figure 1-2: Reaction catalyzed by the malic enzymes

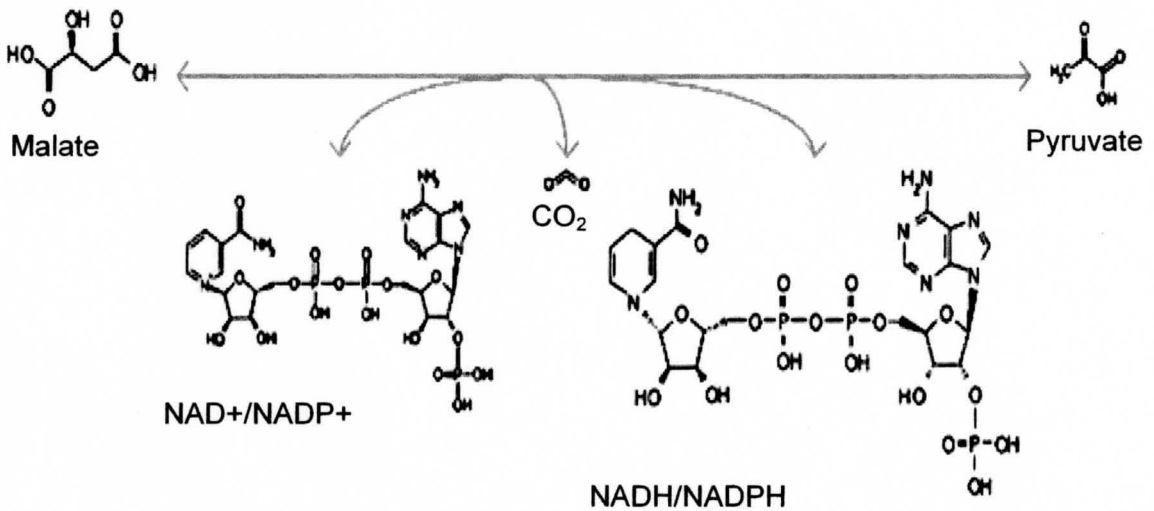


Figure 1-2: Oxidative decarboxylation of malate to pyruvate with the simultaneous reduction of a nicotinamide cofactor, NAD+ or NADP+.

In eukaryotic organisms, malic enzymes are classified based on their location within the cell (As reviewed in Frenkel, 1975). Malic enzymes located within the cytosol of liver cells are thought to generate NADPH- reducing power to fuel lipogenesis (Ma *et al*, 1990). The other two classes are localized to the mitochondrion; one being NAD+-dependent and the other being NADP+-dependent (As reviewed in Frenkel, 1975). The NAD+-dependent isoform is involved in the conversion of amino acids to pyruvate, providing cells with an energy source alternative to glucose (Skorkowski and Storey,

1990). Interestingly, concentrations of this form of the enzyme was found by Loeber *et al* (1991) to be elevated in tissues with high rates of cell division, including human tumor cells. The mitochondrial NADP+ -dependent isoform may be involved in the cycling of NADPH into the mitochondrion (Wedding, 1989).

Prokaryotic malic enzymes

Malic enzymes have been studied extensively in *Escherichia coli.* This species is known to have two distinct malic enzymes which require NAD+ (EC1.1.1.38, *sfcA*) and NADP+ (EC1.1.1.40), respectively as co-factors (Katsuki *et al*, 1967; Sanwal and Smando, 1969; Yamaguchi *et al*, 1973; Mahajan *et al*, 1990;). In 1970, Sanwal hypothesized that the NADP+-dependent malic enzyme in *E. coli* has a biosynthetic role and the NAD+-dependent enzyme has a catabolic role (Sanwal, 1970). The activities of both of these enzymes were repressed by glucose and induced by malate (Murai *et al*, 1971), however, the NADP+-dependent enzyme was found to be allosterically inhibited by acetyl-CoA and OAA (Sanwal *et al*, 1968) whereas the activity of the NAD+ -dependent enzyme was inhibited by CoA and activated by aspartate (Sanwal, 1970b).

Malic enzymes have also been characterized in several other species, including *Bacillus stearothermophilus* (Kobayashi *et al*, 1989), *Clostridium thermocellum* (Lamed and Zeikus, 1981), *Pseudomonas putida,* (Garrido-Pertierra *et al*, 1983), *Pseudomonas diminuta* (Suye *et al*, 1992), *Sulfolobus solfataricus* (Bartolucci *et al*, 1987) and *Streptococcus bovis* (Kawai et al, 1996). More recently, an NADP+-dependent malic enzyme has been purified from the hyperthermophilic archaeon *Thermococcus*

kodakaraensis KOD1 (Fukuda *et al,* 2005) and one of each NAD+-dependent and NADP+ dependent enzymes have been purified from *Mesorhizobium ciceri* CC 1192 (Tabrett and Copeland, 2002),

Within the past few years, the volume of genomic sequence data available on the National Centre for Biotechnology Information (NCBI) website has increased exponentially. This extensive library of information has identified hypothetical malic enzyme genes in many species of bacteria and archaea including *Bacillus subtilis* (Kunst *et al,* 1997), *Rickettsia prowasekii* (Andersson *et al,* 1998), *Haloarcula marismortui* (Baliga *et al*, 2004) and many others.

Malic enzymes in Rhizobial species

Bradyrhizobium japonicum, a species involved in the nodulation of soybean plants, possesses both NAD+ and NADP+-dependent malic enzymes (Copeland *et al,* 1988; Kouchi *et al*, 1988). Activity of the NADP+-dependent enzyme was found to be inhibited by acetyl CoA and OAA (Chen *et al*, 1997). *Rhizobium leguminosarum,* a species capable of pea plant nodulation, was likewise found to have two malic enzymes, one of each NAD+-dependent and NADP+-dependent (McKay *et al,* 1988).

S. meliloti has two malic enzymes, diphosphopyridine-dependent malic enzyme (DME) and triphosphopyridine-dependent malic enzyme (TME) which require NAD+ and NADP+, respectively, as co-factors (Driscoll and Finan, 1993; Driscoll and Finan, 1996). The pyruvate generated through the malic enzyme reaction can be converted into

acetyl-CoA which can be shunted into the TCA cycle to generate the energy required to sustain the nitrogenase reaction (Driscoll and Finan, 1993). Driscoll and Finan, (1993) found that the presence of a functional *dme* gene is required for the process of nitrogen fixation in *S. melitoti*. Plants infected with *dme* knockout mutants of *S. meliloti* are small and chlorotic (Fix-) (Driscoll and Finan, 1993). However, despite encoding an enzyme with an ability to execute the same chemical reaction as the *dme* gene product, the *tme* gene is not essential for nitrogen fixation in *S. meliloti* (Driscoll and Finan, 1996) and the reason for this difference is largely unknown. It has been shown that *tme* gene expression is reduced by 80% in bacteroids compared to free-living cells whereas *dme* gene expression remains unchanged (Driscoll and Finan, 1997). However, DME cannot be functionally replaced by TME as overexpression of the TME enzyme in a *dme* knockout background still yields plants of a Fix- phenotype (Cowie, 1998).

Driscoll and Finan, (1993, 1996) demonstrated that the DME and TME enzymes share many biochemical properties, however, it was noted that DME is subject to greater inhibition compared to TME in the presence of acetyl-CoA (Driscoll, 1995). This result lends evidence to the hypothesis that DME plays a role in the synthesis of acetyl CoA to power the TCA cycle thereby fueling nitrogen fixation. The lack of inhibition of TME by acetyl-CoA suggests that TME may participate in another pathway such as the generation of NADPH for biosynthetic processes (Voegele *et al*, 1999). To address the role of DME and TME enzymes in central carbon metabolism in *S. meliloti,* we have begun an analysis of the polar metabolome of malic enzyme mutants using gas chromatography-mass

spectrometry (GC-MS). It is hoped that these studies will give insight into the requirement of DME for nitrogen fixation.

Metabolite profiling

Metabolomics has been defined as "the study of the metabolite profiles of a given system (cell, tissue or organism) generated under a defined set of conditions" (as reviewed in Rochfort, 2005). Such studies are more representative of the phenotype of a given system compared to transcriptomic and proteomic studies. Metabolites represent the last products generated in the sequence of "DNA to RNA to protein to metabolite", and are thus representative of all influences of regulation that have occurred along the way (as reviewed in Rochfort, 2005).

Metabolite profiling is broadly applicable and has been used in studies of toxicology, plant metabolism and functional genomics (as reviewed in Rochfort, 2005). Notable recent examples of the application of metabolomics include: the study of flavonoid biosynthesis in *Populus* trees (Morreel *et al,* 2006); the study of heat stress in *Drosophila* (Malnedal *et al,* 2006); understanding mouse models of diseases (ie: cardiovascular and neurodegenerative) using high-throughput metabolite profiling (Griffin, 2006) and clinical applications for humans which may soon help tailor food, drug and lifestyle choices to meet the needs of an individual (German, 2005).

Although the field of metabolomics is still in its infancy, a few studies which involve intracellular metabolic profiling have been conducted in bacterial species.

Studies of salt stress in *Desulfovibrio vulgaris* using metabolite profiling have shown increases in branched fatty acids, proline, glycine-betaine and glutamate (Mukhopadhyay *et al*, 2006). Ecotype differentiation in *Bacillus cereus* between avirulent lab strains and clinical isolates was conducted using NMR (nuclear magnetic resonance imaging)-generated metabolic profiles (Bundy *et al*, 2005). Finally, employing methods very similar to those used in this study, Barsch *et al* (2004) have conducted an extensive study of the metabolome of *S. meliloti* in order to help understand its symbiosis with the plant *Medicago truncatula.* This group compared profiles of free-living cells that were grown in a variety of carbon sources and found many common amino acids, sugars, sugar alcohols and organic acids to be present intracellularly.

Studies of the extracellular metabolome (referred to as an exometabolome or a footprint) have been less common but possess a clear advantage over intracellular studies as the intracellular environment is dynamic (Villas *et al,* 2006) and thus quenching of metabolism must be performed quickly and efficiently. This is not always possible due to human error and procedural constraints (ie: centrifugation time) and these issues likely lead to inconsistencies and artifacts in observed metabolite profiles. The first exometabolome studies were reported in yeasts by Allen *et al* (2003) using electrospray ionization-mass spectrometry to differentiate between mutant and wild type strains metabolically. The only other published footprinting study was conducted quite recently to assess fiber degradation in *Clostridium proteoclasticum* and *S. bovis* (Villas *et al,* 2006). Villas *et al* (2006) used microwave radiation to increase the derivatization rate of samples thereby ameliorating yield.

This Work

As previously discussed, *S. meliloti* has two malic enzymes, DME and TME which require NAD+ and NADP+, respectively, as co-factors. The DME enzyme is required for symbiotic nitrogen fixation whereas the TME enzyme is not and the reason for this difference is unknown (Driscoll and Finan 1993; Driscoll and Finan 1996). More generally, the roles of malic enzymes in central carbon metabolism are not clear. Both *dme* and *tme* mutants show no free-living phenotypes and the roles attributed to these enzymes are largely inferred by their properties rather then *in vivo* measurements. To address why DME, but not TME is required for nitrogen fixation, we initiated an analysis of the polar metabolome of malic enzyme mutants using gas chromatography-mass spectrometry (GC-MS).

We began with an analysis of the intracellular polar metabolome of cultures grown in M9-succinate; several differences were noted between the *dme* and *tme* mutant strains most of which hinted at osmotic stress. We then assessed to see if the observed phenotypes were growth-phase or carbon source dependent.

Next, a study of the extracellular polar metabolite profiles of the malic enzyme mutants was conducted and we found that *dme* cells excrete 10-fold greater concentrations of both malate and fumarate into their extracellular environment compared to *tme* mutants.

To date there have been no published reports of exometabolome analysis in *S. meliloti*. In fact, supernatant studies are difficult to execute due to the high phosphate

content of both rich and minimal growth media. Organic phosphate molecules are derivatized and will produce very large peaks on the gas chromatogram, occluding surrounding peaks and causing problems with detection. The relatively high concentration of carbon sources present in the minimal media may also result in large peaks. As this is one of the first studies to employ analysis of the extracellular environment, the techniques used in this study could prove useful to research in many other fields of microbiology. Understanding the extracellular environment of other microorganisms will advance research in medical microbiology and provide useful applications to industry in the overproduction of compounds which are useful for human consumption (commercial or otherwise).

Both the intracellular and extracellular data suggest that the DME enzyme is primarily responsible for the conversion of malate to pryruvate whereas the TME enzyme must be executing another function in the cell.

CHAPTER 2

Materials and methods: metabolite profiling

Plasmids and Bacterial Strains

All plasmids and bacterial strains used in this study are described in Table 2-1. Frozen stocks of bacterial strains were made in glass vials by adding 600uL of dense overnight culture to 600uL of LB (Luria Broth) with 14% dimethylsulfoxide (DMSO) as a cryoprotectant. The vials were stored at -80°C.

Table 2-1: Bacterial strains and plasmids used in this study

Sinorhizobium meliloti strains

Strain	Genotype	Reference
Rm1021	SU47, str-21	Meade *et al.* 1982
RmG454	Rm1021, *dme-2*::Tn5	Driscoll and Finan 1993
RmG455	Rm1021 *dme-3*::Tn5	Driscoll and Finan 1993
RmG456	Rm1021, *dme1*::Tn5	Driscoll and Finan 1993
RmG927	Rm1021, *tme1*::Tn5	Driscoll and Finan 1996
RmG994	Rm1021, *dme-3*::Tn5 *tme-4*::ΩSp	Driscoll and Finan 1996
RmG995	Rm1021, *tme-4*::ΩSp	Driscoll and Finan 1996
RmH215	Rm1021, *tme-2*::Tn5	Driscoll and Finan 1996
RmK990	Rm1021 (pTH1565)	strain collection
RmK991	Rm1021 (pTH1566)	strain collection
RmP110	Rm1021 with wild type *Pst*C	Yuan et al 2006
RmP738	RmP110 (pTH1918)	This study
RmP739	RmP110 (pTH1930)	This study
RmP740	RmP110 (pTH1931)	This study
RmP741	RmP110(pTH1932)	This study
RmP844	RmP110 *dme-3*::Tn5	This study
RmP848	RmP844 (pTH1918)	This study
RmP849	RmP844 (pTH1930)	This study
RmP850	RmP844 (pTH1981)	This study
RmP851	RmP110 (pTH1981)	This study
RmP1126	Rm1021 (pFL383)	This study
RmP1127	Rm1021 (pFL634)	This study
RmP1128	Rm1021 (pFL1372)	This study
RmP1129	Rm1021 (pFL1690)	This study
RmP1130	Rm1021 (pFL1839)	This study
RmP1131	Rm1021 (pFL2028)	This study
RmP1132	Rm1021 (pFL2337)	This study

Sinorhizobium meliloti strains

Strain	Genotype	Reference
RmP1133	Rm1021 (pFL2728)	This study
RmP1134	Rm1021 (pFL2754)	This study
RmP1135	Rm1021 (pFL2807)	This study
RmP1136	Rm1021 (pFL4122)	This study
RmP1137	Rm1021 (pFL4215)	This study
RmP1138	Rm1021 (pFL4767)	This study
RmP1139	Rm1021 (pFL4903)	This study
RmP1140	Rm1021 (pFL5382)	This study
RmP1141	Rm1021 (pFL5643)	This study
RmP1142	Rm1021 (pTH2068)	This study
RmP1143	Rm1021 (pTH2069)	This study
RmP1144	Rm1021 (pTH2070)	This study
RmP1145	Rm1021 (pTH2081)	This study
RmP1146	Rm1021 (pTH2082)	This study
RmP1147	RmG455 (pFL383)	This study
RmP1148	RmG455 (pFL634)	This study
RmP1149	RmG455 (pFL1372)	This study
RmP1150	RmG455 (pFL1690)	This study
RmP1151	RmG455 (pFL1839)	This study
RmP1152	RmG455 (pFL2028)	This study
RmP1153	RmG455 (pFL2337)	This study
RmP1154	RmG455 (pFL2728)	This study
RmP1155	RmG455 (pFL2754)	This study
RmP1156	RmG455 (pFL2807)	This study
RmP1157	RmG455 (pFL4122)	This study
RmP1158	RmG455 (pFL4215)	This study
RmP1159	RmG455 (pFL4767)	This study
RmP1160	RmG455 (pFL4903)	This study
RmP1161	RmG455 (pFL5382)	This study

Sinorhizobium meliloti strains

Strain	Genotype	Reference
RmP1162	RmG455 (pFL5643)	This study
RmP1163	RmG455 (pTH2068)	This study
RmP1164	RmG455 (pTH2069)	This study
RmP1165	RmG455 (pTH2070)	This study
RmP1166	RmG455 (pTH2081)	This study
RmP1167	RmG455 (pTH2082)	This study
RmP1168	RmG995 (pFL383)	This study
RmP1169	RmG995 (pFL634)	This study
RmP1170	RmG995 (pFL1372)	This study
RmP1171	RmG995 (pFL1690)	This study
RmP1172	RmG995 (pFL1839)	This study
RmP1173	RmG995 (pFL2028)	This study
RmP1174	RmG995 (pFL2337)	This study
RmP1175	RmG995 (pFL2728)	This study
RmP1176	RmG995 (pFL2754)	This study
RmP1177	RmG995 (pFL2807)	This study
RmP1178	RmG995 (pFL4122)	This study
RmP1179	RmG995 (pFL4215)	This study
RmP1180	RmG995 (pFL4767)	This study
RmP1181	RmG995 (pFL4903)	This study
RmP1182	RmG995 (pFL5382)	This study
RmP1183	RmG995 (pFL5643)	This study
RmP1184	RmG995 (pTH2068)	This study
RmP1185	RmG995 (pTH2069)	This study
RmP1186	RmG995 (pTH2070)	This study
RmP1187	RmG995 (pTH2081)	This study
RmP1188	RmG995 (pTH2082)	This study
RmP1189	RmG994 (pFL383)	This study
RmP1190	RmG994 (pFL634)	This study

Sinorhizobium meliloti strains

Strain	Genotype	Reference
RmP1191	RmG994 (pFL1372)	This study
RmP1192	RmG994 (pFL1690)	This study
RmP1193	RmG994 (pFL1839)	This study
RmP1194	RmG994 (pFL2028)	This study
RmP1195	RmG994 (pFL2337)	This study
RmP1196	RmG994 (pFL2728)	This study
RmP1197	RmG994 (pFL2754)	This study
RmP1198	RmG994 (pFL2807)	This study
RmP1199	RmG994 (pFL4122)	This study
RmP1200	RmG994 (pFL4215)	This study
RmP1201	RmG994 (pFL4767)	This study
RmP1202	RmG994 (pFL4903)	This study
RmP1203	RmG994 (pFL5382)	This study
RmP1204	RmG994 (pFL5643)	This study
RmP1205	RmG994 (pTH2068)	This study
RmP1206	RmG994 (pTH2069)	This study
RmP1207	RmG994 (pTH2070)	This study
RmP1208	RmG994 (pTH2081)	This study
RmP1209	RmG994 (pTH2082)	This study

Escherichia coli strains

Strain	Genotype	Reference
DH5α	F-, *endA1*, *hsd*R17, (r_k-, m_k-), *sup*E44, *thi*-1, *recA*1, *gyr*A96, *relA1*, *Δ(arg*F- *lac*ZYA) U169, φ80d*lac*Z, ΔM15	BRL Inc
J291	DH5α (pTH452)	Cowie, 1998
MT607	MM294A *recA*-56	Finan *et al.* 1986
MT616	MT607/pRK2013 *npt*::Tn9	Finan *et al.* 1986
EJ1321	*gal*K2, *pck*, *dme*, *tme*	Hansen and Juni 1975
M820	DH5α (pTH1918)	This study
M832	DH5α (pTH1930)	This study
M833	DH5α (pTH1931)	This study
M886	DH5α (pTH1981)	This study
M887	EJ1321 (pTH1918)	This study
M888	EJ1321 (pTH1930)	This study
M889	EJ1321 (pTH1931)	This study
M890	EJ1321 (pTH1981)	This study
M976	DH5α (pTH2068)	This study
M977	DH5α (pTH2069)	This study
M978	DH5α (pTH2070)	This study
M992	DH5α (pTH2081)	This study
M993	DH5α (pTH2082)	This study
EcFL383	DH5α (pFL383)	Cowie *et al* 2006
EcFL634	DH5α (pFL634)	Cowie *et al* 2006
EcFL1372	DH5α (pFL372)	Cowie *et al* 2006
EcFL1690	DH5α (pFL1690)	Cowie *et al* 2006
EcFL1839	DH5α (pFL1839)	Cowie *et al* 2006
EcFL2028	DH5α (pFL2028)	Cowie *et al* 2006
EcFL2337	DH5α (pFL2337)	Cowie *et al* 2006
EcFL2728	DH5α (pFL2728)	Cowie *et al* 2006
EcFL2754	DH5α (pFL2754)	Cowie *et al* 2006
EcFL2807	DH5α (pFL2807)	Cowie *et al* 2006

Escherichia coli strains

Strain	Genotype	Reference
EcFL4122	DH5α (pFL4122)	Cowie *et al* 2006
EcFL4215	DH5α (pFL4215)	Cowie *et al* 2006
EcFL4767	DH5α (pFL4767)	Cowie *et al* 2006
EcFL4903	DH5α (pFL4903)	Cowie *et al* 2006
EcFL5382	DH5α (pFL5382)	Cowie *et al* 2006
EcFL5643	DH5α (pFL5643)	Cowie *et al* 2006

Plasmids

Plasmid	Genotype	Reference
pRK600	pRK2013 *npt*::Tn9, Cm^R	Finan *et al* 1986
pTH1703	MCS gfp in pTH1591 Gm^R	Cowie *et al* 2006
pTH1918	C-terminus of *dme* cloned into pTrcSC via PacI	This study
pTH1930	N-terminus of *dme* cloned into pTrcSC via PacI	This study
pTH1931	pTrcSC, expression vector, strepII and CBP tags, $SmSp^R$	A gift from Dr. Baron
pTH1981	full *dme* gene into pTrcSC via PacI	This study
pTH2068	*edd* gene into pTH1703 via ApaI/XhoI	This study
pTH2069	*eda2* gene into pTH1703 via ApaI/XhoI	This study
pTH2070	*cbbF* gene into pTH1703 via ApaI/XhoI	This study
pTH2081	*gnd* gene into pTH1703 via XhoI/NsiI	This study
pTH2082	*gap* gene into pTH1703 via XhoI/ApaI	This study
pFL383	pTH1522 mcs: Smc02037::*gusA rfp* (*edaI* contained in operon)	Cowie *et al* 2006
pFL634	pTH1522 mcs: *edaI*::*gusA rfp*	Cowie *et al* 2006
pFL1372	pTH1522 mcs: *eno*::*gpf lacZ*	Cowie *et al* 2006
pFL1690	pTH1522 mcs: *glk*::*gusA rfp*	Cowie *et al* 2006
pFL1839	pTH1522 mcs: Smc0680::*gfp lacZ*	Cowie *et al* 2006
pFL2028	pTH1522 mcs: *glgC*::*gfp lacZ* (*pgm* contained in operon)	Cowie *et al* 2006
pFL2337	pTH1522 mcs: Smc02982::gfp *lacZ*	Cowie *et al* 2006
pFL2728	pTH1522 mcs: Smc01615::gfp *lacZ*	Cowie *et al* 2006
pFL2754	pTH1522 mcs: Smb20574::*gusA rfp*	Cowie *et al* 2006
pFL2807	pTH1522 mcs: *glgP*::*gusA*	Cowie *et al* 2006
pFL4122	pTH1522 mcs: *tpiA1*:: *gfp lacZ*	Cowie *et al* 2006
pFL4215	pTH1522 mcs: *pckA*::*gusA rfp*	Cowie *et al* 2006
pFL4767	pTH1522 mcs: *cbbF*:: *gusA rfp*	Cowie *et al* 2006
pFL4903	pTH1522 mcs: Sma0682::*gfp lacZ*	Cowie *et al* 2006
pFL5382	pTH1522 mcs: *pykA*::*gusA rfp*	Cowie *et al* 2006
pFL5643	pTH1522 mcs: *zwf*::*gfp lacZ*	Cowie *et al* 2006

Media preparation and conditions for growth of bacterial strains

S. meliloti and *E.coli* strains were grown in shaking or rotating incubators at 30°C and 37°C respectively.

Luria Bertani broth (LB) which consists of 10g/L Difco Bacto Tryptone, 5g/L Difco Yeast extract and 5g/L NaCl, was used as a rich media source for both *S. meliloti* and *E. coli.* The pH of the solution was then adjusted to 7.0 using 1N NaOH added drop wise. If solid media was required, Difco Agar to 1.5% was added for each liter of the solution following pH adjustment. The media was then autoclaved to ensure adequate sterilization. Growth of *S. meliloti* in liquid LB media required supplementation with 2.5mM of each $MgSO_4$ and $CaCl_2$ (LBmc)

To make minimal medium for *S. meliloti,* the following compounds were added to autoclaved, distilled water: 1mM $MgSO_4$, 0.25mM $CaCl_2$, 1μg/mL D-biotin, and 0.01μg/mL $CoCl_2$. Carbon sources (i.e.: glucose (1M), succinate (1M), glycerol (30%), malate (1M) or arabinose (1M); stock solution concentrations are given in brackets) were filter sterilized through 0.45μm supor membrane (Acrodisc) and then added to the media to final concentrations of 15mM. A stock solution of 5x Difco M9-salts (33.9g/L Na_2HPO_4, 15g/L NaH_2PO_4, 2.5g/L sodium chloride NaCl and 5 g/L ammonium chloride NH_4Cl) was autoclaved separately and then diluted to 1x when added to the minimal media solution.

Antibiotics were purchased from Sigma and stock solutions were made in H_2O or ethanol and stored at -20°C. The final concentrations of antibiotics in μg/mL for *S. meliloti* are as follows: gentamicin (60), neomycin (200-300), rifampicin (20),

spectinomycin (200), streptomycin (200). In *E. coli:* chloramphenicol (10), gentamicin (10), ampicillin (100), spectinomycin (100), streptomycin (100) and Tetracylcine (5). For liquid media, all concentrations were halved, with the exception of ampicillin which remains at 100μg/mL. This is due to the presence of β-lactamase which is secreted by *E. coli* and reduces the concentration of the ampicillin in solution.

Bacterial Matings

Overnight cultures of donor strain (generally *E. coli*), recipient strain (*S. meliloti*) and mobilizing strain (MT616) were centrifuged at 16,100g for one minute and then washed in 0.85% NaCl. The pellet was then resuspended in 0.85% NaCl and 20μL of each donor, mobilizer and recipient was spotted onto LB and incubated face-up at 30°C overnight. The following day, the entire spot was resuspended in 0.85% NaCl and a 10-fold dilution series was made in 0.85% NaCl, plated onto selective media and incubated at 30°C.

Growth curves

2mL cultures of the strain in question were inoculated in tripicate in LBmc media and grown overnight. The following day, replicates were subcultured into 5mL of specific media (i.e.: M9-succinate) to an OD_{600nm} between 0.01 and 0.05. Starting optical density (OD_{600nm}) readings were taken with the Baush and Lomb Spectronic 20

spectrophotometer. OD_{600nm} readings were taken at one to two hour intervals during the exponential phase of the culture's growth. Once the cultures reached the log phase, OD_{600nm} readings were taken every four to six hours. OD_{600nm} readings were plotted log (OD_{600nm}) versus time using Excel.

Growth conditions for intracellular metabolite analysis.

Three to five single colony replicates taken from agar plates <10 days old were grown as overnight cultures in 2mL of LBmc and an appropriate antibiotic (if required). These replicates were subcultured into 50mL of the desired minimal media, generally a 1:200 dilution into a 250mL flask which had been rinsed 10 times in distilled water and subsequently autoclaved. The 50mL cultures were then incubated at 30°C with shaking (200rpm) and grown to the desired OD_{600nm}, generally between 0.2 to 0.8 (typically 16 hours). Following incubation, 30-50mL of the cells (depending on OD_{600nm}) were harvested by centrifugation for 20 minutes at 3730g (SX4250 rotor) at 4°C The cell pellets were resuspended in 1mL of distilled water, flash frozen in liquid nitrogen and stored at -80°C.

Extraction of intracellular polar metabolites.

Procedures for both extraction and derivatization were developed by Weretilnyk *et al* using methods reported by Fiehn *et al* (2000) as a template. Frozen cells were thawed

and an internal standard, ribitol was added to a final concentration of approximately 3ng/mL (50µL of a 83ng/µL stock of ribitol was added to the thawed cells). NaCl was also added to a final concentration of 29mM (40µL of a 1M stock solution) and the samples were vortexed to mix thoroughly. The cells were then transferred into screw-cap tubes containing 0.75mL of 0.1mm glass beads (Biospec products), the tube was washed with 400µL of 100% methanol and then the cells were lysed using the Mini Beadbeater-8 three times (one minute each time) with one minute incubation on ice following each beating. After lysis, the screw cap tubes were centrifuged at 14 000g and 4°C for five minutes and the supernatants were transferred to chilled 15mL glass Corex tubes. The beads were then washed with 1mL of 60% methanol and then with 400µL of 100% methanol. After each wash, the tubes were centrifuged at 14 000g and 4°C for five minutes and the supernatants removed and added to the chilled Corex tubes. The resulting lysates were incubated at 70°C with gentle shaking in the Boeker Grant ORS 200 Shaker for 15 minutes and then centrifuged at 14000g (JA-20 Rotor) at 4°C for 15 minutes. Subsequently, the supernatant was transferred to a chilled 16x100mm glass culture tube and placed on ice. To the pellet in the Corex tube, 750µL of 100% chloroform was added, the tubes were vortexed and the resulting supernatant was added to the 16x100mm glass culture tube and vortexed once more. This mixture was then centrifuged in a swing-out table top centrifuge for five minutes at 226g in a table top centrifuge with a swing-out rotor at room temperature. The aqueous phase (~1.75mL) of the mixture was transferred into a new 16x100mm glass culture tube and 1mL was put into a 1mL Wheaton V-vial with a Teflon liner with 7 small holes made using a B-D

precision glide needle 25G5/8. The volume in the V-vial was reduced by half by drying with a stream of nitrogen gas using the N-Evap analytical evaporator. The extracted samples were lyophilized overnight on the Labconco Freezone Plus 6. Following freeze-drying, the caps on the V-vial were replaced with screw caps lined with teflon with no hole in the top.

Derivatization of *S. meliloti* polar metabolites.

5μL of 5x diluted fatty acid (FA) standards (10mg/mL of each FA) (Table 2-2) (Sigma) were added to the lyophilized samples which were subsequently dried down with a stream of nitrogen gas using the N-Evap. 50μL of 20mg/mL Methoxyamine hydrochloride (Sigma) dissolved in pyridine was then added to the samples which were vortexed and incubated at 30°C for 90 minutes. Next, 50μL of N-methyl-N-trimethylsilyltrifluoroacetamide (MSTFA) (Pierce) was added and the samples were vortexed and incubated for 30 minutes at 37°C. The samples were then transferred into GC vials and run on the GC-MS instrument (Trace GC 2000/Trace DSQ (Thermo Finigan, Gerstel)). Between the addition of the fatty acid standards, methoxyamine or MSTFA or the transfer of separate samples into GC vials, the syringes were rinsed ten times in the solvent of the component that was added, (tetrahydrofuran for the fatty acid standards and pyridine for methoxyamine and MSTFA).

Table 2-2: Characteristics of fatty acid standards used in this study

Carbon Number	m/z ratio for major molecular ion
C7	187
C9	215
C11	243
C13	271
C15	299
C19	355
C23	411
C27	467
C31	523

Running of samples on the GCMS instrument

All samples were loaded onto the Trace GC 2000 instrument which is equipped with a quadrupole mass spectrometry detector (Trace DSQ) functioning through electron impact (EI) and operating at 70eV. Helium was used as a carrier gas, heated to 280°C and passed through the column at a rate of 1mL per minute and 113 psi. The column operating temperature started at 50°C which was held for two and a half minutes. The first ramp was to 70°C which was then held for seven and a half minutes and the second ramp was to 310°C, which climbed at a rate of 5°C per minute, holding each pause for six minutes. The column used for all experiments was the Rtx-5MS with IntegraGuard (Restek). It was a non-polar column consisting of 5% diphenyl and 95% dimethyl polysiloxane. All samples were run using the following specifications: Excalibur method Fiehn_left_ssL2129; Excalibur sequence ssp1; Gerstel method ssinslaw; and Gerstel sequence dilute_02.

Analysis of intracellular GC-MS data.

Raw data files generated by the GC-MS instrument and saved into the Excalibur program were opened in the Automated Mass Spectral Deconvolution and Identification System (AMDIS) (Stein, 1999), where the files were deconvoluted using the settings described below in Table 2-3. Retention times for the fatty acids standards were identified and the response factor for ribitol was recorded for each replicate. Response factors are derived from the area under the peaks on the gas chromatogram and are proportional to the concentration of the compound within the sample.

AMDIS files were then imported by the GC-MS Data Analysis Software Program (GASP) (Nuin *et al,* registered) where retention times (RT) are converted into retention indexes (RI) to correct for variability in column performance and where relative response factors for all peaks in the chromatogram are calculated by dividing the response factor of the peak by the response factor for ribitol. For conversion to GASP, a signal to noise ratio of 20 was used in all experiments. Peaks are aligned by RI in a spreadsheet. Two checks are subsequently performed on the aligned files, first a check using RT and RI values and then a check using the mass spectra.

Next, using GASP, a statistical analysis of the data using ANOVA (ANalysis Of VAriance) was performed and average values for set of mutants or treatment type were calculated. The files that were generated in GASP were then imported by Excel and sorted according to "P"-value where values less than 0.05 were denoted as significantly different. Significant compounds were identified by retention time using the NIST

(National Institute of Standards and Technology) library where mass spectra are compared to databases at McMaster University and the Max Planck Institute. The GASP program occasionally has difficulties aligning according to retention time index in ascending order. When this problem arose, it was necessary to open the file in the Microsoft NotePad program and to manually align the RT values in proper numerical order.

Table 2-3: Settings used for deconvolution inAMDIS of all samples run on the GCMS.

Instrument settings:	
Low m/z	50
High m/z	700
Threshold	High
Instrument Type	Quadrupole
Scan direction	Low to High
Data file format	Xcalibur Raw file
Ions	50 000
Deconvolution settings:	
Component width	10
Adjacent peak subtraction	one
Resolution	Low
Sensitivity	Medium
Shape Requirements	Low

Growth of cultures and media composition for study of extracellular metabolites

Overnight cultures were inoculated as previously described (growth conditions for metabolite analysis). The following day, the replicates were subcultured into 250mL flasks containing 50mL of M9 media with two carbon sources: 7.5mM glucose and 7.5mM of succinate or arabinose and grown for approximately 16 hours or until an OD_{600nm} of around 0.5-0.6 was reached. Next, the cultures were spun down at 3730g (SX4250 rotor) and 4°C for 20 minutes, the supernatants were removed and the pellet was resuspended in 30mL of 0.85% NaCl and spun down once more. After the second centrifugation step, the supernatant was removed and the cells were resuspended in 50mL of a minimal media containing low concentrations of phosphate and carbon sources (20mM NH_4Cl, 2mM $MgSO_4$, 1.2mM $CaCl_2$, 100mM NaCl, 0.5μg/mL D-biotin, 0.01μg/mL $CoCl_2$, 2.5mM KH_2PO_4 and 5mM of the desired carbon source). The replicates of cell cultures were incubated, shaking, for one to three hours depending on the nature of the experiment. Following incubation, the cells were pelleted as previously described. After centrifugation, 1mL of each of the supernatants was removed, flash frozen in liquid nitrogen and stored at -80°C until needed. The cells were prepared for freezing as previously described.

Preparation of extracellular metabolites for derivatization

Prior to derivatization, large molecular weight compounds (>10,000Da) needed to be removed from the samples to prevent damage from occurring to the GC-MS instrumentation. This was accomplished by passing the supernatants through Amicon Ultra Millipore columns which have a molecular weight cut-off of 10,000Da. First, the columns were washed three times by adding 4mL of distilled water to the filter and centrifuging for 30 minutes at 1204g (SX4250 rotor) and 4°C. Next, the supernatants were thawed, ribitol was added to a final concentration of approximately 3ng/mL (50µL of a 83ng/µL solution) and the supernatants were vortexed. The supernatants were then applied to the columns and spun for 20 minutes at 1204g (SX4250 rotor) and 4°C. Next, 200µL of each sample was transferred into a V-Vial and lyophilized overnight on the Labconco Freezone Plus 6. The remaining 500-800µL of supernatant was transferred into a 1.5mL microfuge tube, flash frozen and stored at -80°C until needed.

Derivatization and GC-MS of extracellular metabolites

Supernatant samples were derivatized and run on the GC-MS instrumentation according to the standard protocols as previously described.

Analysis of extracellular GC-MS data

As before, the raw data files generated by the GC-MS instrument and saved into the Excalibur program were opened in AMDIS, where the files were deconvoluted using the settings described above in Table 2-3. The response factor for ribitol was recorded for each replicate. Since the supernatant samples did not contain many detectable metabolites, the GASP program was omitted from the data analysis process and instead, the chromatograms were analyzed manually by finding the retention time of the compound in question and then dividing the RF value by the RF value for ribitol. This calculation was repeated for all the replicates of each mutant/condition and the averages and standard deviations were calculated using Excel.

Quantification of malate and fumarate in both intracellular and extracellular regions

50mM stocks of malate and fumarate were made, divided into 1mL aliquots and stored at -80°C until needed. Samples of intracellular and extracellular metabolites were prepared and derivatized as described above. 1μL of each 5mM, 10mM and 20mM malate and fumarate was derivatized (in triplicate) according to the previously outlined protocol and run on the GC-MS. Note that 50μL of a 83ng/μL dilution of ribitol was added to the malate and fumarate samples as an internal standard. All GC-MS conditions were consistent with those previously described. A standard curve was constructed from the relative response factors of the malate and fumarate at each concentration. The concentrations of malate and fumarate in the samples were calculated using the equation of the line derived from the standard curve (Appendix).

Co-injection to verify compound identity

For co-injection, 5mg/mL stock solutions were made for the compounds in question. Following lyophilization, but prior to derivatization, 1-3μL of the 5mg/mL solution was added to the samples and was dried down with the N-evap. The samples were then derivatized according to standard protocol. Control samples containing only the compound in question or simply the intracellular or extracellular mixture (blank sample) were also run for comparison. RRF values were calculated manually and compared between the blank samples and the samples which had the compounds added.

β-galactosidase and β-glucoronidase enzyme assays

Strains for *lacZ* and *gusA* gene expression analysis were inoculated into 600μL of LBmc in 96-well deep well plates (ABgene) and grown at 30°C while shaking for two days to ensure adequate growth. Following two days of incubation, the strains were subcultured into specific minimal media by pipetting 6μL of dense culture into 600μL of the specific media in 96-well deep well plates. These cultures were then incubated for 1-2 days, shaking at 30°C. After the appropriate incubation period, 20μL of culture for each strain and condition was added to 96-well microtiter plates, this was used to conduct the GusA or Lac*Z* assays. 100μL of culture was transferred to separate microtiter plates to record OD_{600nm}. To the 20μL of culture, 80μL of the appropriate buffer (see below) was added. The reactions were timed and allowed to proceed until there was a change in

colour or for 30 minutes and then were stopped by the addition of 100μL of 1M Na_2CO_3 (Stop buffer). Optical densities were then measured for the LacZ and GusA assays at 420nm and 405nm respectively.

The specific activities for each strain under each condition were calculated using the following formula: $(1000xA_{420 \text{ or } 405})/(\text{volume of culture added (mL)} \times \text{time (minutes)} \times OD_{600nm})$. The LacZ (ONPG) buffer was made from Z buffer (0.06M Na_2HPO_4 * $7H_20$, 0.04M NaH_2PO_4, 0.01M KCl and 0.01M $MgSO_4$), ONPG (7.2 mg/mL) (Sigma), 0.0125% SDS (10% stock) and 37.8mM β-mercaptoethanol (14M stock). The GusA buffer consisted of 50mM sodium phosphate buffer (1M stock, pH 7.0), 50mM DTT, 1mM EDTA, 0.0125% SDS and 438μg/mL PNPG. Note that information regarding the cloning of fusions to central carbon metabolism genes is contained in Chapter 6.

CHAPTER 3

Results: metabolite profiling

Analysis of intracellular polar metabolites

The two malic enzyme mutants of *S. meliloti*, *dme* and *tme*, display markedly different symbiotic phenotypes; *S. meliloti dme* mutants are deficient in nitrogen fixation whereas *S. meliloti tme* mutants are not (Driscoll and Finan, 1993; Driscoll and Finan, 1996). In order to assess these differences, an analysis of the intracellular polar metabolites of the two mutants was conducted. Three cultures of each *Sinorhizobium* strain, Rm1021, RmG455 (*dme3*::Tn5), RmG995 (*tme4*Ω*Sp)* and RmG994 (*dme3*::Tn5, *tme4*Ω*Sp*) were grown in M9-succinate to an OD_{600nm} of 0.8, their polar metabolites were extracted and derivatized as described in Materials and Methods and the resulting profiles were compared. Metabolites that were significantly different (where $P \leq 0.05$) are shown in Figures 3-1a, b and c. For simplicity, RmG455, RmG995 and RmG994 will be referred to as DME, TME and double, respectively, throughout this document with respect to graphical representation.

Figure 3-1a: Metabolic profile derived from malic enzyme mutants grown in M9-succinate (OD_{600nm} 0.8)

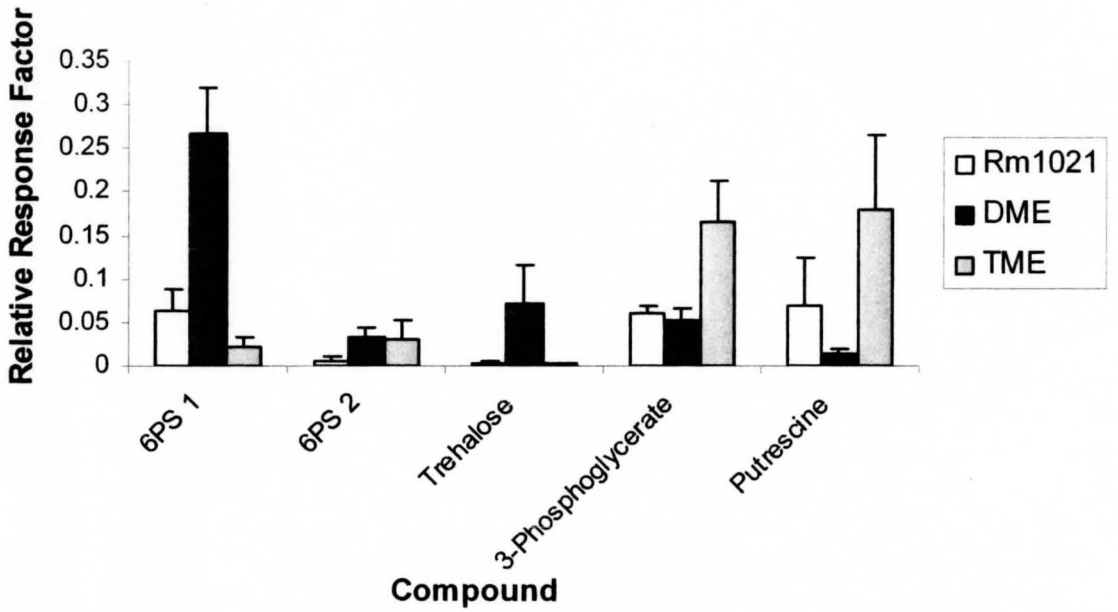


Figure 3-1a: Relative response factors (RRF) for compounds significantly differentially expressed between extracts of Rm1021, *dme* mutants and *tme* mutant strains. Significant compounds were those possessing P values of <0.05 in ANOVA. Note that the most prominent compounds include 6-phosphate sugars (6PS), trehalose and putrescine. Error bars were calculated using standard deviation where N=3.

Figure 3-1b: Gas chromatograms derived from malic enzyme mutants grown in M9-succinate (OD_{600nm} 0.8)

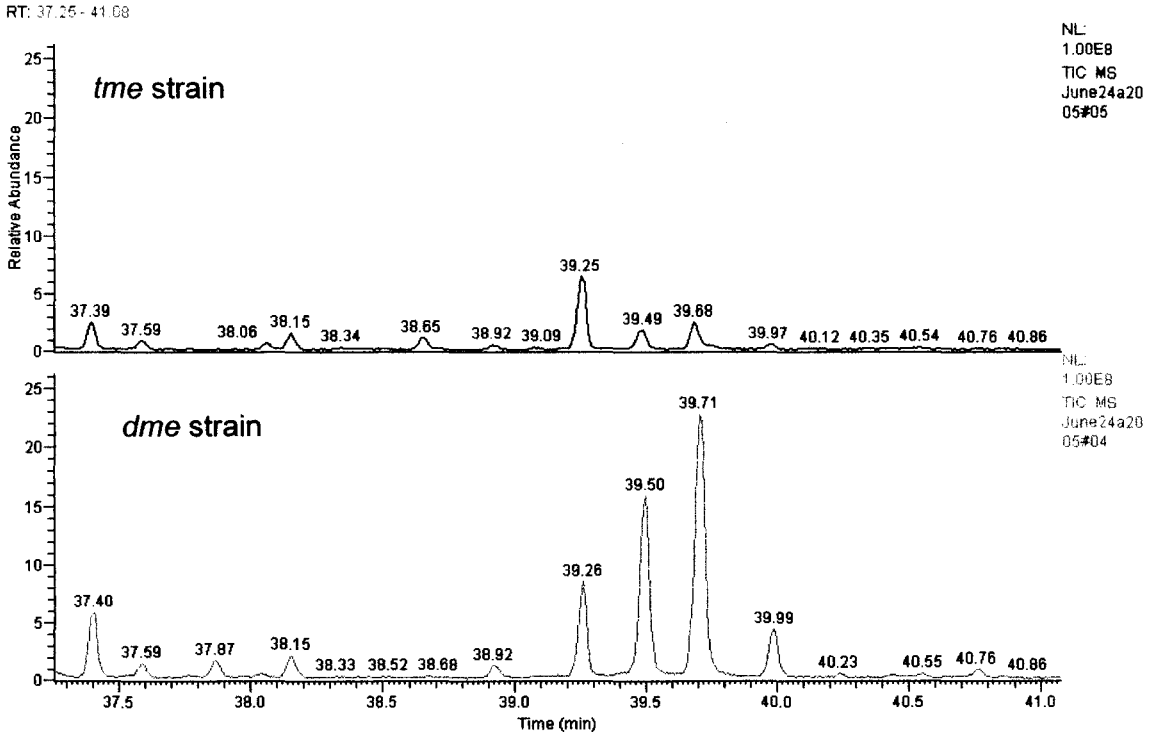


Figure 3-1b: Gas chromatograms for t*me* and *dme* mutant strains, top and bottom respectively, grown in M9-succinate. Note the 6-phosphate sugars at retention times 39.5, 39.71 and 39.99. Images were taken from the Xcalibur program's Qual Browser function.

Figure 3-1c: Gas chromatograms derived from malic enzyme mutants grown in M9-succinate (OD_{600nm} 0.8)

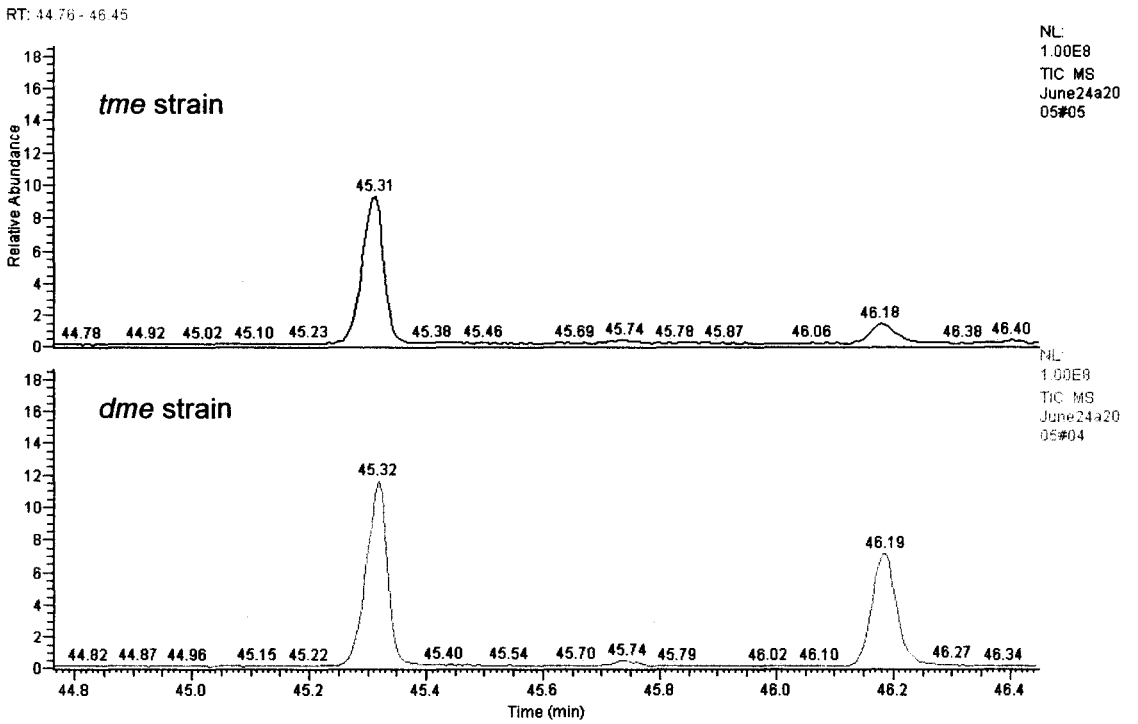


Figure 3-1b: Gas chromatograms for *tme* and *dme* mutant strains, top and bottom respectively, grown in M9-succinate. Note trehalose at retention time 46.19. Images were taken from the Xcalibur program's Qual Browser function. (The peak at RT 45.32 is the C23 fatty acid standard).

When cultures were grown in M9-succinate to an OD_{600nm} of 0.8, it was found that 6-phosphate sugars, likely fructose-6-phophate (6PS1), mannose-6-phosphate (6PS2) and glucose-6-phosphate (6PS3, not shown) had greater RRF values in metabolic profiles of the *dme* mutant strain compared to profiles of both the *tme* mutant strain and wild type Rm1021. As 6PSs play roles in both the glycolytic and gluconeogenic pathways, their presence suggests a possible disturbance in the central carbon metabolism pathways of the *dme* mutant strain. In addition, trehalose was present at much higher levels in the

polar extracts of *dme* mutant cells when compared to the polar extracts of Rm1021 and *tme* mutant cells. Trehalose is a disaccharide and a known osmoprotectant in *S. meliloti* (Gouffi *et al,* 1999) and its presence on the metabolic profiles of the *dme* mutant strain suggests osmotic stress. Moreover, putrescine, a common polyamine in rhizobia (Fujihara and Harada 1989), was found to be greatly reduced in *dme* mutant cell extracts. The biological role of putrescine is currently unknown, but it is believed to be a form of nitrogen storage and notably, decreases in putrescine concentrations in *Rhizobium* have been correlated with decreases in nitrogenase activity (Lahiri *et al* 2004). Lastly, we saw that 3-phosphoglycerate (3PG), which is also part of the glycolytic and gluconeogenic pathways, was present in greater amounts in *tme* cell extracts compared with those of wild type Rm1021 and the *dme* mutant strain. Possible reasons for the accumulation of the 6PSs, trehalose and putrescine will be outlined in the discussion section.

Note that metabolite profiles for the double mutant are not shown for this experiment as only two replicates were tested, making the data difficult to align. It is however important to mention that the two double mutant replicates were similar to those of the *dme* mutant strain replicates as the extracts contained greater amounts of 6PSs and trehalose when compared to the wild type and *tme* mutant strain extracts.

Observed metabolic phenotypes are not growth phase dependent

The presence of trehalose on the metabolic profiles of the *dme* and double mutant strains suggested that these strains may be experiencing a form of osmotic stress and since trehalose concentrations increase with increasing optical density (Miller and Wood, 1996), cultures were grown to an OD_{600nm} of 0.2 and 0.4 in M9-succinate to determine whether the production of these compounds was growth phase dependent (Figure 3-2). Only the data generated from the replicates grown to an OD_{600nm} of 0.2 is included in this document to avoid redundancy. Trehalose is shown on a different graph due to the magnitude of the RRF values.

Figure 3-2: Metabolic profile derived from malic enzyme mutants grown in M9-succinate (OD_{600nm} 0.2)

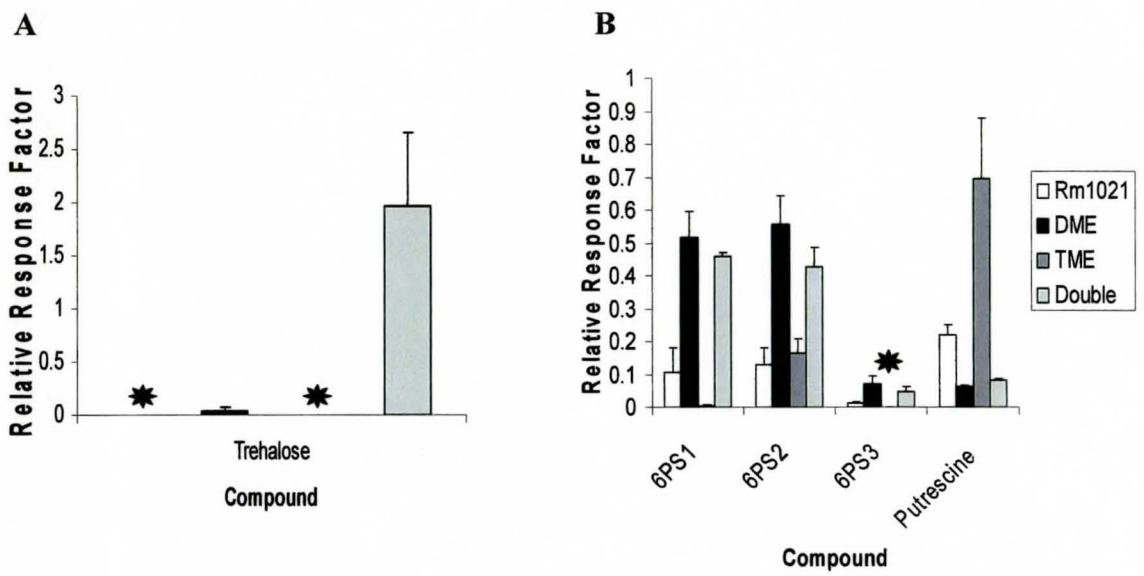


Figure 3-2: A) Relative response factors for trehalose in cultures grown to an OD_{600nm} of 0.2 in M9-succinate. * Trehalose production is quite high in the double mutant also appears in the *dme* mutant strain but is not detected on chromatograms derived from the Rm1021 and the *tme* mutant strains. B) Relative response factors of 6-phosphate sugars and putrescine. Note that 6-phosphate sugars are again produced to higher levels in the DME mutant and the double mutant. Also that RRF values for putrescine are much greater in the *tme* mutant and Rm1021 compared to the *dme* and double mutant strains. * 6PS3 was not detected on the chromatograms derived from the *tme* mutant strain. Error bars were calculated using standard deviation where N=3.

As was found in the cultures grown to an OD_{600nm} of 0.8, the RRF values for the 6-phosphate sugars and trehalose were elevated in the *dme* and double mutant extracts compared to Rm1021 and *tme* mutant extracts. Moreover, the RRF values for putrescine were greatly reduced in the *dme* and double mutant profiles compared to the *tme* mutant and Rm1021 profiles. Hence, the production of these compounds does not appear to depend on the optical density of the culture.

Metabolic differences are not allele-specific

To rule out the possibility that the metabolic differences observed between the *dme* and *tme* mutant strains could be dependent on the individual mutant allele, or on the nature of the insertion creating the knock-out, a second set of alleles for both the *dme* and *tme* mutant strains was grown in M9-succinate to an OD_{600nm} between 0.5 and 0.6 (Figure 3-3). The strains tested were RmG454 (Rm1021 *dme2*::Tn5) and RmH215 (Rm1021 *tme2*::Tn5), both of which have Tn5 insertions whereas *dme3* and *tme4* had Tn5 and ΩSp insertions respectively.

Figure 3-3: Metabolic profile derived from malic enzyme mutants, *dme2* and *tme2* grown in M9-succinate

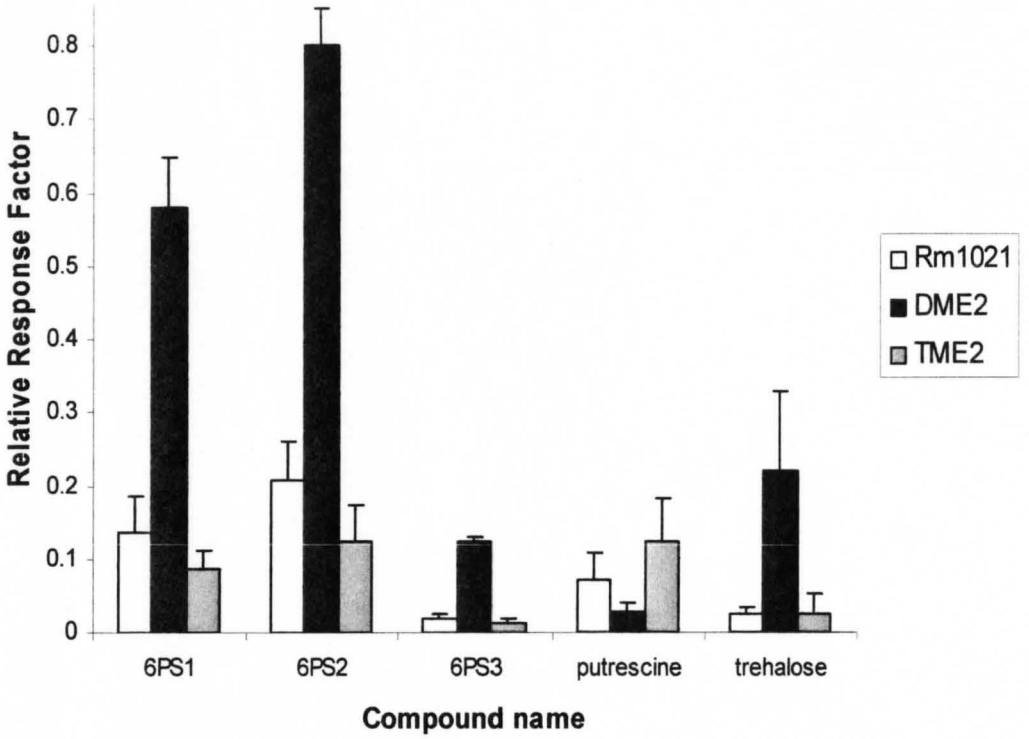


Figure 3-3: Relative response factors for comparison of *dme* and *tme* mutant strains (*dme2::Tn5* and *tme2::Tn5)* of alleles differing from those previously tested. Note the elevated relative response factors for the 6-phoshate sugars (6PS) and trehalose in the *dme* mutant strain in addition to the lower RRF for putrescine compared to Rm1021and the TME mutant. Error bars were calculated using standard deviation where N=3.

The results shown above in Figure 3-3 are consistent with those previously observed on the profiles of *dme*3 and *tme*4 which consisted of elevated RRF values for the 6PSs and trehalose and reduced RRF values for putrescine displayed the *dme* mutant profiles compared to the *tme* mutant and Rm1021 profiles (Figure 3-1). These results support the hypothesis that the observed metabolic phenotypes are not dependent on the allele of the mutation or on the nature of the insertion that has caused the mutation.

Observed metabolic phenotypes are dependent on the presence of succinate in the media

All of the strains used in this study (including the wild type strain, Rm1021) grow more slowly in M9-succinate and M9-glycerol compared to M9-glucose and LB (data not shown). One reason for the difference in growth rate is that the metabolism of succinate or glycerol as a sole carbon source necessitates the induction of gluconeogenesis whereas the use of glucose does not. In order to examine whether the metabolic differences observed between the *dme* and *tme* mutants were dependent on the structure of the carbon source or upon growth rate, cultures were grown in M9-glucose and M9-glycerol to OD_{600nm} values of 0.5-0.6 (Figures 3-4 and 3-5).

Figure 3-4: Metabolic profile derived from malic enzyme mutants grown in M9-glucose

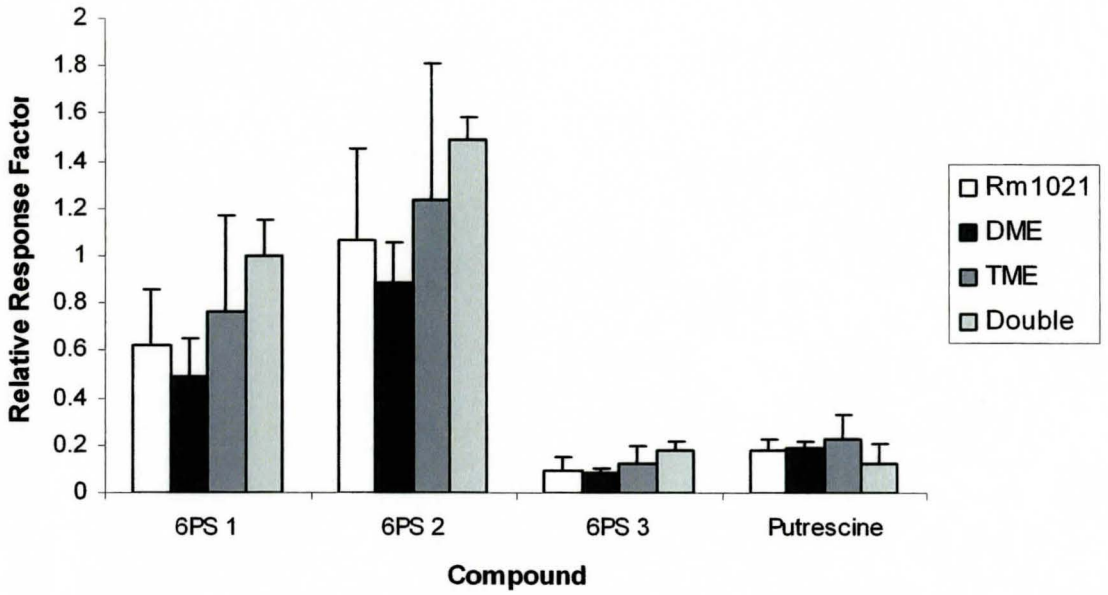


Figure 3-4: Relative response factors for 6-phosphate sugars and putrescine for polar extracts of cultures grown to an OD_{600nm} between 0.5 and 0.6 in M9-glucose. There does not appear to be any significant differences between the relative response factors for the different strains for each of these compounds. Trehalose RRFs were negligible in all extracts when the strains were grown in M9-glucose. Error bars were calculated using standard deviation where N=3.

Figure 3-5: Metabolic profile derived from malic enzyme mutants grown in M9-glycerol

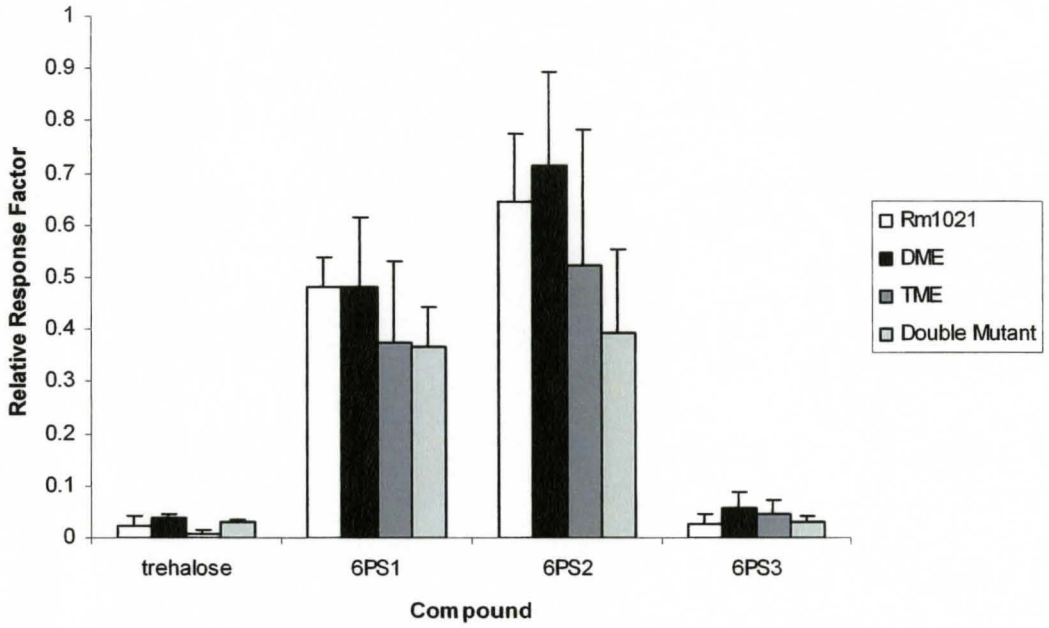


Figure 3-5: Relative response factor versus compound for cultures grown to an OD_{600nm} between 0.5 and 0.6 in M9-glycerol. RRF values for the 6-phosphate sugars and trehalose are comparable for each of the strains. Putrescine RRFs varied extensively between replicates of each strain and thus were not included in this graph. Error bars were calculated using standard deviation where N=3.

As seen in Figures 3-4 and 3-5, chromatograms of extracts of cultures grown in glucose and glycerol showed that there was no substantial difference between the four tested strains in the RRF values for the 6-phosphate sugars or putrescine. In addition, trehalose peaks were undetectable on the chromatograms for all strains grown in glucose and were very small for cultures grown in glycerol. These results confirm the hypothesis that the observed metabolic phenotypes are dependent on the presence of succinate in the media and that they are not dependent on the growth rate of the culture.

Succinate and malate pulse experiments

To account for the increased production of 6-phosphate sugars in the *dme* mutant strain, we hypothesized that 6-phosphate sugar production may be a downstream effect of a disturbance in central carbon metabolism. Two sets of "pulse" experiments were conducted in an attempt to visualize more immediate effects of the disruption of malate metabolism in the *dme* mutant strain. In both of these experiments, LBmc cultures were subcultured into M9-glucose and incubated with shaking overnight until an optical density between 0.5 and 0.6 was reached. For the first experiment, the cultures were centrifuged, washed in 0.85% NaCl and resuspended in M9-succinate. For the second experiment, the cultures were washed in 0.85% NaCl and resuspended in 0.85% NaCl with 5mM malate. The second pulse experiment was conducted in 0.85% NaCl with 5mM malate to allow for analysis of the supernatant as M9 salts contain a high concentration of phosphates and so cannot be adequately analyzed by the GC-MS in the lab.

Interestingly the results yielded from these experiments (Figure 3-6) were strikingly similar to those seen following a full incubation in M9-succinate media.

Figure 3-6: Metabolic profile derived from malic enzyme mutants grown in M9-glucose and pulsed for two hours in M9-succinate

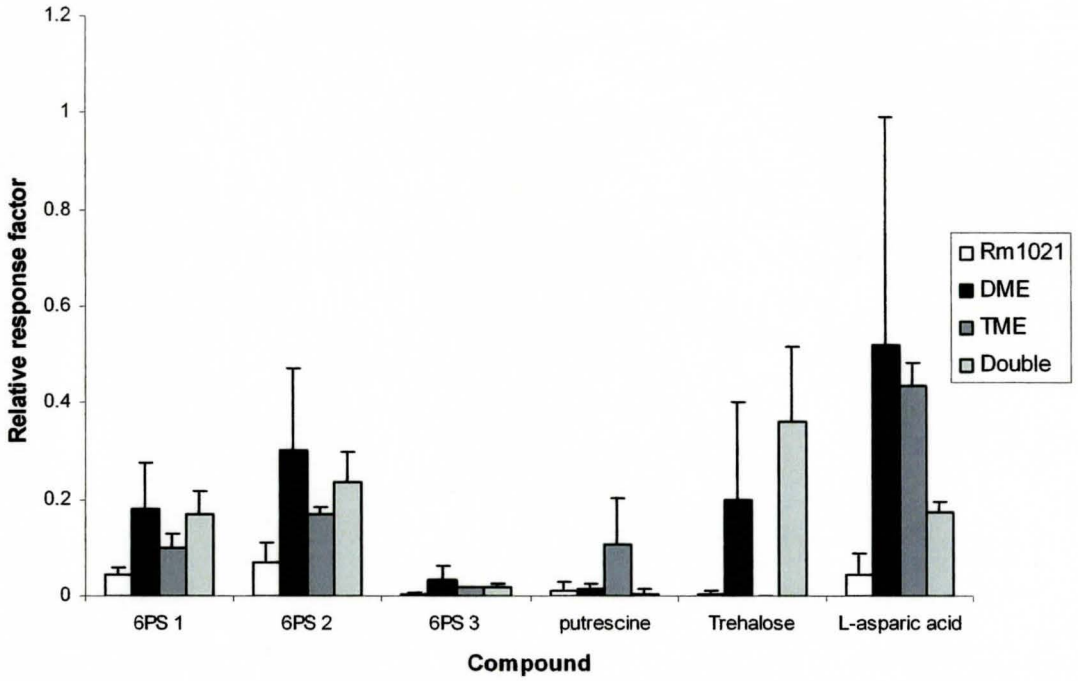


Figure 3-6: Relative response factors for polar extracts of strains grown in M9-glucose overnight and then pulsed for two hours with M9-succinate. 6-phosphate sugars are elevated in the appropriate strains as is trehalose and putrescine is reduced. In addition, aspartic acid RRF values are much greater in the *dme* and *tme* mutants compared to wild type. Error bars were calculated using standard deviation where N=3.

Results from the "succinate-pulse" experiment show similar phenotypes to the previous experiments (6PSs, trehalose and putrescine) with the exception of the increased production of aspartic acid which was observed in all malic enzyme mutant strains. This increased production of aspartic acid may be easily explained as aspartic acid is synthesized from oxaloacetate (OAA) and OAA is synthesized from malate (Figure 1-1). However, the RRF value for the double mutant is not consistent with this hypothesis.

A second pulse experiment was performed where the cultures were grown overnight in M9-glucose and centrifuged, washed in 0.85% NaCl and finally resuspended in 0.85% NaCl with 5mM malate for two hours (Figure 3-7). Malate was used as a carbon source to observe more immediate effects of the *dme* mutation on the metabolic profiles. Only Rm1021 and the *dme* mutant strain were compared for this experiment.

Figure 3-7: Metabolic profile derived from malic enzyme mutants grown in M9-glucose and pulsed for two hours in 5mM malate in 0.85% NaCl

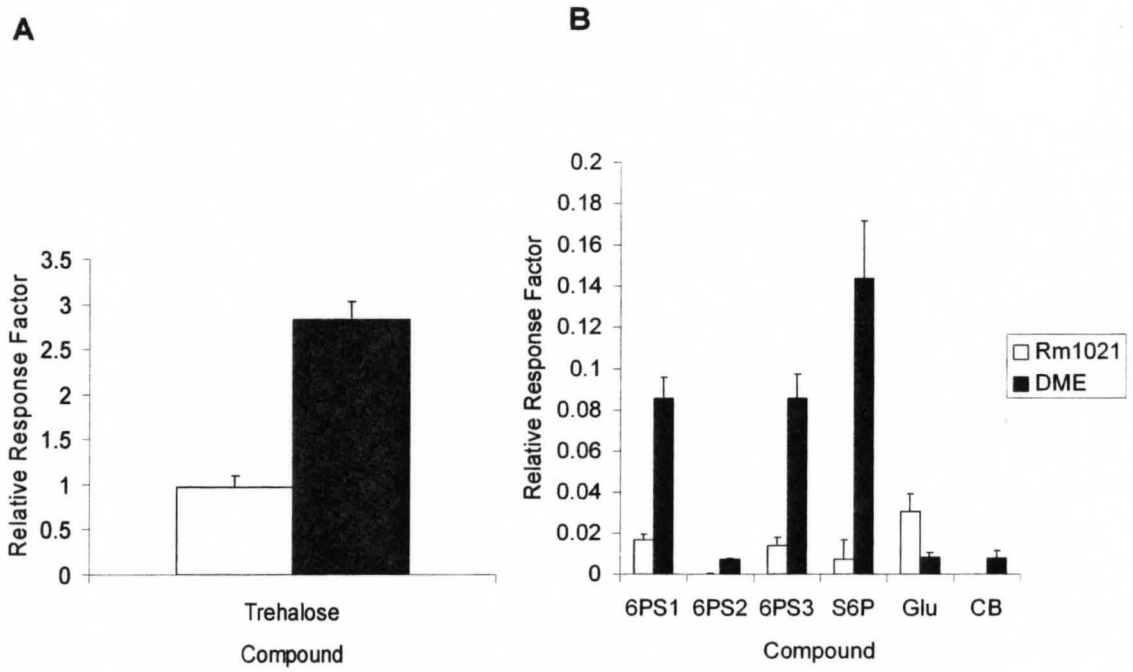


Figure 3-7: A) Relative response factors for trehalose from polar metabolite extracts of Rm1021 and the DME mutant grown overnight in M9-glucose and then pulsed for two hours the following day in 0.85% saline with 5mM malate. Note the 3-fold greater RRF value for trehalose in the *dme* mutant strain compared to Rm1021. B) Relative response factors for 6PSs, sucrose-6-phosphate (S6P) and a cellobiose-like compound (CB) all of which are of greater concentrations in the *dme* mutant profiles compared to Rm1021 profiles. Glutamic acid (Glu) has a greater RRF value in the Rm1021 profiles compared to the *dme* mutant profiles. Error bars were calculated using standard deviation where N=5.

As in the previous experiments, RRF values for 6-phosphate sugars and trehalose on the profiles of the *dme* mutant strain were greater relative to those of Rm1021. In addition, we see that sucrose-6-phosphate and a cellobiose-like compound are present at elevated concentrations in the *dme* mutant extracts relative to those of the wild type strain. Sucrose-6-phosphate may be used to make UDP-glucose which is a precursor of trehalose and as previously noted, trehalose accumulates in the *dme* mutant strain. Both sucrose and cellobiose are also a known osmoprotectants in *S. meliloti* (Gouffi *et al*, 1999) which lends support to the hypothesis that the *dme* mutant strain may be experiencing osmotic stress. Glutamic acid is slightly elevated in the wild type strain, Rm1021, which seems reasonable as glutamate is synthesized from α-ketoglutarate (Figure 1-1).

Analysis of supernatants

Supernatants derived from the cultures pulsed in 0.85% NaCl with 5mM malate were harvested, passed through Amicon filters, derivatized and run on the GC-MS as described in materials and methods. Results from this experiment are depicted below in Figure 3-8.

Figure 3-8: Extracellular metabolic profile derived from malic enzyme mutants grown in M9-glucose and pulsed for two hours in 5mM malate in 0.85% NaCl

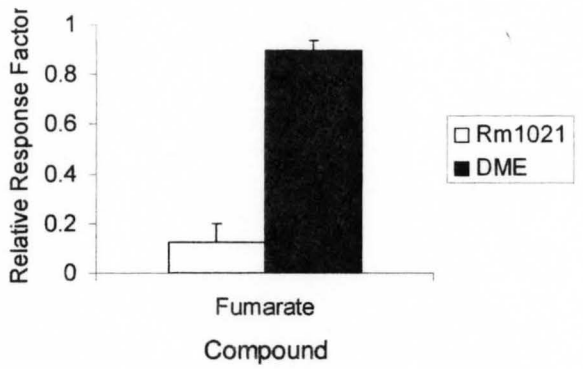


Figure 3-8: Relative response factors for fumarate derived from the supernatants of Rm1021 and the DME mutant grown overnight in M9-glucose and shocked for two hours the following day in 0.85% saline with 5mM malate. Note the 10-fold difference in the RRF value for fumarate between the *dme* mutant and Rm1021 strains. Error bars were calculated using standard deviation where N=5.

Figure 3-8 above shows that the *dme* mutant strain, when provided with malate as a sole carbon source, excretes approximately a ten-fold greater quantity of fumarate into their extracellular environment compared to the wild type strain. Fumarate was the only compound that was visible on the chromatogram that was not visible on chromatograms of the original, unused 5mM Malate/0.85% NaCl media.

The experiment described above provided us with a glimpse of the metabolic profile of the supernatant of a malic enzyme mutant but the experimental set-up was inadequate as cells would not survive long in an environment lacking all nutrients aside from carbon source. Therefore, it was necessary to examine the supernatant from cultures grown in minimal media. However, due to the high phosphate content of both rich and minimal growth media, supernatant studies are difficult to execute. Organic phosphate molecules are derivatized by MSTFA and these trimethylsilylized phosphate molecules

will produce very large peaks on the gas chromatogram, occluding surrounding peaks and causing problems with detection. The concentration of the carbon sources present in the minimal media will also result in large peaks.

To avoid these issues, cultures were grown overnight in minimal media containing 7.5mM glucose and 7.5mM succinate as carbon sources (glucose was added to ensure adequate growth; succinate was added to permit adequate time for induction of the *dct* transport system which is required for uptake of succinate, malate and fumarate), the cultures were washed the following day and transferred into a media with 2.5mM phosphate (pH 7.0) and 5mM succinate for a period of two hours. The supernatants were then passed through filters which removed compounds with molecular weights greater than 10000Da, dervatized according to standard protocol and run on the GCMS. The results from this experiment are shown below in Figures 3-9a and b.

Figure 3-9a: Extracellular metabolic profile derived from malic enzyme mutants grown overnight in M9-glucose + succinate and pulsed for two hours in 5mM succinate media

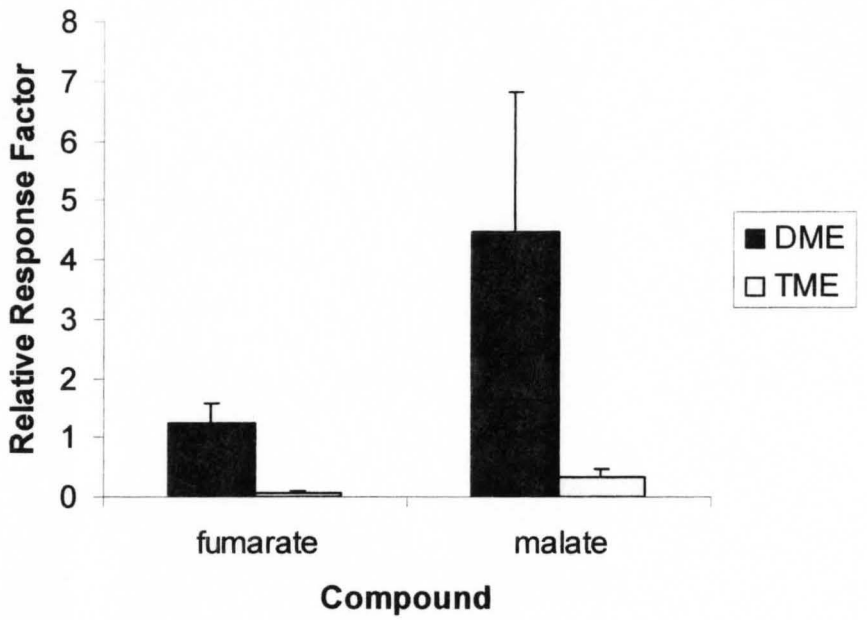


Figure 3-9a: Relative response factors for fumarate and malate derived from supernatants of DME and TME mutants grown overnight in M9-glucose plus succinate (7.5mM each) and then shocked for two hours the following day in minimal media with 2.5mM phosphate and 5mM succinate. Error bars were calculated using standard deviation where N=5.

Figure 3-9b: Gas chromatograms derived from supernatants of malic enzyme mutants grown in 5mM succinate media

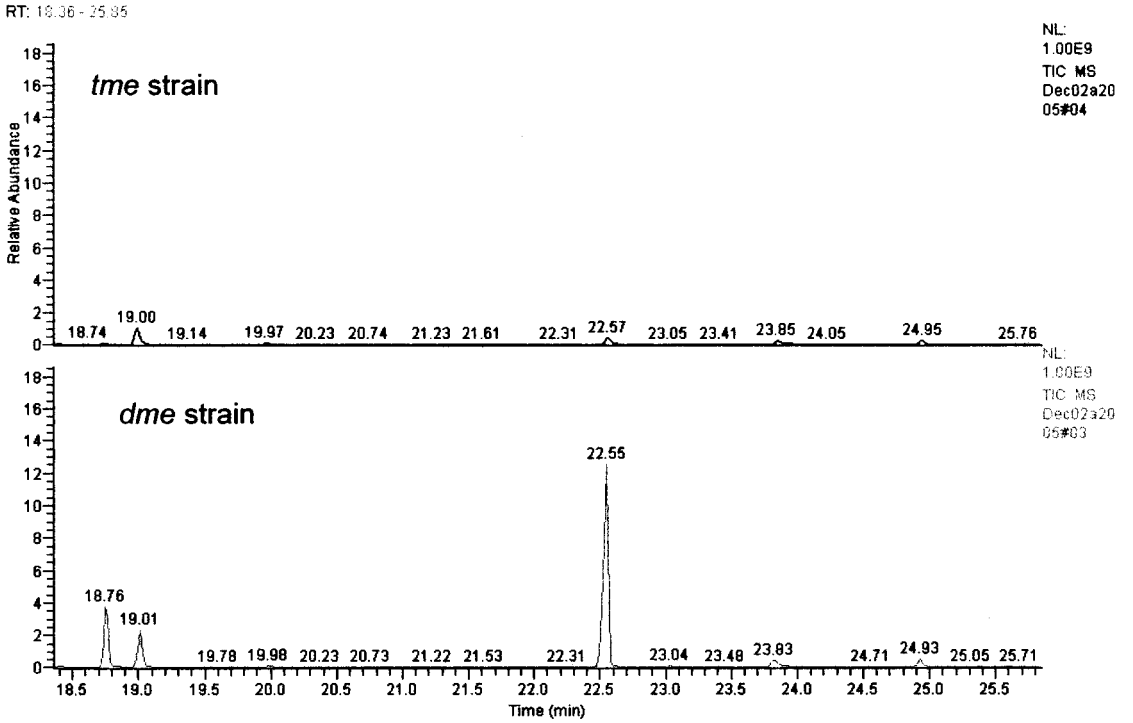


Figure 3-9b: Gas chromatograms for supernatants of *tme* and *dme* mutant strains, top and bottom respectively, grown in 5mM succinate media. Note fumarate and malate at retention times 18.76 and 22.55 respectively. Images were taken from the Xcalibur program's Qual Browser function.

Table 3-1: RRF values for malate and fumarate in the extracellular environment of *dme* and *tme* mutants

Sample	Malate RRF	Fumarate RRF
dme supernatant	4.46 +/- 2.36	1.24 +/- 0.34
tme supernatant	0.346 +/- 0.131	0.068 +/- 0.042

As shown in Figures 3-9 a and b and Table 3-1, malate and fumarate RRF values in the supernatant of the DME mutant were respectively 15 and 19 fold greater than those

of the *tme* mutant strain, thus indicating that malate is not efficiently metabolized in *dme* mutant *S. meliloti*.

Once it was established that the mass spectrometry detector could be switched off during the time that the derivatized phosphate molecules were flowing through the instrument, a second experiment was conducted where supernatants derived from cultures grown in M9-succinate (15mM succinate) for two hours were derivatized and run on the GC-MS. Interestingly, the results from this experiment were identical (data not shown) to those from the experiment where the supernatants were derived from cultures grown in media with 2.5mM phosphate and 5mM succinate (Figure 3-9a) and so the remainder of the supernatant experiments were conducted in media containing low concentrations of phosphate and carbon source.

Accumulation of fumarate and malate over time

A time-course experiment was performed where cultures were grown overnight in M9 media with glucose and succinate and then transferred the following day into media with low concentrations of succinate and phosphate. Three 30mL aliquots were taken from a flask containing 200mL of culture at one, two and three and one half hours post transfer. Both malate and fumarate were found to accumulate over time in the supernatant of the DME mutant strain grown in media containing succinate (Figure 3-10).

Figure 3-10: Accumulation of fumarate and malate in the supernatants of DME and TME mutants over time

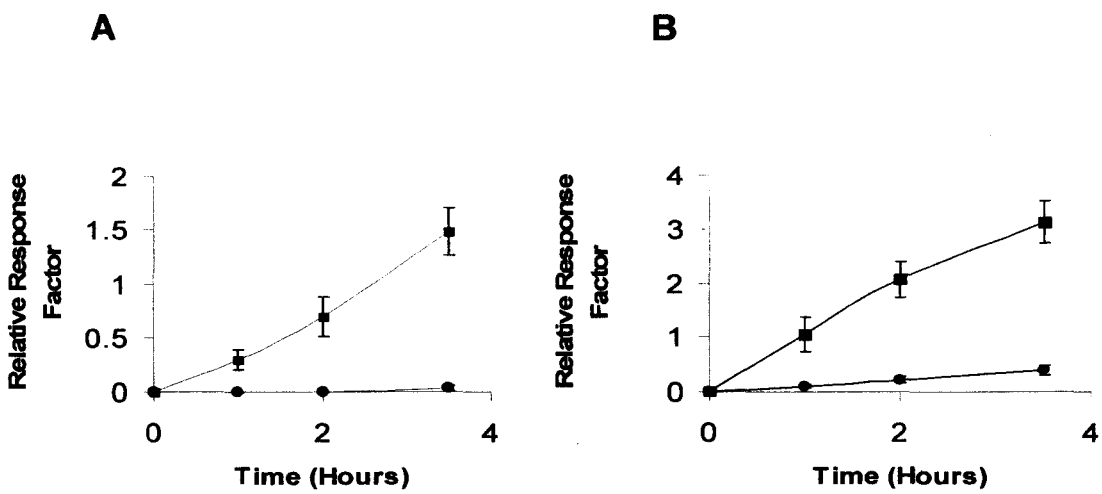


Figure 3-10: A) Relative response factors for fumarate in the supernatants of *dme* and *tme* mutant strains grown overnight in M9-glucose and succinate media then transferred the following morning into media containing 2.5mM phosphate and 5mM succinate. B) Relative response factors for malate in the supernatants of *dme* and *tme* mutant strains. Note that both malate and fumarate concentrations increase in the supernatant of the *dme* mutant strain over time. Samples were taken 1, 2 and 3.5 hours post transfer. Error bars were calculated using standard deviation. For each sample, N=3. ■ DME ● TME

It is important to note that malate and fumarate were the only compounds which appeared on the gas chromatogram of the supernatants of the malic enzyme mutants which were not already present in the chromatogram of a derivatized sample of the original media.

Since malate and fumarate were excreted into the surrounding media, it may be hypothesized that these compounds also accumulate within the cell, however, the levels of neither compound were significantly different between the two malic enzyme mutants (defined by $p<0.05$) for any of the experiments previously conducted (Figure 3-11, below).

Figure 3-11: Intracellular metabolic profile derived from malic enzyme mutants grown overnight in M9-glucose and succinate and pulsed for two hours in 5mM succinate media

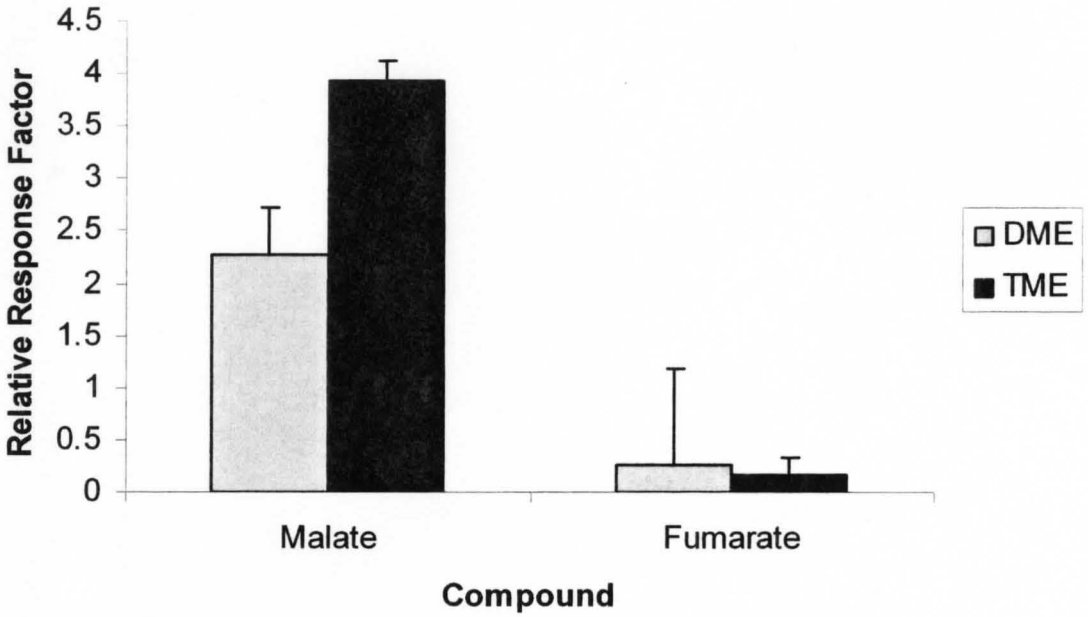


Figure 3-11: Relative response factors for fumarate and malate from polar metabolites contained in the cell extracts of DME and TME mutants grown overnight in M9-glucose plus succinate (7.5mM each) and then shocked for two hours the following day in minimal media with 2.5mM phosphate and 5mM succinate. Error bars were calculated using standard deviation where N=5.

Quantification of malate and fumarate

The molarity of the malate and fumarate which is excreted into the surrounding media by both the *tme* and the *dme* mutant strain was estimated through the construction of standard curves (Appendix) using the GC-MS. The Table 3-2 below shows the concentrations of malate and fumarate both inside and outside the cell.

Table 3-2: Molarity of malate and fumarate in the cells extracts and supernatants of malic enzyme mutants

Sample	Malate Concentration (mM)	Fumarate Concentration (mM)
dme cell extract	0.0187 +/- 0.000907mM	0.0102 +/- 0.00132 mM
tme cell extract	0.0115 +/- 0.00416 mM	0.00416 +/- 0.0007 mM
dme supernatant	4.32 +/- 0.6 mM	2.33 +/- 0.184 mM
tme supernatant	0.247 +/- 0.0303 mM	0.146 +/- 0.0071 mM

NB: To obtain the cell extracts, cells were grown in M9-succinate until a final OD_{600nm} of 0.5-0.6 was reached. To obtain the supernatants, cultures were grown overnight in M9-glucose and succinate, transferred the following day into media with 2.5mM phosphate and 5mM succinate for two hours.

The data in table 3-2 shows that while the concentrations of malate and fumarate inside the cell are in the nanomolar range, the *dme* mutant excretes malate and fumarate into the extracellular environment in the millimolar range. It is important to note that the values in Table 3-2 are estimates and hence this experiment will require repetition.

Growth of cultures in M9-glucose and succinate

Interestingly, when *dme* and *tme* mutant strains were grown in M9-glucose and succinate overnight prior to transfer into media containing 2.5mM phosphate and 5mM succinate, a difference in growth rate was observed. As demonstrated in Figure 3-12, the *dme* mutant strain grows at approximately one half the rate of both the *tme* mutant and wild type strains in media containing equal concentrations of glucose and succinate (7.5mM of each).

Figure 3-12: Growth curves of malic enzyme mutants in M9-succinate and in M9-glucose and succinate

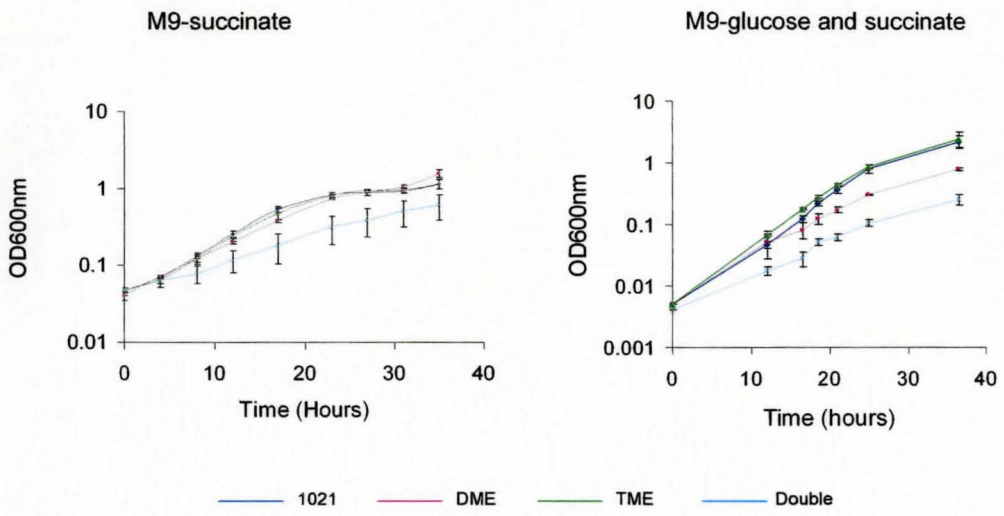


Figure 3-12: Growth curves of Rm1021, DME, TME and double mutants grown in M9-glucose and succinate (left) and in M9-succinate (right). Note the reduced growth rate of the *dme* mutant in M9-glucose and succinate compared to the *tme* mutant and Rm1021. The double mutant has a reduced growth rate in all types of media.

Because of the difference in growth rate between *dme* and *tme* mutant strain cultured in M9-glucose and succinate, it was hypothesized this may be a reflection of metabolic differences. To explore these differences, extracts were taken from cells grown in M9-glucose and succinate which were subsequently derivatized and run on the GC-MS instrument; refer to figure 3-13.

Figure 3-13: Intracellular metabolic profile derived from malic enzyme mutants grown in M9-glucose and succinate

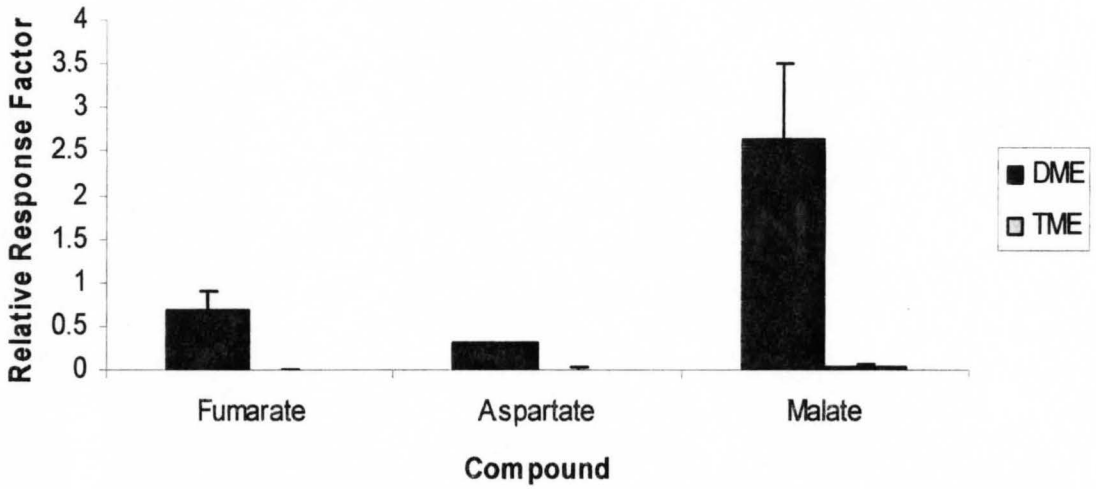


Figure 3-13: Relative response factors for fumarate, asparatate and malate contained in the polar extracts of DME and TME mutant cells grown in M9-glucose and succinate. Note that all three of the compounds shown are of greater concentration in the *dme* mutant strain compared to the *tme* mutant strain: 171, 23 and 78- fold respectively for fumarate, aspartate and malate. Error bars were calculated using standard deviation where N=5.

As shown in figure 3-13, the polar extracts of *dme* mutants grown in M9-glucose and succinate contain 171, 23, and 78-fold greater concentrations of fumarate, aspartate and malate, respectively when compared to those of the *tme* mutant; this was not observed when cells were grown in M9 containing only succinate as a carbon source. These results suggest that the addition of glucose to a media containing succinate may cause an increase in the accumulation of unused malate and fumarate within the cells of the *dme* mutant. Possible reasons for this accumulation will be outlined in the discussion section.

Supernatants of cultures transferred from M9-glucose and succinate into media with 2.5mM of each phosphate, glucose and succinate were derivatized and run on the GC-MS instrument. The metabolic profiles of these samples were identical to those of cells transferred into media with 2.5mM phosphate and 5mM succinate, with approximately 10-fold higher RRF values for malate and fumarate in the *dme* mutant supernatants compared to the *tme* mutant supernatants (data not shown).

Metabolic profiles of cultures grown in M9 glucose and arabinose

Through a series of enzymatic reactions, arabinose may be converted into α-ketoglutarate which is a TCA cycle intermediate located upstream of succinate in the cycle (Figure 1-1). Accordingly, it was hypothesized that succinate and possibly fumarate and malate may be excreted into the surrounding environment of *dme* mutants and hence may also accumulate within the cell if cultures were grown in M9-glucose and arabinose (similar to what was observed when cells were grown in M9-glucose and succinate). However, there was no significant difference in levels of succinate, malate or fumarate between *dme* and *tme* mutant cell extracts from cultures grown in M9-glucose and arabinose (data not shown). In addition, aliquots of these cultures were subsequently transferred into media containing 2.5mM phosphate and 5mM arabinose and the supernatants were derivatized, run on the GC-MS and analyzed. There was no significant difference in the RRF values of malate, fumarate or succinate between the two mutant strains under these conditions.

There were several peaks present in the chromatograms of both *dme* and *tme* cell extracts which were not seen in the previous experiments (not involving arabinose), including uracil and xylulose and polyamine-like compounds. Note that these compounds were identified using only the NIST library program.

Analysis of the effect of malic enzyme mutations on the expression of the genes of central carbon metabolism

The appearance of glucose-6-phosphate and fructose-6-phosphate in *dme* mutant extracts, both of which play roles in the Etner-Douderfoff pathway in *S. meliloti,* lead to the hypothesis that such an accumulation may be resulting from disturbances in the central carbon metabolism of the organism. To test this hypothesis, gene expression of many carbon metabolism genes was determined by analysis using either *lacZ* or *gusA* reporter fusion clones from an *S. meliloti* library (Cowie *et al,* 2006). In addition, reporter fusions to specific genes were constructed in this study. The fusion plasmids (non-replicating) were mated into Rm1021 and the three mutant backgrounds (*dme3*, *tme4* and *dme3tme4*). The resulting strains were then assayed for *gusA* or *lacZ* gene expression depending on the nature of the fusion. Gene expression was tested in four media types- M9-succinate, M9-glucose, M9-glycerol and M9-glucose and succinate.

The expression level of the majority of the central carbon metabolism genes remained consistent between mutant backgrounds despite differences in media type. However, the original round of assays did yield two potentially interesting results. The first was a two-fold increase in enolase (*eno)* gene expression in M9-succinate-grown

tme and double mutants over wild type and the *dme* mutant (Figure 3-14). However, when these strains were re-tested, the two-fold difference was no longer apparent. The second interesting result was a six-fold increase in glyceraldehyde-3-phosphate dehydrogenase (*gap*) gene expression in the double and *tme* mutants compared to Rm1021 (Figure 3-14) and interestingly, this difference in expression was apparent in all media types.

Figure 3-14: β-galactosidase specific activity for a selection of fusions to central carbon metabolism genes

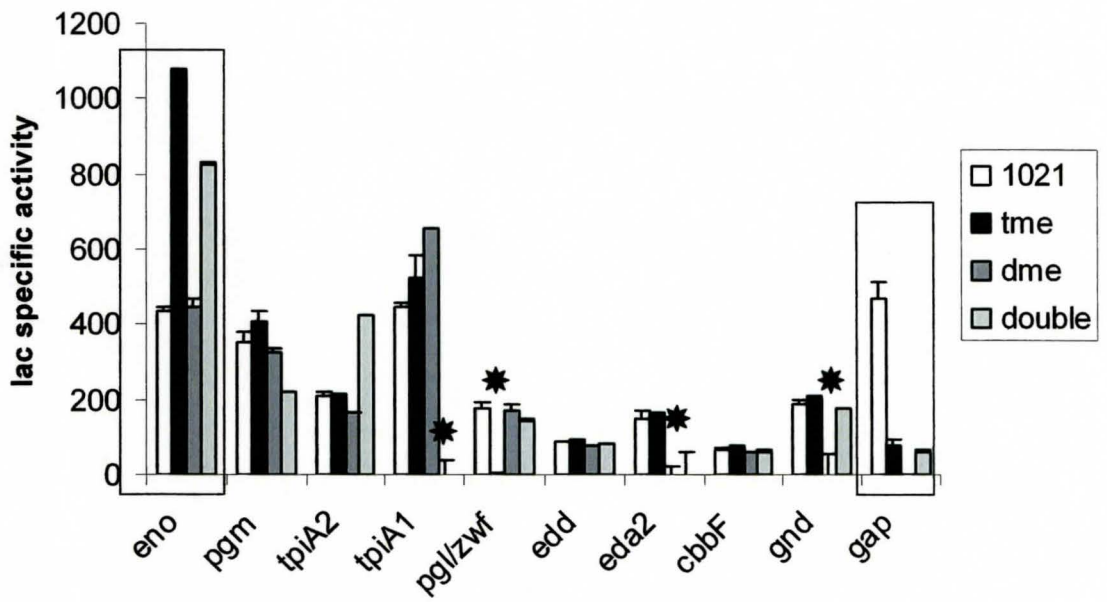


Figure 3-14: Graph of *lacZ* gene expression data for selected carbon metabolism genes where strains were grown in M9-succinate. Assays were conducted in triplicate and the error bars were calculated from the standard deviation between the three replicates. *eno* and *gap* expression profiles are boxed as they show a difference in expression that is greater than two-fold. * Note that some strains appear to have zero activity, this is a result of an inability of the strain in question to grow. Error bars were calculated using standard deviation where N=3.

As shown in Figure 3-14 above, there were a number of strains that showed low *lacZ* gene expression in M9-succinate as they did not grow in any of the media types tested. These included fusions to *tpiA1* and *eda2* in a double mutant background, to *pgl* in a *tme* mutant background and to *eda2* in a *dme* background. This inability to grow may be indicative of a novel phenotype and so should be further investigated.

Upon re-test, fusion to the *gap* gene was found to have an eight-fold increase in the double and *tme* mutants compared to Rm1021 and the *dme* mutant (Figure 3-15).

Figure 3-15: β-galactosidase specific activity for a *gap* fusion in malic enzyme mutant backgrounds

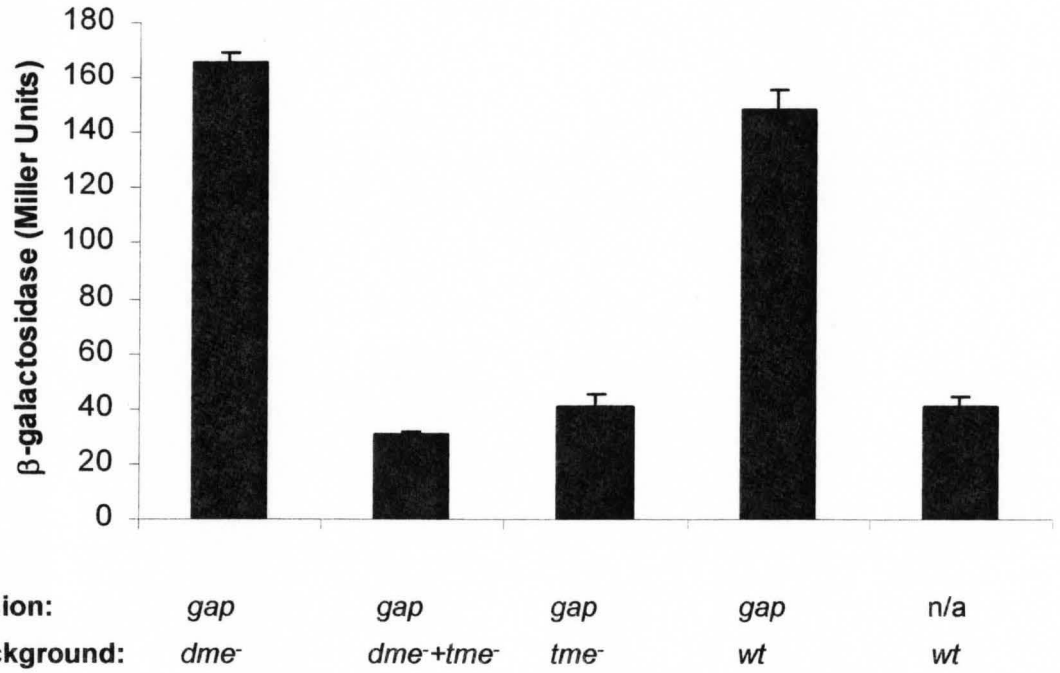


Figure 3-15: Graph of *lacZ* gene expression data for *gap* gene in malic enzyme mutant backgrounds where strain were grown in M9-succinate. Assays were conducted in triplicate and the error bars were calculated from the standard deviation where N=3.

From Figure 3-15 it is clear that the *gap* gene fusion is expressed at decreased levels in the *tme* mutant strains.

Additional data for the experiments of gene fusion expression are included in the appendix.

CHAPTER 4

Discussion: metabolite profiling

Previous work conducted by Driscoll and Finan, (1993, 1996) demonstrated that the symbiotic phenotypes of the two *S. meliloti* malic enzyme mutant strains, *dme* and *tme,* are distinctly different as the *dme* mutant strain is deficient in nitrogen fixation whereas the *tme* mutant strain is not. Despite this, there are no obvious free-living phenotypes which distinguish these two strains. In order to assess these differences, an analysis of the intracellular and extracellular polar metabolomes of the two strains was conducted.

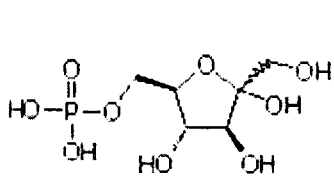
dme strains have increased intracellular concentrations of 6-phosphate sugars and trehalose

When *dme* mutant cultures were grown in M9-succinate, the 6-phosphate sugars (6PS): fructose-6-phosphate, mannose-6-phosphate, glucose-6-phosphate and sucrose-6-phosphate (appears when malate is provided as a sole carbon source for a 2 hour pulse) accumulated intracellularly. As 6PSs (Figure 4-1) play roles in both the glycolytic and gluconeogenic pathways, their presence suggests a disturbance in the central carbon metabolism of the *dme* mutant. In fact, it may be hypothesized that an increased flux is occurring through these pathways (excess malate might be converted to OAA; OAA to phosphoenolpyruvate (PEP); PEP to 2-phosphoglycerate and so on) thus generating an excess of 6-phosphate sugars leading to their accumulation within the cell. However,

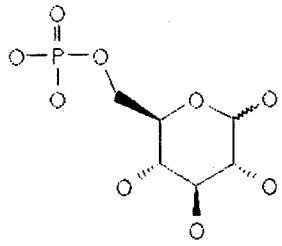
when expression of central carbon metabolism genes was examined by gene fusions to *lacZ* and *gus*A genes, no change in gene expression was detected when *dme* and *tme* strains were grown in M9-succinate and compared; additional information might be obtained from a comparison of these mutant strains using microarray technology.

In addition, it was observed that 3-phosphoglycerate (3PG), which is also part of the gluconeogenic pathway, was present in greater amounts in cell extracts of the *tme* mutant strain compared with those of Rm1021 and the *dme* mutant strains. This observation corresponds to the decreased expression of the *gap* gene in *tme* mutant strains. These results may be considered contradictory to the previous hypothesis which stated that there is overflow through the gluconeogenic pathway in the *dme* mutant strain (which did not show a 3PG accumulation). However, 3PG is also a precursor for 3-phosphate-hydroxypyruvate which is a precursor for phosphoserine. Phosphotidylserine is a lipid that can be used to make membranes. Sanwal (1970) hypothesized that the NADP+-dependent malic enzyme in *E. coli* may have a biosynthetic role and if this is true for the TME enzyme of *S. meliloti,* perhaps the TME enzyme functions to generate NADPH for lipid biosynthesis and thus absence of TME may result in an accumulation of lipid precursors.

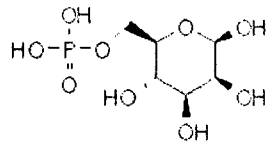
Figure 4-1: 6-phosphate sugars present in intracellular extracts of *dme* mutants



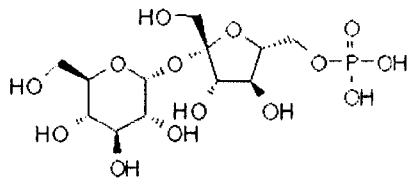
Fructose-6-phosphate



Glucose-6-phosphate



Mannose-6-phosphate



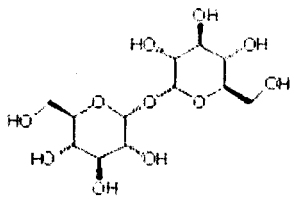
Sucrose-6-phosphate

Figure 4-1: Molecular structures for the 6-phosphate sugars: fructose-6-phosphate, glucose-6-phosphate, mannose-6-phosphate and sucrose-6-phosphate.

Interestingly, trehalose, a disaccharide and a known osmoprotectant in *S. meliloti,* (Gouffi *et al,* 1999) (Figure 4-2) was also found to accumulate (>6 fold) in the polar extracts of *dme* and double mutant cells when compared to the polar extracts of Rm1021 and *tme* mutant cells. The presence of trehalose on the metabolic profiles of the *dme* mutant strain hints at either a disturbance in central carbon metabolism or osmotic stress as glucose-6-phosphate, via trehalose-6-phosphate synthase (*ots*AB) (Kassen *et* al, 1991) and sucrose-6-phosphate through conversion to UDP-glucose, may be used to synthesize trehalose. Sucrose is also a known osmoprotectant in *S. meliloti* (Gouffi *et al,* 1999) and

also lends further support to the hypothesis that the *dme* mutant is experiencing osmotic stress (sucrose-6-phosphate is easily converted to sucrose).

Figure 4-2: Molecular structure the of disaccharide, α-α-trehalose.



Trehalose

Osmotic stress is generally brought on by an increase in the amount of solutes in the extracellular environment of a cell. This increase in extracellular osmolarity leads to an efflux of intracellular water thus abolishing turgor pressure which is the driving force for cell elongation (as reviewed in Csonka and Hanson 1991). Osmotic stress also induces protein mis-folding (Hecker *et al*, 2001). Accumulation of compatible solutes including betaines, ectoines and amino acid derivatives like pipecolate reverses the water efflux (as reviewed in Csonka and Hanson 1991)

The soil environment is particularly susceptible to changes in osmolarity due to the effects of weather such as rain and drought (as reviewed in Csonka and Hanson 1991). In addition, the osmolarity of rhizosphere soil water is said to be elevated when compared to that of bulk soil due to the release of solutes through root exudation and the production of exopolymers by roots and rhizobia (as reviewed in Miller and Wood, 1996). Thus, not surprisingly, rhizobia adapt well to mild osmotic stress by accumulating potassium ions,

glutamate, N-acetyl glutaminyl glutamine amide (NAGGN) and trehalose (as reviewed in Miller and Wood, 1996). Unlike other bacteria, *S. meliloti* accumulates glycine-betaine only transiently (Pichereau *et al,* 1998) and does not accumulate ectoine and pipecolate but instead uses these as precursors to synthesize glutamate and NAGGN (Gouffi *et al*, 1999). Moreover, *S. meliloti* is unique as it uses disaccharides as osmoprotectants including maltose, cellobiose, gentibiose, turanose, palatinose, trehalose and sucrose among others (Gouffi *et al*, 1999). These sugars are not accumulated to osmotically significant levels, however like ectoine and pipecolate, the sugars stimulate a rise in glutamate and NAGGN levels (Gouffi *et al*, 1998; Gouffi *et al*, 1999). Since the *dme* mutant strain accumulates trehalose (in a growth-phase independent manner), sucrose-6-phosphate (which is easily converted to sucrose) and a cellobiose-like compound (when malate is provided as a sole carbon source for a 2 hour pulse), it seems reasonable to hypothesize that this strain becomes osmotically stressed when grown in succinate.

Interestingly, inactivation of the *zwf* gene (glucose-6-phosphate 1-dehydrogenase (G6PD)) in *S. meliloti* induces an osmosensitive phenotype and a loss in the osmoprotection provided by sucrose and trehalose (Barra *et al*, 2003). Moreover, this loss is not due to a defect in the biosynthetic pathways of these compatible solutes (Barra *et al*, 2003). It was also observed that *zwf* expression is induced by high osmolarity and the presence of sucrose and trehalose (Barra *et al*, 2003). Therefore it would be valuable to determine the phenotype of a *dme- zwf-* double knock-out mutant as this may support the hypothesis that *dme* mutant cells are osmotically stressed.

dme mutants have decreased intracellular concentrations of putrescine

Putrescine (Figure 4-3), a common polyamine found in rhizobia (Fujihara and Harada 1989), was present at decreased concentrations (< 4-fold) in intracellular extracts of the *dme* mutant strain. Putrescine can be synthesized by either the ornithine decarboxylase or arginine decarboxylase enzymes and is the precursor to both spermidine and spermine (as reviewed in Tabor and Tabor, 1984). Other polyamines include homospermidine and aminobutylhomospermidine (Fujihara and Harada 1989).

Figure 4-3: Molecular structure of the diamine, putrescine

Putrescine

Although correlated with cell growth and differentiation, the biological roles of putrescine and other polyamines are still under investigation (as reviewed in Tabor and Tabor, 1999). Polyamines are cationic and thus binds negatively charged molecules such as DNA and RNA (tRNA, rRNA, mRNA) enabling modulation of DNA conformation and RNA translation (as reviewed in Tabor and Tabor, 1999) permitting gene expression control. Other known roles of polyamines include protection of DNA from nucleases by

formation of nuclear aggregates (D'Agostino *et al,* 2006); stabilization of proteins; interaction with phospholipids to help stabilize membranes; and regulation of protein kinases (Bachrach *et al*, 2005).

Notably, decreases in putrescine concentration in the nodules *Lotus japanicus* and *Vigna mungo* have been correlated with decreases in nitrogenase activity (Lahiri *et al* 2004) and it has been stated that endogenous free polyamines may be considered as genotypic markers for nodule senescence (Lahiri *et al* 2004). *dme* mutant bacteroids have a Fix- phenotype and free-living *dme* mutant cells have a low cellular concentration of putrescine compared to wild type and *tme* mutants (which have Fix+ phenotype). It is possible that the reduced concentration of intracellular putrescine in the *dme* mutant strain is contributing to the Fix- phenotype in addition to other factors (including carbon source usage). Further investigation of this hypothesis includes inoculation of plants with a *dme* mutant strain overexpressing genes that are known to synthesize putrescine. It would also be interesting to examine the effect that knocking out putrescine synthesis genes has on nitrogenase activity in the bacteroid.

In *E. coli,* cells grown in media of high osmolarity (high concentrations of sodium chloride, potassium chloride, magnesium chloride or sucrose) showed deceases in cellular putrescine content due to excretion (Munro *et al*, 1972). In *Vibrio parahaemolyticus,* situations of low osmolarity resulted in a rapid production of putrescine (Yamamoto *et al*, 1989) and in *Rhizobium fredii,* the concentration of homospermidine was inversely proportional to sodium chloride concentration (Fujihara *et al,* 1993). (Homospermidine is synthesized from putrescine via the homospermidine synthase enzyme in *S. meliloti*

(*hss*)). All the points discussed above indicate that an osmotically stressed cell would have low intracellular putrescine and as discussed, primarily due to the accumulation of trehalose, we suspect that the *dme* mutant strain may be osmotically stressed when grown in M9-succinate. Thus, the reduced concentration of putrescine in *dme* mutant cells grown in succinate supports the osmotic stress hypothesis.

dme mutants excrete malate and fumarate into their extracellular environment

When the supernatants of cells pulsed for two hours with 5mM succinate were examined, those of the *dme* mutant strain were found to contain at least 10-fold more excreted malate and fumarate compared to those of the *tme* mutant strain. Surprisingly, the *dme* mutant cell excretes malate and fumarate in the millimolar range (and this excretion is time-dependent) whereas the concentration in the cell extract is in the nanomolar range. From this data it is clear that the *dme* mutant strain is unable to metabolize most of the malate that is synthesized from the succinate provided. The previous observation that DME activity (but not TME) is strongly inhibited by acetyl-CoA (Voegele *et al*, 1999) suggested that DME plays an important role in regulating flux through the TCA cycle. The excretion of fumarate and malate by succinate fed *dme* cells is consistent with this suggestion.

Since excretion of fumarate and malate is not observed in the *tme* mutant strain, it suggests that the TME enzyme may be executing another function in the cell, perhaps a biosynthetic one as Sanwal (1970) hypothesized for the NADP+-dependent malic enzyme

in *E. coli.* To test the anabolic-TME hypothesis, non-polar metabolic profiles could be generated to compare the phospholipid and fatty acid content of *dme* and *tme* mutant strain cellular extracts. If the role of the TME enzyme is to generate NADPH for lipogenesis, non-polar metabolite profiles may be different for *dme* and *tme* mutant cells.

dme mutants show a growth defect in M9 media containing glucose and succinate

When *dme* and *tme* mutant strains, were grown in M9 media containing equal concentrations of glucose and succinate, a difference in growth rate was observed; the *dme* mutant strain grew at approximately one half the rate of both the *tme* mutant and wild type strains (Figure 3-13). Intracellular metabolic profiles of the *dme* mutant strain grown in M9-glucose and succinate showed at least 20-fold greater levels of fumarate, aspartate and malate, respectively when compared to those of the *tme* mutant strain (Figure 3-14). These results suggest that the addition of glucose to the media may cause the *dme* mutant to use the ED (Entner-Doudoroff) pathway preferentially over TCA cycle leading to an intracellular accumulation of malate and fumarate as well as aspartate (which is derived from OAA). Extracellularly, the concentrations of malate and fumarate were ten-fold greater in the *dme* mutant strain compared to the *tme* mutant strain. The intracellular accumulation of malate and fumarate results from a reduction in flux through the TCA cycle and since the glycolytic pathway generates much less ATP compared to that generated by the TCA cycle, a slowed growth rate is observed.

Summary of metabolic data

From the data discussed above, a couple of hypotheses may be proposed. The first (Figure 4-4) is that when the *dme* mutant cell is grown in M9-succinate media, malate cannot be converted to pyruvate and so malate is converted to OAA. Excess OAA causes an increased flux through gluconeogenesis which creates an excess of the fructose, mannose and glucose-6-phosphates. Glucose-6-phosphate can then be used to make sucrose-6-phosphate and trehalose. Malate that is not converted to OAA is directly excreted or converted to fumarate and excreted. It is probable that succinate is also made in the reverse reaction but that the change in extracellular concentration is negligible as it is present in the original media.

Figure 4-4: Diagrammatic summary of the metabolic situation of *dme* mutants grown in M9-succinate

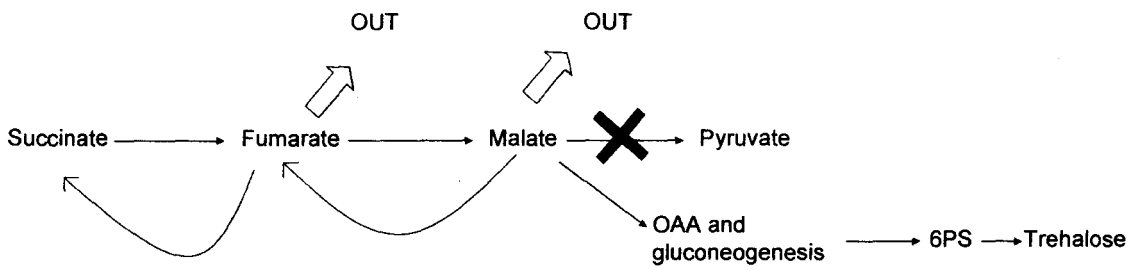


Figure 4-4: Schematic representation of metabolic situation of the *dme* mutant strain grown in M9-succinate. OUT represents excretion into media. X represents non-functioning enzyme.

When *dme* cells are grown in M9-glucose and succinate, there is increased flux through the ED pathway. Since much less ATP is generated through the ED pathway compared to the TCA cycle, if *S. meliloti* cells were to live preferentially via this pathway, cell growth might be retarded.

It is also possible that growth is slowed due to the accumulation of TCA cycle intermediates. When cells are grown in M9-succinate (no glucose) excess malate may be converted secondarily to intermediates of the gluconeogenic pathway. However, the presence of glucose in the media results in a primary accumulation of gluconeogenic intermediates preventing the flux of intermediates derived from malate thus causing malate and fumarate to build up inside the cell in addition to being excreted.

A second hypothesis is that perhaps the excretion of millimolar amounts of malate and fumarate is strenuous for the cell causing a response similar to that executed under situations of osmotic stress. Coincidently, osmolytes not only protect the cell from

osmotic stress but also from other physical stresses such as high temperature, desiccation and freezing (as reviewed in Csonka *et al*, 1991). The accumulated 6-phosphate sugars, trehalose and the decreased putrescine concentration in the *dme* mutant cells, could all be considered to be symptoms of stress in the cell.

Conclusion

The aim of this project was to determine if a metabolic difference existed between the *dme* and *tme* strains which might explain the different symbiotic phenotypes of these mutants. When considered as a whole, the metabolic data clearly suggests that the DME enzyme is principally responsible for the conversion of malate to pyruvate to generate acetyl-CoA whereas the TME enzyme must serve a secondary function within the cell; perhaps through the generation of NADPH for anabolic purposes.

CHAPTER 5

Introduction: protein interactions

Modular structure of malic enzymes

As previously outlined, *S. meliloti* has two malic enzymes, DME and TME which require NAD+ and NADP+, respectively, as co-factors. The DME enzyme is required for symbiotic nitrogen fixation whereas the TME enzyme is not. (Driscoll and Finan 1993; Driscoll and Finan 1996). However, despite differences in symbiotic phenotype, the two malic enzymes of *S. meliloti* share a unique modular structure (Figure 5-1) (Mitsch *et al* 1998.)

Figure 5-1: Schematic diagram depicting the modular structure of the DME and TME malic enzymes

Figure 5-1: Schematic diagram depicting the modular structure of the *dme* and *tme* malic enzymes of *S. meliloti*. N and C represent the amino and carboxy termini of the protein, respectively. PTA stands for phosphotransacetylase. Pta enzymes catalyze the phosphorylation of acetyl-CoA.

The 456 amino acid (a.a.) N-terminal domain is responsible for malic enzyme activity and the 328 a.a. C-terminal domain shows some homology at the amino acid level (32.5% identity) to phosphotransacetylase (PTA) enzymes, which catalyze the phosphorylation of acetyl-CoA (Hahm *et al*, 1994) (acetyl-CoA + phosphate = CoA + acetyl phosphate). Mitsch *et al* (1998) showed that the PTA-like domain is not essential for DME enzyme activity and is not required for nitrogen fixation. Attempts to detect PTA activity from the C-terminal domains of DME and TME have been unsuccessful and thus the function of this domain remains unknown. However, the modular structure of malic enzymes is conserved across a broad spectrum of α, β and γ proteobacteria (Mitsch *et al,* 1998 and this work). Interestingly, this conserved PTA-like domain is also present in some archaebacteria including *Halobacterium salinarum* and *Haloarcula marismortui* (this work). The broad conservation across bacterial species of this C-terminal PTA-like domain suggests that it may serve an important, albeit unknown role.

pTrcSC vector

To investigate this phenomenon, we have constructed C-terminal translational fusions with each of the malic enzyme domains to the Streptavidin (StrepII) (Korndorfer *et al*, 2002) and the Calmodulin Binding (CBP) peptides using the pTrcSC vector (Baron *et al*, unpublished). pTrcSC was built by the cloning of both strep tag polylinkers and Tobacco Etch Virus-CBP polylinkers into the pTrc200 vector (Schmidt-Eisenlohr *et al*, 1999) which is a pVS1 derivative possessing an RK2 origin of replication. The pTrcSC

vector (Figure 5-2) is a replicating vector that carries a hybrid *trp/lac* promoter which is IPTG inducible. The vector also possesses a ribosome binding site and ATG start codon and if desired, the StrepII and CBP tags may be removed via Tobacco Etch Virus (TEV) cleavage.

The methodology of affinity tagging is more commonly known as TAP-tagging, where TAP stands for Tandem Affinity Purification (Rigaut *et al*, 1999). The standard TAP tag consists of CBP followed by the ProteinA IgG-binding domain with a TEV cleavage site located in between the two tags (Riguat *et al,* 1999). Characteristics of a good tag include: small size (ie: 2-26 amino acids), ability to be excised enzymatically, allowance for detection of the recombinant product using antisera and minimal interference with protein structure and/or activity (Terpe, 2003). The StrepII tag (Korndorfer *et al*, 2002) used in this work is eight amino acids in length (WSHPQFEK) with a mass of 1.06kDa (Schmidt and Skerra, 1993; Terpe, 2003). The CBP tag is 26 amino acid residues in length (KRRWKKNFIAVSAANRFKKISSSGAL) with a mass of 2.96kDa. CBP is derived from skeletal muscle myosin light chain kinase and has the ability to bind Calmodulin in the presence of calcium chloride (Blementhal *et al,* 1985)

Due to the non-denaturing conditions of the purification procedure, this two-step method allows for the purification of protein complexes without any prior structural knowledge. Potential interacting partners may then be dissociated via SDS-PAGE, excised and analyzed using mass spectroscopy.

Figure 5-2: pTrcSC vector

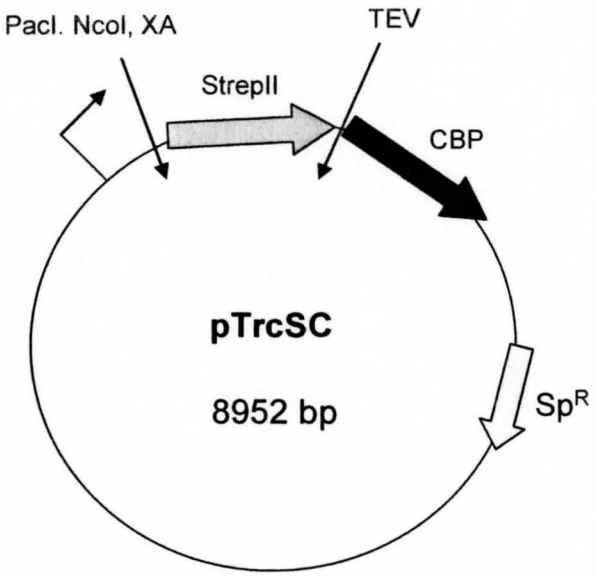


Figure 5-2: pTrcSC vector derived from plasmid pTrc200 which is derived from pVS1 which has an RK2 origin of replication. pTrcSC contains a hybrid *trp/lac* promoter which is IPTG inducible as well as a ribosome binding site and a start codon. When DNA fragments are cloned into this vector via PacI, a translational fusion to the StrepII and Calmodulin binding peptides is created. (Baron *et al*, unpublished)

THIS WORK

To investigate a possible role for the C-terminal domain in protein-protein interactions, we constructed C-terminal translational fusions with each of the malic enzyme domains (Figure 5-1) to the Streptavidin (StrepII) (Korndorfer *et al*, 2002) and the Calmodulin Binding (CBP) peptides using the pTrcSC vector (Baron et al, unpublished) (Figure 5-2)

We tested the expression of all fusions and confirmed that the full-length DME fusion and the fusion to the malic enzyme domain are biologically active. The fusions to the full

DME peptide and to the malic enzyme domain have been successfully purified. We have also identified proteins that co-purify with both fusions. Unfortunately, western blot analysis with antibodies specific for the DME protein indicates majority of the co-purifying proteins consist of multimers or degradation products of DME.

CHAPTER 6

Materials and methods: protein interactions

Primer design and polymerase chain reaction (PCR)

Primers for PCR were designed using the Gene Runner 3.05 program (Hastings Software 1994) and melting temperatures and molarity were estimated using the Oligo Calculator program (Biomedical Research Support Facilities, http://www.pitt.edu/~rsup/OligoCalc.html). Primers were ordered from Sigma-Genosys, dissolved upon arrival in autoclaved, distilled water to final concentrations of 100pmol/μL and stored at -20°C until required. All primers used in this study are listed below in Table 6-1.

PCR was conducted in 100μL volumes consisting of the following in distilled, deionized water: 1x Platinum *Taq* polymerase buffer (Invitrogen), 0.2mM of each deoxynucleotide-triphosphate (dATP, dTTP, dCTP and dGTP), 100 pmol of each forward and reverse primers, 1.0-2.0mM $MgCl_2$ and 2.5 units of Platinum *Taq* polymerase (Invitrogen). In some cases 1%-5% DMSO was added. The pUC119 derivative, pTH408 was used as DNA template at 1ng/μL for construction of DME fusions. Rm1021 genomic DNA at 2ng/μL was used as a template for the construction reporter gene fusions in pTH1703. 5μL of the appropriate template DNA (1-2ng/μL) was used for each 100μL reaction.

Using the Eppendorf Mastercycler epgradient S, the reaction tubes were heated at 94°C for two minutes then each reaction was cycled 25 times where denaturation took place at 94°C for 30 seconds, primer annealing at 55 to 65°C for 45 seconds and extension at 72°C for one minute per each kb of the resulting product. Note that annealing time was calculated at 5°C less than the melting temperature of the primer pair. Following the 25 cycles, the reactions were incubated at 72°C for 7 minutes and then cooled to 4°C.

10μL aliquots of the resulting products were subsequently electrophoresed through agarose gels (0.8-1.3% depending on the product size) in 0.5xTBE to verify the success of the PCR. Products were purified using a commercially available kit (Qiagen, QAIquick quickspin PCR purification kit) following the manufacturers protocol.

Table 6-1. Names and sequences of primers used to construct the listed plasmids

Plasmid Name	Primer names	Primer Sequences	Restriction sites
pTH1918	10LAS	GG**TTAATTAA**CATGAAGCCGGTCTTTGCTGC	PacI
	8LAS	CC**TTAATTAA**GACGGGATGGCTTGATTCGACG	PacI
pTH1930	7LAS	CC**TTAATTAA**CATGAACACGGGCGATAAA	PacI
	9LAS	CG**TTAATTAA**GTTCTTCGCCGCGTTCTTCGC	PacI
pTH1957	7LAS	CC**TTAATTAA**CATGAACACGGGCGATAAA	PacI
	8LAS	CC**TTAATTAA**GACGGGATGGCTTGATTCGACG	PacI
pTH2068	14LAS	GC**CTCGAG**CAGCGTCAAGCGTGAGC	XhoI
	15LAS	AC**GGGCCC**GCCTAATGGAACAGCACGC	ApaI
pTH2069	16LAS	GG**CTCGAG**TCAATGCTGCCCAGTTCG	XhoI
	17LAS	GA**GGGCCC**TAGGATCAGCCCTTGAGC	ApaI
pTH2070	18LAS	CG**CTCGAG**GATGAGGCTGTCCTCG	XhoI
	19LAS	TA**GGGCCC**TTGGTCGTCGCACAGG	ApaI
pTH2081	22LAS	TA**CTCGAG**CGAGAAGGTCTTGAAGGCCGC	XhoI
	23LAS	CC**ATGCAT**TCAAGCGTTCAGACCGCTGCC	NsiI
pTH2082	24LAS	TG**CTCGAG**CTGCACGACCAACTGCCT	XhoI
	25LAS	AA**GGGCCC**AGCCCGGAACATCGGTCT	ApaI

Manipulation of DNA, restriction endonuclease digests and ligations

The restriction enzymes used were purchased from Invitrogen, Fermentas and Roche and used according to specifications listed by the product's manufacture. Generally 1x reaction buffer and 10 units of enzyme were used. If the digest was conducted to verify the presence of a plasmid or to determine the success of a ligation, the reaction products were subjected to agarose gel electrophoresis in 0.5xTBE, stained in ethidium bromide, destained in distilled, deionized water and photographed on the Pharmacia Biotech image master VDS UV transilluminator.

PCR products required for ligations were purified using a commercially available kit (Qiagen, QAIquick quickspin PCR purification kit) and the PCR purification. If multiple products were present following PCR amplification, the desired DNA fragment was purified by agarose gel electrophoresis and eluted using Qiagen, QAIquick quickspin PCR purification kit.

Three PCR products, 2311bp, 1369bp and 985bp in size amplifying the full DME protein, the N-terminal domain and the C-terminal domain respectively were cloned non-directionally into the pTrcSC vector at the PacI restriction site generating N-terminal translational fusions to the strep/CBP tag. The pTrcSC vector was treated with alkaline phosphatase following digestion with PacI but prior to ligation. The treatment was performed using one unit of calf intestine alkaline phosphatase (Roche), 800-1000ng of DNA and 10x alkaline phosphatase buffer (Roche). This mixture was incubated for one hour at 37°C.

Ligations were conducted using approximately 200ng of vector DNA, approximately 600ng of insert DNA, 400 units T_4 DNA ligase (New England Biolabs) and 1x T_4 ligase buffer (New England Biolabs). This was generally done in a 10μL volume at 16°C overnight and then transformed into competent *E. coli* DH5α.

Preparation of competent *E. coli*

E. coli cells were made competent by treatment with calcium chloride. Overnight cultures of DH5α were inoculated into 2mL of LB then subcultured the following day into 100mL of LB to an OD_{600nm} of approximately 0.05 and grown for about two hours to an OD_{600nm} of 0.4. The cells were immediately placed on ice for 15 minutes and then divided into two 50mL aliquots. The two aliquots were spun at 3730g (SX4250 rotor) at 4°C for 20 minutes and the supernatants were decanted. The pellets were resuspended in 25mL of ice-cold 0.1M $CaCl_2$ and incubated on ice in the cold room overnight. The following day, the aliquots were centrifuged using the same conditions described above. The cell pellets were each resuspended in 5mL ice cold 0.1M $CaCl_2$ with 15% glycerol. Competent cells were stored long term by flash freezing 150μL aliquots in liquid nitrogen and then placing at -80°C until required.

Transformation of competent *E. coli* DH5α

Aliquots of frozen competent cells were thawed and added to ligation mixtures or plasmid DNA (< 1μg) and placed on ice for 30 minutes. The mixture was then heat shocked at 40-42°C for one minute and subsequently placed on ice for five minutes. Next, 500μL of LB broth was added to the mixture which was then incubated at 37°C for one to two hours with shaking. Following incubation, the cells were centrifuged at 16,100g for one minute, the supernatant was decanted and the cells were resuspended in the remaining media (~100μL). This was spread plated onto LB agar containing the appropriate selective antibiotic and incubated overnight at 37°C.

Preparation of *S. meliloti* electro-competent cells

2mL overnight cultures of *S. meliloti* were grown in LBmc, subcultured the following day into 100mL of LBmc (containing 1.25mM $MgSO_4$ and $CaCl_2$) and grown to an OD_{600nm} of ~0.8. The cells were then pelleted by centrifugatioin at 4°C and 3730g (SX4250 rotor) for 20 minutes. The cell pellets were washed twice in 10mL of cold 10% glycerol (in ddH_20), each time the pellet was resuspended and the cells centrifuged using the conditions outlined above. The cells were then resuspended in 10mL of cold 10% glycerol (in ddH_20), flash frozen in 100μL aliquots and placed at -80°C until required.

Electroporation of S. meliloti competent cells

100μL aliquots of electro-competent *S. meliloti* cells were thawed on ice and then added to 1μL (~200ng) of plasmid DNA. The DNA and cell mixture was transferred to a chilled cuvette (Biorad) and then pulsed on the BioRad Pulse Controller Plus at 2.5 V, a capacitance of 25μF and a resistance of 200ohms. 900μL of LBmc was immediately added, the cells were allowed to recover for at least two hours, and then plated onto selective medium and incubated at 30°C.

Preparation of plasmid and genomic DNA for sequencing

Plasmid DNA was isolated from 2-5mL overnight cultures using commercially available kits (Sigma) according to manufacturers protocol.

To isolate genomic DNA from *S. meliloti* for sequencing, cells were inoculated into 4mL of LBmc and grown overnight at 30°C. The cultures were transferred to 2mL microfuge tubes and spun at 16 100g for one minute. The pellet was resuspended in 1mL of 0.85% NaCl and centrifuged for one minute at 16 100g. The pellet was resuspended in 700μL $T_{10}E_{25}$ (10mM Tris-HCl pH 7.5 and 25mM EDTA). 35μL of 25% SDS was added and the tube was inverted to mix the solution, 35μL of 10mg/mL proteinase K was added and the tube was inverted to mix the solution, 175μL of 5M NaCl was added and the tube was swirled gently to mix and the resulting mixture was incubated at 65°C for two hours. After cooling to room temperature, the mixture was extracted once with buffer saturated

phenol, twice with phenol:chloroform (1:1, v:v) and then once with chloroform. 70μL of 5M NH_4CH_3COO and 770μL of isopropanol were added to the aqueous supernatant then centrifuged at 16,100g for 10 minutes. The subsequent pellet was washed with 0.5mL of 70% ethanol and dried at 37°C for 10 minutes. The pellet was dissolved in 400μL of 10mM Tris-HCl pH8.0 with 1mM EDTA and 20mg/mL RNaseA and incubated at 37°C for 30 minutes. Another extraction was performed with an equal volume of 1:1 Phenol: Chloroform. To this, 400μL of 5M NH_4CH_3COO and 40μL of absolute ethanol were added, vortexed and the mixture was centrifuged for 10 minutes as before. The pellet was washed with 0.5mL of 70% ethanol and dried at 37°C for 10 minutes. The resulting pellet was dissolved in 100μL of distilled, deionized water and 5μL of this solution was run on an agarose gel to ensure that the DNA was not degraded. An additional 900μL of ddH_2O was added and the OD at 260 and 280nm was read and the 260/280 ratio was calculated to find the purity and mass of the DNA. Finally, the sample was dried down in the Eppendorf Vacufuge for 4.5 hours at 45°C and then dissolved in ddH_2O to a concentration of 2μg/μL.

All sequencing was performed in the MOBIX laboratory. For sequencing of constructs built in pTrcSC, the sequence for the 5' primer was (5' GTGTGGAATTGTGAGCGGAT 3') and for the 3' primer was (5' CTACTGCCGCCAGGCAAATT 3'). For all constructs build with pTH1703, the sequencing primer was ML4875 (5' TTAGGACAACTCCAGTGAAAAGTTC 3'). Sequences obtained were compared to the wild type sequence using the AlignQuery program (Genestream network)

Over-expression of DME- small scale

pTrcSC- DME constructs from overnight cultures of *E.coli and S meliloti* were subcultured into 5mL volumes of LB or LBmc containing spectinomycin (50μg/mL for *E. coli* and 100μg/mL for *S. meliloti*) . Cultures were grown to OD_{600nm} of 0.4 then induced by addition of IPTG to 0 0.5mM and grown for four hours. Following induction, 1mL of culture was centrifuged at 16 100g for one minute, the pellet resuspended in 2x SDS-PAGE loading buffer (recipe for 1x= 240mM Tris-HCl pH6.8, 6% SDS, 30% glycerol, 2.24% β-mercaptoethanol, 0.6g/L bromophenol blue) and heated at 95°C for 10 minutes then centrifuged as before. Supernatants were then loaded onto 10% acrylamide SDS-PAGE (Maniatis Sambrook) and expression was visualized by western blot.

Over-expression of DME- large scale

For large scale over-expression of DME from *S. meliloti,* overnight cultures were subcultured into 50mL LBmc and incubated overnight. Cultures were subcultured into 4 x 2.5L flasks containing 1L LBmc with 100μg/mL spectinomycin, grown to OD_{600nm} of 0.4 and then induced by addition of IPTG to a final concentration of 0.5mM and then grown for four hours. Following induction the cells were centrifuged for one hour at 3730g (SX4250 rotor) at 4°C, the pellets were flash frozen in liquid nitrogen and placed at -80°C. As preparation for lysis the cells were resuspended in Strep Wash buffer (10mL for each Liter) (100mM Tris-HCl pH8.0, 150mM NaCl and 1mM EDTA). The four

resuspended pellets were pooled and PMSF dissolved in absolute isopropanol was added to a final concentration of 0.5mM. Cells were lysed by passing through the French Press three times at 10 000psi, after which the lysate was centrifuged at 14 000g (SS34 rotor) and 4°C to remove debris.

Purification of DME constructs in pTrcSC

A) Streptactin column

A Streptactin Superflow column (IBA) was constructed with a final bed volume of 4mL. The column was washed with two volumes of Strep Wash buffer (100mM Tris-HCl pH8.0, 150mM NaCl and 1mM EDTA) prior to addition of the cell lysate. Un-bound proteins were removed by the application five aliquots of one column volume of Strep Wash buffer. 0.5mL aliquots of each cell lysate, flow-through and washes were kept for SDS-PAGE analysis. Bound proteins were removed by applying 6x 0.5 column volumes of Strep elution buffer (100mM Tris-HCl pH 8.0, 150mM NaCl, 1mM EDTA and 2.5mM D-desthiobiotin), and collecting 1mL fractions. The column was regenerated through the application of 3x five column volumes of Strep buffer R (100mM Tris-HCl pH 8.0, 150mM NaCl, 1mM EDTA and HABA, this buffer was purchased through IBA), followed by washing with 2x four column volumes of Strep wash buffer and then storing at 4°C until needed.

The four fractions containing the highest concentrations of the recombinant protein as determined by SDS-PAGE were pooled protein concentration was determined using the BioRad protein assay.

B) Calmodulin column

50µL of suspended Calmodulin beads (Stratagene, calmodulin affinity resin) ca bind 150µg of protein. 2mL bed volume of Calmodulin resin was resuspended in five column volumes of Calmodulin binding buffer (10mM TrisHCl pH8.0, 100mM NaCl, 2mM $CaCl_2$, 0.1%Triton X-100 and 10mM β-mercaptoethanol) and allowed to flow through, this was subsequently repeated. To the pooled fractions from the strep column, 1.2µL of 1M $CaCl_2$ per 0.5mL eluate was added and then the mixture was applied to the Calmodulin beads. The column was resuspended and incubated, rotating at 4°C for three hours. Following incubation, the lysate was allowed to flow through the column which was then washed with 30mL of Calmodulin binding buffer. 0.5mL aliquots of the lysate, flow-through and wash were saved for SDS-PAGE analysis. The bound proteins were eluted using five column volumes of Calmodulin elution buffer (10mM Tris-HCl pH8.0, 100mM ammonium bicarbonate, 3mM EGTA and 10mM β-mercaptoethanol), 0.5mL fractions were collected and all fractions were analyzed by 10% acrylamide SDS-PAGE. The column was regenerated using the protocol outlined in the Stratagene booklet accompanying the affinity resin #214303.

The fractions from the Calmodulin column containing the highest concentrations of recombinant protein as determined by SDS-PAGE were pooled and dialyzed against 1L of cold dialysis buffer (50mM TrisHCl pH8.0, 50mM KCl, 10% glycerol, and 1mM DTT) overnight at 4°C. The dialyzed sample was dispensed into 100µL aliquots, flash frozen in liquid nitrogen and placed at -80°C.

The remaining fractions from the Calmodulin column (containing lower amounts of recombinant protein) were precipitated by addition of an equal volume of 30% trichloroacetic acid, incubating for 30 minutes on ice and then centrifuging at 16 100g for 15 minutes. Following centrifugation, the supernatant was removed, 300μL of cold acetone was added and the tube was centrifuged as before. The resulting pellet was dried in the fume hood and resuspended in 4X SDS-PAGE loading buffer and stored at -80°c until required.

Western blot Analysis

SDS-PAGE gels were transferred to Hybond-P membrane (Amersham Pharmacia Biotech) according to manufacturers' protocols, using a Biorad Transblot SD semi dry transfer cell. The transfer buffer consisted of 20% methanol, 50mM Tris and 39mM glycine and the transfer was carried out for one hour. The membrane was then placed into 1x PBST (phosphate-buffered saline with 0.1% Tween20) with 3% BSA and incubated, shaking overnight at 4°C.

The membrane was washed three times in PBST, five minutes per wash. Primary antibodies were diluted as follows: rabbit αDME antibody (Mike Mitsch) was diluted 1 in 2000 with PBST; StrepII antibodies (IBA) were diluted 1 in 200 with PBST. Diluted antibodies were then added to the membrane and incubated for one hour at room temperature. The membrane was then washed three times in PBST, five minutes for each wash. The secondary antibody for the rabbit αDME antibody was the alkaline phosphatase conjugated mouse α-rabbit antibody (Biorad) which was diluted 1 in 3000

with PBST. The secondary antibody for the strepII antibodies was an alkaline phosphatase conjugated goat α-mouse antibody (Biorad) diluted 1 in 3000 with PBST. Membranes were incubated at room temperature for 45 minutes then washed twice with PBST and once with PBS (phosphate-buffered saline).

Blots were developed in alkaline phosphatase buffer (100mM Tris-HCl pH 9.5, 100mM NaCl, 5mM $MgCl_2$). To 30mL of alkaline phosphatase buffer, 1mL of 10mg/mL Nitro blue tetrazolium (NBT) (sigma) and 250μL of 15mg/mL BCIP (Biorad) dissolved in 100% N-N-dimethylformamide were added. The membrane was soaked in this solution until adequate colour had developed and the reaction was stopped by the addition of ddH_2O.

Antibody purification

Several aliquots of purified tagged-purified DME protein were subjected to 10% acrylamide SDS-PAGE followed by transfer to a PVDF membrane (as done for western blotting). The membrane containing the transferred protein was stained in Ponceau Red (Sigma) (2.5% Ponceau, 20% methanol) for approximately ten minutes. Fragments known to contain tagged DME were cut out and blocked with 3% BSA in TBST (Tris-buffered saline with 0.1% Tween20) for one hour. The blocked DME was incubated overnight 4°C in 1mL of TBST with 100μL of DME antibodies. The following day, the membrane was washed five times with TBST, where each wash was conducted at room

temperature for five minutes. The purified antibodies were eluted with 500µL of 0.1M glycine and 22µL of 1M Tris so the final pH was between 7.0-8.0.

Protein determination

The concentration of protein in crude cell extracts was determined using the Bradford method (Bradford, 1976). A standard curve was constructed using quantities of BSA ranging from 1µg to 20µg. 5 and 10µL of eluate was tested for each protein that was purified. For both the standard curve and the samples, 20% Coomassie blue R250 Biorad Protein Assay Dye Reagent was used.

Malic enzyme assays

Malic enzyme assays were performed as outlined by Driscoll (thesis, 1995). 5µL of purified protein or cell lysate was added to 1mL of malic enzyme buffer (100mM Tris-HCl pH 7.8, 30mM K-L malate pH 7.8, 3mM $MnCl_2$, 50mM KCl and 0.25mM NAD+), the 3mL cuvette was inverted to mix and readings were taken over a five minute period in the Cary Varian spectrophotometer. Change in absorbance was read at 340nm which is a result of the reduction of NAD+ to NADH. The specific activity was calculated in nmoles of NADH produced per minute per mg protein using the following formula: slope(ΔA_{340})/0.00622/mg protein.

Bioinformatic analysis of DME

To highlight the level of conservation in the PTA-like domain of malic enzymes across a diversity of species, multisequence analysis was conducted and a phylogenetic tree was constructed. ClustalX was used to construct the multisequence alignment. The MRBAYES program (Huelsenbeck and Ronquist, 2001) was used to construct phylogenetic tree for malic enzymes.

CHAPTER 7

Results: protein interactions

Bioinformatic analysis of DME

Despite a difference in function which can be inferred from the metabolic profiling as discussed in Chapter 3, both DME and TME share a unique modular structure where the N-terminal domain of the enzyme is responsible for malic enzyme activity and the C-terminal domain is somewhat similar in sequence to phophotransacetylase enzymes (Pta) (Mitsch *et al*, 1998). Attempts to detect Pta activity from the C-terminal domains of DME and TME have been unsuccessful and the function of this domain remains unknown.

Here we show that the modular structure of the malic enzymes of *S. meliloti* is also shared across a diverse spectrum of α, β and γ proteobacteria including *Agrobacterium tumefaciens*, *Bordetella pertusis and Escherichia coli* (Figures 7-1 and 7-2). Interestingly, this conserved Pta-like domain is also present in some archaebacteria including *Halobacterium salinarum* and *Haloarcula marismortui*.

Figure 7-2: : Multiple alignment constructed in ClustalX showing conservation between Pta-like domains of malic enzymes of 10 selected microorganisms

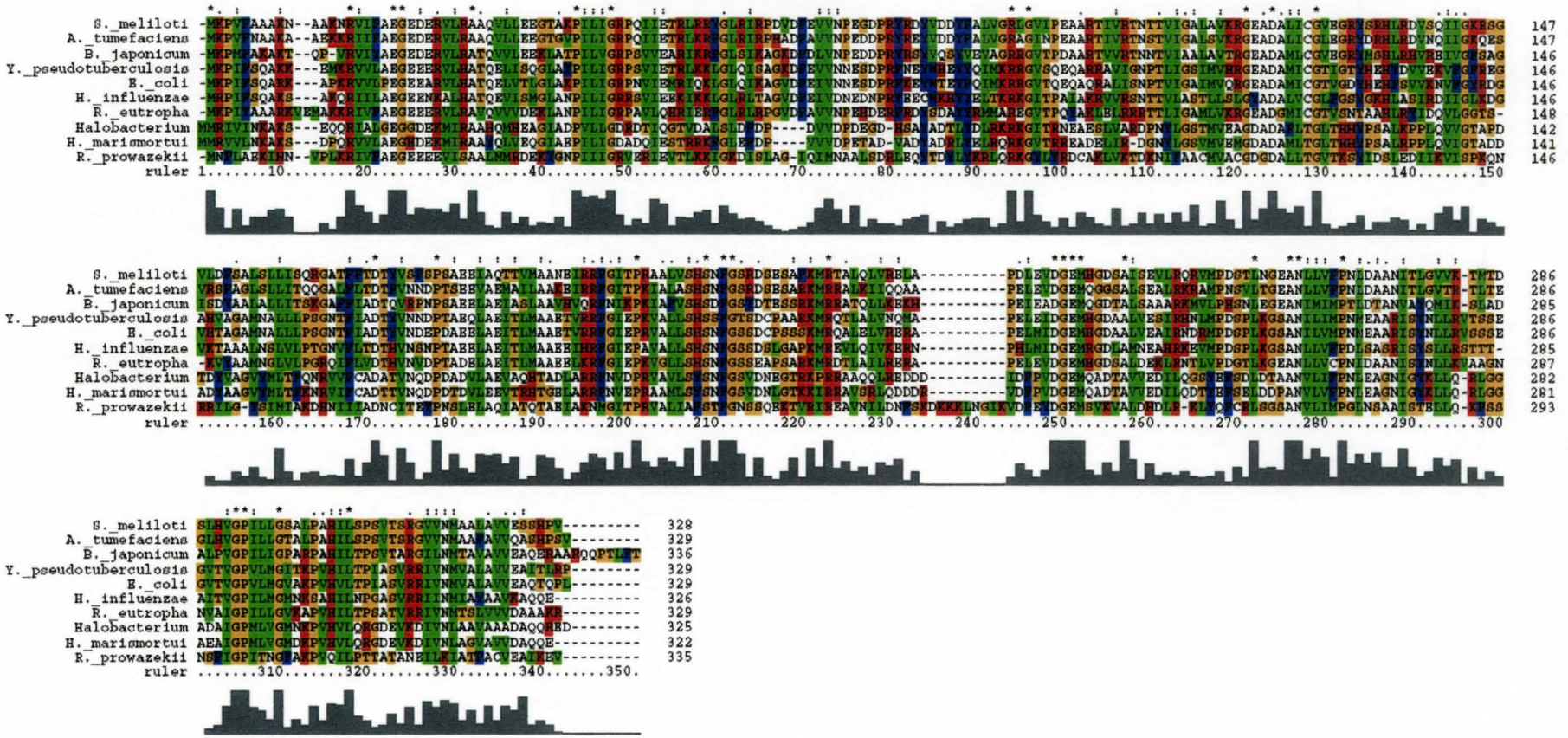


Figure 7-2: Multiple alignment constructed in ClustalX (Jeanmougin, 1998) showing conservation between Pta-like domains of malic enzymes of 10 selected microorganisms. Note the striking similarity between the Pta-like regions of modular malic enzymes. Highlighted amino acids are those belonging to characteristic groups: Green= methionine,leucine, isoleucine, valine, Red= lysine, arginine, histidine , Yellow= serine, threonine, proline or glycine, Blue= phenylalanine, tyrosine and tryptophan and white = alanine, asparagine, glutamic acid, aspartic acid, cysteine and glutamine. The degree of conservation is indicated by the bar graph located under the alignment, where higher bars represent greater conservation.

Figure 7-2: : Phylogenetic tree of Malic enzymes from a range of bacterial species generated through ClustalW alignment and the MRBAYES program

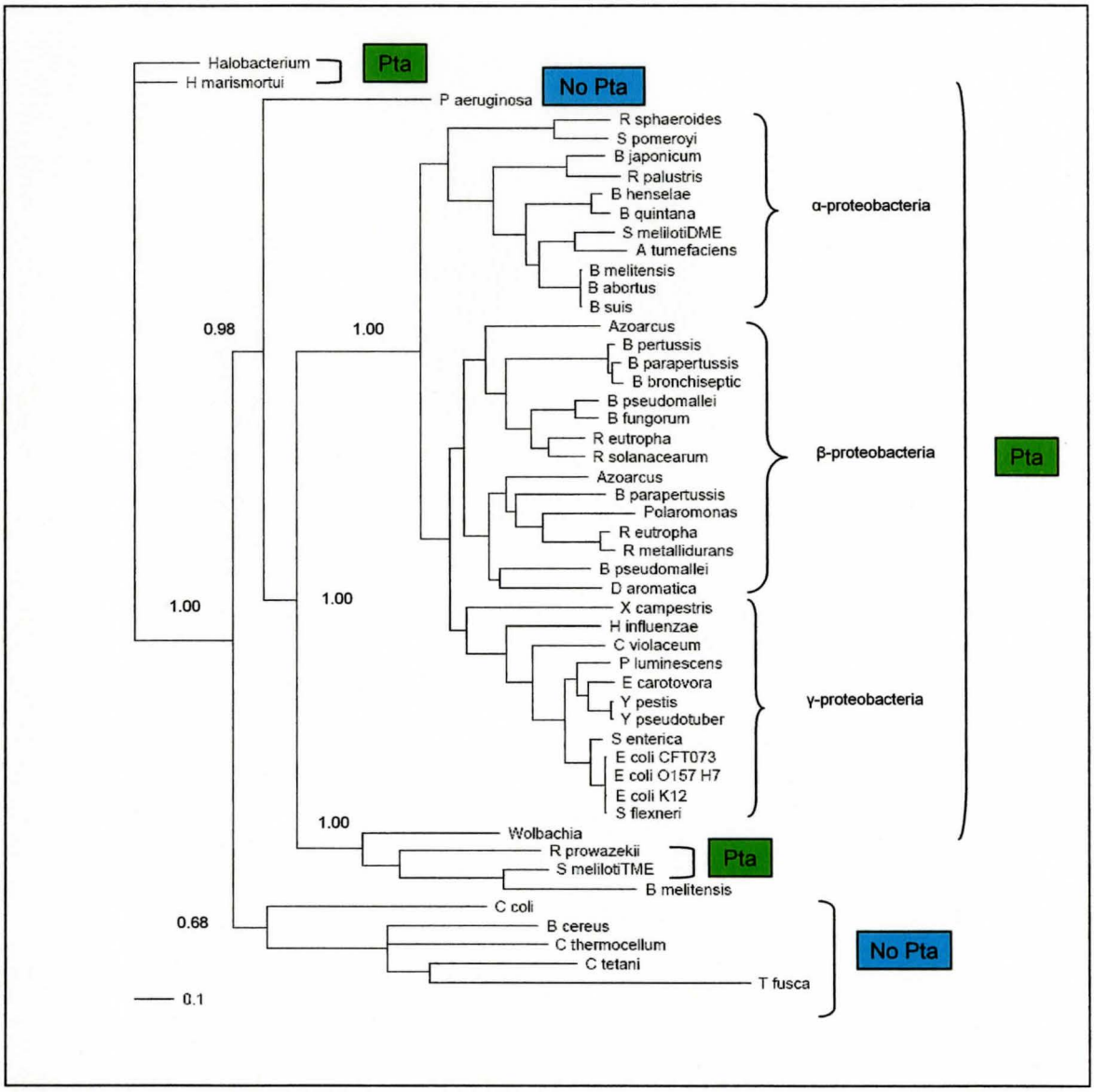


Figure 7-2: Phylogenetic tree of malic enzymes (full sequence) from a range of bacterial species generated through ClustalW alignment and the MRBAYES program (Huelsenbeck and Ronquist, 2001). The C-terminal Pta-like domain is conserved across many species of proteobacteria including *Bradyrhizobium japonicum Bordetella pertussis* and *Escherichia coli.* In addition, this structure is observed in the malic enzymes of *Halobacterium salinarum* and *Haloarcula marismortui.* Some species lacking the C-terminal Pta-like domain from their malic enzymes were also included in this tree: *Campylobacter coli, Bacillus cereus, Clostridium tetani, Clostridium thermocellum, Thermobifida fusca, Wolbachia, Brucella melitensis* and *Pseudomonas aeruginosa.* See appendix for accession numbers.

From Figure 7-1, it is obvious that the Pta-like domain of the malic enzyme is highly conserved. The crystal structure of the Pta enzyme has been elucidated in both *Bacillus subtilis* and *Methanosarcina thermophila* (Xu *et al*, 2005; Iyer *et al*, 2004). The structural data and sequence data for Pta enzymes has identified eight amino acid residues that are strictly conserved (Xu *et al*, 2005). These include G124, D166, D240, D251, N271, K277, S303 and R304. A multiple alignment was conducted using ClustalW which included protein sequences for DME and six other modular malic enzymes in addition to Pta sequences from *S. meliloti* and *B. subtilis* (data not shown). Using this alignment, we were able to see that five of the eight strictly conserved residues were also conserved or partially conserved in the Pta-like region of modular malic enzymes. Table 7-1 below outlines which amino acids are conserved.

Table 7-1: Conservation of strictly conserved residues of the Pta enzyme in modular malic enzymes

Residue	Conservation in malic enzymes
G124	Yes- complete
D166	Yes- complete
D240	All but *A. tumefaciens* (D→G)
K251	No
N271	Yes
K227	No
S303	Patially- either S or T
S304	No

Purification of DME fusions and co-eluting proteins

To investigate a possible role for the C-terminal domain in protein-protein interaction, we have constructed translational fusions with each of the malic enzyme domains to the Streptavidin (StrepII) (Korndorfer *et al*, 2002) and the Calmodulin Binding (CBP) peptides using the pTrcSC vector (Baron *et al*, unpublished). Figure 7-3 below depicts the fusions constructed using pTrcSC.

Figure 7-3: : Schematic depicting protein fusions between a small TAG consisting of the Streptavidin and Calmodulin binding peptides

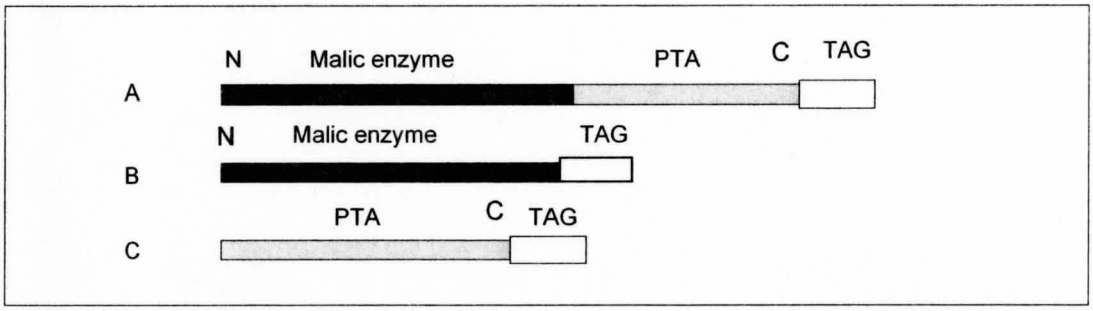


Figure 7-3: Schematic depicting protein fusions between a small TAG consisting of the Streptavidin and Calmodulin binding peptides and a) the full length malic enzyme, b) the N-terminal domain of malic enzyme, and c) the C-terminal PTA-like domain.

Expression of these C-terminally tagged constructs in pTrcSC was successful for translational fusions to the full length DME peptide and to its N-terminus (Figure 7.3, A and B). However expression of the tagged C-terminus of DME (i.e. PTA-like domain) could only be visualized with DME-specific antibodies and not strep II specific antibodies. This was observed when fusions were expressed in either *Escherichia coli* or *S. meliloti* (Figure 7-4). It is possible that this construct was unstable, degraded, or

insoluble. Interestingly, the C-terminal domain of DME has many lysine and arginine residues and both lysine and arginine are good substrates for protease II *(ptrB* in *S. meliloti*). This may explain the lack of detection of the fusion to the C-terminus of DME. Due to time constraints, further experiments to investigate the instability of the C-terminus of DME were not conducted.

Figure 7-4: Blots showing expression of DME fusions in (A) *S. meliloti* and (B) *E. coli*

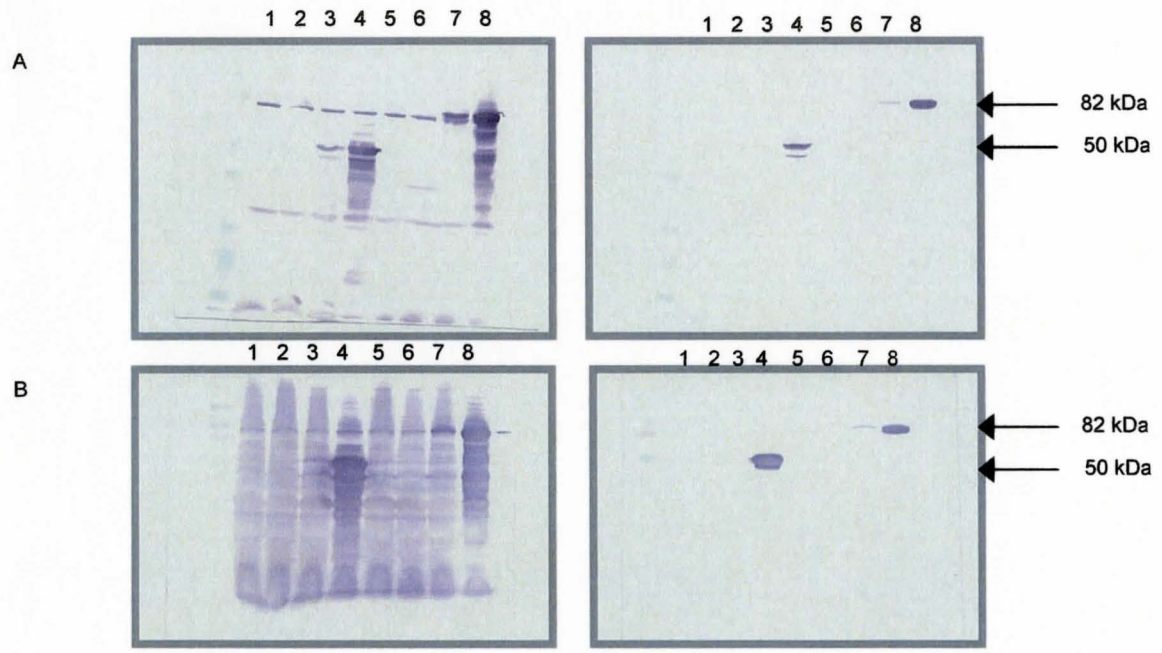


Figure 7-4: Western blots showing expression of DME fusions in (A) *S. meliloti* and (B) *E. coli*. Cultures were induced with 0.5mM IPTG for four hours, run on a 10% SDS-PAGE and blotted onto PVDF then incubated with either (Left) DME-specific antibodies or (Right) Strep-tag specific antibodies. The ladder used was benchmark protein ladder, which is not clearly visible in this photo. Lanes 1 and 2 contain pTrcSC induced and uninduced crude cell lysate. Lanes 3 and 4 contain the fusion to the N-terminus induced and uninduced (50 kDa). Lanes 5 and 6 contain the fusion to the C-terminus induced and uninduced (36 kDa). Lanes 7 and 8 contain the fusion to the full DME enzyme, induced and uninduced (82 kDa).

The biological activity of the *dme* fusions was tested using complementation of the *E. coli* malic enzyme mutant, EJ1321. This mutant is unable to grow using succinate as a sole carbon source unless it is complemented with a functional malic enzyme. Fusion plasmids, pTH1930 and pTH1981 containing the N-terminal fusion and the full *dme* fusion respectively were transformed into the EJ1321 background and the tranformants were patched onto M9-succinate with and without 0.5mM IPTG. In both cases, with and without IPTG, patches grew when the mutant *E. coli* strain was transformed with either the full *dme* fusion or the fusion to the N-terminus only. As a positive control, the *E. coli* strain J291 with harbours a plasmid (pTH452) carrying a truncated, functional version of the *dme* gene (Cowie, 1998) was grown in parallel. The negative controls consisted of the pTrcSC vector with no insert. The fusion to the C-terminal domain of the *dme* gene also did not complement the growth defect.

The fusion products to both the full *dme* gene and the 456 amino acid N-terminal region were purified according to the protocol outlined in the methods section and as seen in Figure 7-5, there are some proteins that co-purified with the fusions. Following purification of the fusion products, Bradford assays and malic enzyme assays were conducted to determine the amount and the activity of the purified protein. Malic enzyme assays (NAD+ reduction) were conducted using the methods described by (Brian Driscoll, 1995).

A negative control, pTrcSC with no insert was also purified using the same procedure as the full *dme* gene and the N-terminal region. Following passage of the cell lysate through the streptavidin column, several protein fragments were apparent following

SDS-PAGE analysis. However, once the aliquots were pooled and passed through the CBP column, there were no longer any protein fragments apparent (gel not shown); they were eluted with the un-bound proteins following CBP column washing.

Figure 7-5: Purified DME fusions on 10% SDS-PAGE

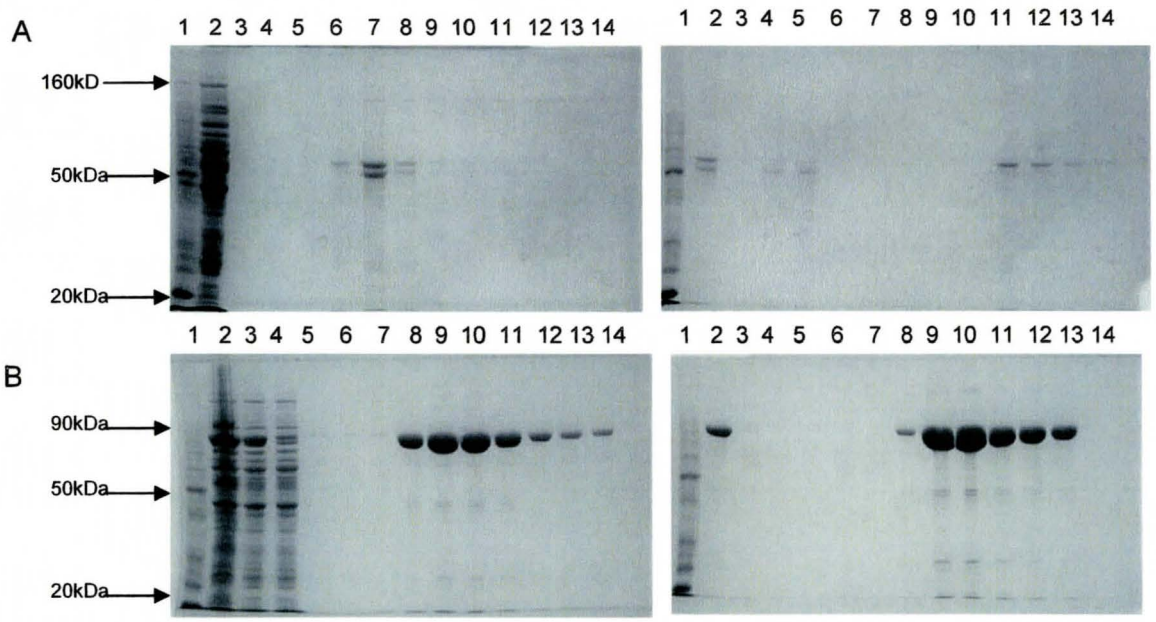


Figure 7-5: Purified fusions run on 10% SDS-PAGE. Gels in A) resulted from the purification of the strepII-CBP fusion to the N-terminal "malic enzyme domain" of DME. The gel on the left was from Strep-Tactin column elutions. Lane 1 is the benchmark ladder, Lane 2 is the lysate flow-through, lane 3 is the wash and lanes 4 through 14 are fractions. Fractions in lanes 6 through 9 were pooled and run on the CBP column. The gel on the right of A resulted from purification using the CBP column. Lane 1 is the benchmark ladder, lane 2 is the pooled strep-tactin column fractions, lane 3 is the wash and lanes 4 through 14 are the fractions. B shows gels from the purification of the fusion to the full *dme* gene. On the left is the gel from the strep purification, lane 1 is the benchmark ladder, lane 2 is the crude cell lysate, lane 3 is the lysate flowthrough, lane 4 is the wash, lanes 5 through 15 are the fractions. Fractions in lanes 8 through 11 were pooled and run on the CBP column. On the right is the gel for the CBP column. Lane one is the benchmark ladder, lane 2 is an aliquot from the strep purification, lane 3 is the wash and lanes 4 through 14 are the fractions.

The purified peptides and other co-eluting proteins were run on a gradient SDS-PAGE gel and subjected to western blotting using antibodies affinity-purified for the DME protein. All co-eluting proteins appeared to be interacting with the antibodies that were purified for the DME protein. Possible reasons for this interaction will be outlined in the discussion section of this thesis. It is important to note that these experiments were not continued due to time constraints and the success of the metabolic profiling.

CHAPTER 8

Discussion: protein interactions

The N-terminal domain of the DME enzyme is responsible for malic enzyme activity and the C-terminal domain shows some homology at the amino acid level to phosphotransacetylase (PTA) enzymes. Moreover, it has been shown that the PTA-like domain is not essential for DME enzyme activity and is not required for nitrogen fixation (Mitsch *et al* 1998). The function of this domain has remained unknown as attempts to detect PTA activity from the C-terminal domains have been unsuccessful; however, there is a high degree of conservation between the C-terminal domains of the malic enzymes of many bacteria (Mitsch *et al*, 1998). This conservation suggests that this unique extension possesses a function within the cell.

Bioinformatic analysis

Recently there has been an explosion of sequence data available for public use via the NCBI database. Previously, there were five malic enzymes known to possess the modular structure (DME, TME, *Rickettsia prowazekii*, *Haemophilus influenzae* and *E. coli* f759) (Mitsch *et al*, 1998). The tree constructed in this study consists of 50 malic enzyme sequences, 44 of which have the modular structure. Both the phylogenetic tree and the multiple sequence alignment (Figures 7-5 and 7-4) showed that the unique

modular structure possessed by the DME and TME malic enzymes of *S. meliloti,* is shared across a diverse spectrum of α, β and γ proteobacteria and is also present in some archaebacteria. The presence of this PTA-like extension in archaebacteria and its high level of conservation at the amino acid level across many bacterial species hints at functionality.

Comparison of conserved residues in the PTA-like domain of malic enzymes against conserved residues in PTA enzymes showed some interesting results. The fact that more than half of the strictly conserved residues of PTA enzymes (Table 7-1) are present in modular malic enzymes suggests that this PTA-like domain may have had activity at one point in time and because of the presence of another *pta* gene in the genome, may have lost this activity through evolution. In addition, the three residues that are not conserved between modular malic enzymes and PTA enzymes, K251, K277 and R304 were all found to interact with oxygen atoms of the acetyl phosphate molecules (Xu *et al*, 2005) whereas those that remain conserved seem to have primarily structural roles (Xu *et al*, 2005); perhaps the role of the PTA-like domain is to bind acetyl-CoA without converting it to acetyl phosphate.

Purification of DME fusions and co-eluting proteins

Expression and purification of the C-terminal-tagged DME constructs from pTrcSC was successful for fusion peptides of the full length DME and N-terminal domain (Figure 7-5). However, all proteins that co-eluted with these appeared to be interacting

with antibodies specific for the DME protein and so may have been aggregates or break-down products of DME. Nevertheless, these purified constructs could be used for pull-down assays in the future.

The tagged C-terminal domain was likely unstable, degraded, or insoluble. Degradation may have resulted because the C-terminal domain has many lysine and arginine residues which are good substrates for protease II *(ptrB* in *S. meliloti*). It would be beneficial to determine the nature of the instability of this construct to get a purified product that could be used for highthroughput screening for detection of potential enzyme activity. Purified peptide could also be used to conduct pull-down assays to determine the presence of interacting proteins.

Overexpression may not have been the best way to find a low abundance interacting protein as the high abundance of the target (DME) may lead to incorrect folding, protein aggregation and artifactual interactions. Repetition of this experiment would be better accomplished through the use of an integration vector, however, there are no proteins known to interact with malic enzymes and it is possible that none will be found.

Time constraints and the success of the metabolic profiling meant that this project was not pursued further.

REFERECES

Allen J, Davey HM, Broadhurst D, Heald JK, Rowland JJ, Oliver SG, Kell DB. (2003). High-throughput classification of yeast mutants for functional genomics using metabolic footprinting. Nat Biotechnol. **21**: 692-6

Andersson SG, Zomorodipour A, Andersson JO, Sicheritz-Ponetn T, Alsmark UC, Podowski RM, Naslaud AK, Eriksson AS, Winkler HH, Kurland CG. (1998). The genome sequence of Rickettsia prowazekii and the origin of mitochondria. Nature. **396(6707)**: 133-40

Bachrach U. (2005). Naturally occurring polyamines: interaction with macromolecules. Curr Protein Pept Sci. **6**: 559-66.

Baliga NS, Bonneau R, Facciotti MT, Pan M, Glusman G, Deutsch EW, Shannon P, Chiu Y, Wend RS, Gan RR, Hung P, Date SV, Marcotte E, Hodd L, Ng Wv. (2004). Genome sequence of Haloarcula marismortui: a halophilic archaeon from the Dead Sea. Genome Res. **14**: 2221-2234

Barsch A, Patschkowski T, Niehaus K. (2004). Comprehensive metabolite profiling of Sinorhizobium meliloti using gas chromatography-mass spectrometry. Funct Integr Genomics. **4**: 219-30.

Bartolucci S. Rella R, Guagliardi A, Raia CA, Gambacorta A, DeRosa M and Rossi M. (1987). Malic enzyme from *Sulfolobus solftaricus:* purification, structure and kinetic properties. J. Biol. Chem. **262:** 7725-7731.

Blumenthal DK, Edelman AM, Charbonneau H, Titani K, Walsh KA, Krebs EG. (1985). Identification of the calmodulin-binding domain of skeletal muscle myosin light chain kinase. Proc Natl Acad Sci U S A. 82: 3187-91.

Bundy JG, Willey TL, Castell RS, Ellar DJ, Brindle Km. (2005). Discrimination of pathogenic clinical isolates and laboratory strains of Bacillus cereus by NMR-based metabolomic profiling. FEMS Microbiol Lett. **242**: 127-36.

Chen F, Okabe Y, Osano K and Tajima S. (1997). Purification and characterization of the NADP-malic enzyme from *Bradyrhizobium japonicum* A1017. Biosci. Biotech. Biochem. **61:** 384-386.

Copeland L, Vella J, Hong Z. (1988). Enzymes of carbohydrate metabolism in soybean nodules. Phytochemistry. **28:** 57-61.

Cowie A. (1998). The roles of malic enzymes of *rhizobium (Sinorhizobium) meliloti* in symbiotic nitrogen fixation. Thesis (M.Sc.). McMaster University.

Cowie A, Cheng J, Sibley CD, Fong Y, Zaheer R, Patten CL, Morton RA, Golding BG, Finan TM. (2006). An integrated approach to functional genomics: construction of a novel reporter gene fusion library of *Sinorhizobium meliloti.* Appl. Environ. Microbiol. (Submitted)

Csonka LN and Hanson AD. (1991). Prokaryotic osmoregulation: genetics and physiology. Annu Rev Microbiol. **45**: 569-606

D'Agostino L, di Pietro M, Di Luccia A. (2006). Nuclear aggregates of polyamines. IUBMB Life. **58**: 75-82.

Driscoll BT. (1995). Genetic and biochemical characterization of the NAD+ and NADP+ malic enzymes of *Rhizobium meliloti.* Thesis (Ph.D.). McMaster University.

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

Driscoll BT, Finan TM. (1996). NADP+ -dependent malic enzyme of Rhizobium meliloti. J Bacteriol. **178**: 2224-31.

Driscoll BT, Finan TM. (1997). Properties of NAD(+)- and NADP(+)-dependent malic enzymes of Rhizobium (Sinorhizobium) meliloti and differential expression of their genes in nitrogen-fixing bacteroids. Microbiology. **143**: 489-98.

Fiehn O, Kopka J, Trethewey RN, Willmitzer L. (2000). Identification of uncommon plant metabolites based on calculation of elemental compositions using gas chromatography and quadrupole mass spectrometry. Anal Chem. **72**: 3573-80

Finan TM, Oresnik I, Bottacin A. (1988). Mutants of Rhizobium meliloti defective in succinate metabolism. J Bacteriol. **170**: 3396-403

Finan TM, McWinnie E, Driscoll B, Watson RJ. (1991). Complex symbiotic phenotypes result from gluconeogenic mutations in *Rhizobium meliloti.* Mol. Plant-Microbe Interact. **4:** 386-392.

Fischer HM. (1994). Genetic regulation of nitrogen fixation in rhizobia. Microbiol Rev. **58**: 352-86

Fujihara S, Harada Y. (1989). Fast-growing root nodule bacteria produce a novel polyamine, aminobutylhomospermidine. Biochem Biophys Res Commun. **165**: 659-66.

Frenkel. (1975). Regulation and physiological functions of malic enzymes. Curr. Top. Cell. Regul. **9:** 157-181.

Fujihara S, Yoneyama T. (1993). Effects of pH and Osmotic Stress on Cellular Polyamine Contents in the Soybean Rhizobia Rhizobium fredii P220 and Bradyrhizobium japonicum A1017. Appl Environ Microbiol. **59**: 1104-1109.

Fukuda W, Ismail YS, Fukui T, Atomi H and Imanaka T. (2005). Characterization of an archaeal malic enzyme from the hyperthermophilic archaeon Thermococcus kodakaraensis KOD1. Archaea. **1:** 293-301.

Garrido-Pertierra A, Martinez-Marcos C, Martin-Fernandez M and Ruiz-Amil M. (1983). Properties and functions of malic enzyme from *Psuedomonas putida.* Biochemie. **65:** 629-635.

German JB, Hammock BD, Watkins SM. (2006). Metabolomics: building on a century of biochemistry to guide human health. Metabolomics. **1**: 3-9

Gouffi K, Pichereau V, Rolland JP, Thomas D, Bernard T, Blanco C. (1998). Sucrose is a nonaccumulated osmoprotectant in Sinorhizobium meliloti. J Bacteriol. **180**: 5044-51.

Gouffi K, Pica N, Pichereau V. Blanco C. (1999) Disaccharides as a new class of nonaccumulated osmoprotectants for Sinorhizobium meliloti.Appl Environ Microbiol. **65**: 1491-500.

Griffin JL. (2006). Understanding mouse models of disease through metabolomics. Curr Opin Chem Biol. Jun 30; [Epub ahead of print].

Hahm DH, Pan J, Rhee JS. (1994). Characterization and evaluation of a pta (phosphotransacetylase) negative mutant of Escherichia coli HB101 as production host of foreign lipase. Appl Microbiol Biotechnol. **42**: 100-7

Hecker M, Volker U. (2001). General stress response of Bacillus subtilis and other bacteria. Adv Microb Physiol. **44**: 35-91.

Huelsenbeck JP, Ronquist F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics. **17**: 754-5.

Iyer PP, Lawrence SH, Luther KB, Rajashankar KR, Yennawar HP, Ferry JG, Schindelin H. (2004). Crystal structure of phosphotransacetylase from the methanogenic archaeon Methanosarcina thermophila. Structure. **12**: 559-67

Jeanmougin F, Thompson JD, Gouy M, Higgins DG, Gibson TJ. (1998). Multiple sequence alignment with Clustal X. Trends Biochem Sci. **23**: 403-5

Kaasen I, McDougall J, Strom AR. (1991). Analysis of the otsBA operon for osmoregulatory trehalose synthesis in Escherichia coli and homology of the OtsA and OtsB proteins to the yeast trehalose-6-phosphate synthase/phosphatase complex. Gene. **145**: 9-15.

Katsuki I, Takeo K, Kameda K and Tanaka S. (1967). Existence of two malic enzymes in *Esherichia coli.* Biochem. Biophys. Res. Comm. **27:** 331-336.

Kawai S, Suzuki H, Yamamoto K, Inui M, Yukawa H, Kumagai H. (1996). Purification and characterization of a malic enzyme from the ruminal bacterium *Streptococcus bovis* ATCC 15352 and cloning and sequencing of its gene. Appl. and Environ Microbiol. **62:** 2692-2700.

Keele BB Jr, Hamilton PB, Elkan GH. (1969). Glucose catabolism in Rhizobium japonicum. J Bacteriol. **97**:1184-91.

Kobayashi K, Doi S, Negoro S, Urabe I and Okada H. (1989). Structure and properties of malic enzyme from *Bacillus stearothermophilus.* J. Biol. Chem. **264**: 3200-3250.

Korndorfer IP, Skerra A. 2002. Improved affinity of engineered streptavidin for the Strep-tag II peptide is due to a fixed open conformation of the lid-like loop at the binding. Protein Sci. **11**: 883-93.

Kouchi H, Kukai K, Katagiri H, Minamisawa K and Tajima S. (1998). Isolation and enzymological characterization of infected and uninfected cell protoplasts from root nodules of *Glycine max*. Physiologia Plantarum. **73:** 327-334.

Kunst F, Ogasawara N, Moszer I, Albertini AM, Alloni G, Azevedo V, Bertero MG, Bessieres P, Bolotin A, Borchert S, Borriss R, Boursier L, Brans A, Braun M, Brignekk SC, Bron S, Brouillet S, Bruschi CV, Caldwell B, Capuano V, Carter NM, Choi SKm Codani JJ, Connerton IF, Danchin A *et al.* (1997). The complete genome sequence of the gram-positive bacterium Bacillus subtilis. Nature. **390(6657):** 249-56

Lahiri K, Chattopadhyay S, Ghosh B. (2004). Correlation of endogenous free polyamine levels with root nodule senescence in different genotypes in Vigna mungo L. J Plant Physiol. **161**: 563-71.

Lamed R and Zeikus JG. (1981). Thermostable, ammonium-activated malic enzyme of *Clostridium thermocellum.* Biochem. Biophys. Acta. **660:** 251-255.

Loeber G, Infante AA, Maurer-Fogy I, Krystek E and Dworkin MB. (1991). Human NAD+-dependent mitochondrial malic enzyme. J. Biol. Chem. **266:** 3016-3021.

Long SR. (1989). Rhizobium-legume nodulation: life together in the underground. Cell. 56: 203-14.

Ma XJ, Salati LM, Ash SE, Mitchell DA, Klautky SA, Fantozzi DA and Goodridge AG. (1990). Nutritional regulation and tissue-specific expression of the malic enzyme gene in chicken. J. Biol. Chem. **265**: 18435-18441.

Mahajan SK, Chu CC, Willis DK, Templin A and Clark AJ. (1990). Physical analysis of spontaneous and mutagen-induced mutants of *Escherichia coli* K12 expressing DNA exonuclease VIII activity. Genet. **125**: 261-273.

Malmendal A, Overgaard J, Bundy JG, Sorensen JG, Nielsen NC, Loeschcke V, Holmstrup M. (2006). Metabolomic profiling of heat stress: hardening and recovery of homeostasis in Drosophila. Am J Physiol Regul Integr Comp Physiol. **291**: R205-12

McKay IA, Dilworth MJ and Glenn AR. (1988). C_4-dicarboxylate metabolism in free-living and bacteroid forms of *Rhizobium leguminosarum* MNF3841. J Gen. Microbiol. **134**: 1433-1440.

Miller Kj, Wood JM (1996). Osmoadaptation by rhizosphere bacteria. Annu Rev Microbiol. **50**: 101-36.

Mitsch MJ, Voegele RT, Cowie A, Osteras M, Finan TM. (1998). Chimeric structure of the NAD(P)+- and NADP+-dependent malic enzymes of Rhizobium (Sinorhizobium) meliloti. J Biol Chem. **273**: 9330-6.

Moreel K, Goeminne G, Storme V, Sterck L, Ralph J. Coppieters W, Breyne P, Steenackers M, Georges M, Messens E and Boerjan W. (2006). Genetical metabolomics of flavonoid biosynthesis in Populus: a case study. Plant J. Jun 15; [Epub ahead of print].

Mukhopadhyay A, He Z, Alm EJ, Arkin AP, Baidoo EE, Borglin SC, Chen W, Hazen TC, He Q, Holman HY, Huang K, Huang R, Joyner DC, Katz N, Keller M, Oeller P,

Redding A, Sun J, Wall J, Wei J, Yang Z, Yen HC, Zhou J, Keasling JD. (2006). Salt stress in Desulfovibrio vulgaris Hildenborough: an integrated genomics approach . J Bacteriol. **188**: 4068-78

Munro GF, Hercules K, Morgan J, Sauerbier W. (1972). Dependence of the putrescine content of Escherichia coli on the osmotic strength of the medium. J Biol Chem. **247**: 1272-80.

Murai T, Tokushige M Nagai J and Katsuki H. (1971). Physiological functions of NAD and NADP-linked malic enzymes in *Escherichia coli.* Biochem. Biophys. Res. Comm. **43:** 875-881.

Ochoa S., Mehler A. Kornberg A. (1948). Reversible oxidative decarboxylation of malic acid. J. Biol. Chem. 871-872.

Pichereau V, Pocard JA, Pocard JA, Hamelin J, Blanco C, Blanco C, (1998). Differential Effects of Dimethylsulfoniopropionate, Dimethylsulfonioacetate, and Other S-Methylated Compounds on the Growth of Sinorhizobium meliloti at Low and High Osmolarities. Appl Environ Microbiol. **64**: 1420-1429.

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

Rigaut G, Shevchenko A, Rutz B, Wilm M, Mann M, Seraphin B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. Nat Biotechnol. **17**: 1030-2.

Roberts BG, Brill WJ. (1981). Genetics and regulation of nitrogen fixation. Ann. Rev. Microbiol. **35**: 297-319.

Rochfort, S. (2005). Metabolomics reviewed: a new "omics" platform technology for systems biology and implications for natural products research. J Nat Prod. **68**: 1813-20

Sanwal BD. (1970). Allosteric controls of amphibolic pathways in bacteria. Bacteriol. Rev. **34:** 20-39.

Sanwal BD. (1970b). Regulatory characteristics of the diphosphopyridine nucleotide specific malic enzyme of *Escherichia coli*. J. Biol. Chem. **245:** 1212-1216.

Sanwal BD and Smando R. (1969). Malic enzyme of *Escherichia coli:* diversity of the effectors controlling enzyme activity. J. Biol. Chem. **244:** 1817-1823.

Sanwal BD, Wright JA and Smando R. (1968). Allosteric control of the activity of malic enzyme in *Escherichia coli.* Biochem. Biophys. Res. Comm. **31**: 623-627.

Schmidt TGM and Skerra A. (1993). The random peptide library-assisted engineering of a C-terminal affinity peptide, useful for the detection and purification of a functional Ig FV fragment. Protein Eng. **6**: 109-122.

Schmidt-Eisenlohr HDomke N, Baron C. (1999). TraC of IncN Plasmid pKM101 Associates with Membranes and Extracellular High-Molecular-Weight Structures in *Escherichia coli.* J Bacteriol. **181**: 5563–5571.

Skorkowski EF, Storey KB. (1990). Regulation of coenzyme utilization by mitochondrial NAD(P)+-dependent malic enzyme. Int. I. Biochem. **22**: 471-475.

Stein, SE. (1999). An integrated method for spectrum extraction and compound identification from GC/MS data. Journal of the American Society for Mass Spectrometry, **10**: 770.

Suye SI, Okada Y, Funanda A, Kawagoe M and Inuta S. (1992). Purification and properties of malic enzyme from *Pseudomonas diminuta* IFO-13182. J. Ferm. and Bioeng. **73:** 343-347.

Tabor CW, Tabor H. (1984). Polyamines. Annu Rev Biochem. **53**: 749-90.

Tabor CW, Tabor H. (1999). It all started on a streetcar in Boston. Annu Rev Biochem. **68**: 1-32.

Tabrett CA and Copeland L. (2002). Enzymes of malate metabolism in Mesorhizobium ciceri CC 1192. **:** Can J Microbiol. **48**: 279-84

Terpe K. (2003). Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. Appl Microbiol Biotechnol. **60**: 523-33.

Vance, C.P.and Heichel GH. (1998). Carbon in N_2-fixation: limitation or exquisite adaptation. Ann. Rev. Plant Physiol. Plant Mol. Biol. **42:** 373-392.

Villas-Boas SG, Noel S, Lane GA, Attwood G, Cookson A. (2006). Extracellular metabolomics: a metabolic footprinting approach to assess fiber degradation in complex media. Anal Biochem. **349**: 297-305

Vasse J, deBilly F, Camut S, Truchet G. (1990). Correlation between ultrastructural differentiation of bacteroids and nitrogen fixation in alfalfa nodules. J. Bacteriol. **172**: 4295-4306.

Voegele RT, Mitsch MJ, Finan TM. (1999). Characterization of two members of a novel malic enzyme class. Biochim Biophys Acta. **1432:** 275-85

Wedding RT. (1989). Malic enzymes of higher plants: characteristics, regulation and physiological function. Plant Physiol. **90:** 367-371.

Xu QS, Jancarik J, Lou Y, Kuznetsova K Yakunin AF, Yokota H, Adams P, Kim R, Kim SH. (2005). Crystal structures of a phosphotransacetylase from Bacillus subtilis and its complex with acetyl phosphate. J Struct Funct Genomics. **6**: 269-79.

Yamaguchi M, Tokushige M and Katsuki H. (1973). Studies on regulatory functions of malic enzymes II purification and molecular properties of nicotinamide adenine dinucleotide-linked malice enzymes from *Escherichia coli.* J. Biochem. **73:** 169-180.

Yamamoto S, Yamasaki K, Takashina K, Katsu T, Shinoda S. (1989). Characterization of putrescine production in nongrowing Vibrio parahaemolyticus cells in response to external osmolality. Microbiol Immunol. **33**: 11-21.

Yarosh OK, Charles TC, Finan TM. (1989). Analysis of C4-dicarboxylate transport genes in Rhizobium meliloti. Mol. Microbiol. **3:** 813-823.

APPENDIX I

Quantification of malate and fumarate

In order to assess the amount of malate and fumarate that was excreted into the extracellular environment by *dme* and *tme* mutants (Table 3-2), 1μL of each 5mM; 10mM and 20mM malate and fumarate (in triplicate) was derivatized and run on the GC-MS. A standard curve was constructed from the relative response factors of malate and fumarate at each concentration (Figures A-1 and A-2). The concentrations of malate and fumarate in the *dme* and *tme* samples were calculated using the equation of the line derived from the standard curve. Note that intracellular samples were concentrated 300 times and extracellular samples 2 times and that these concentration steps were accounted for in the calculations.

Figure A-1: Standard curve for malate

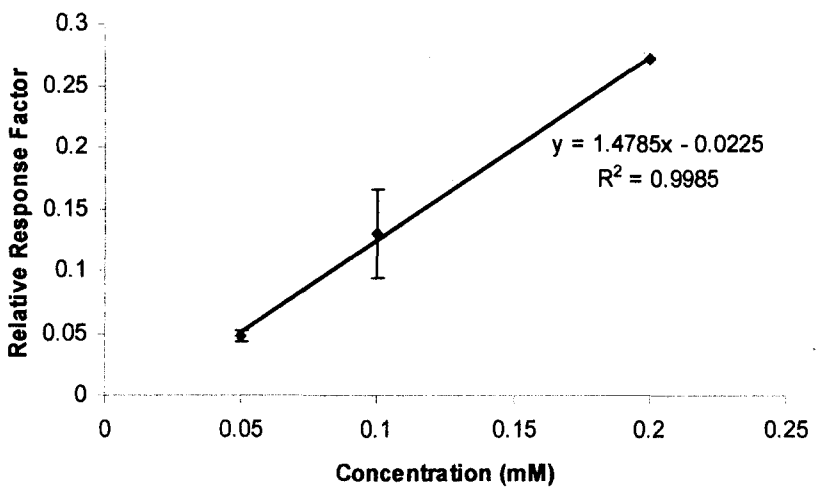
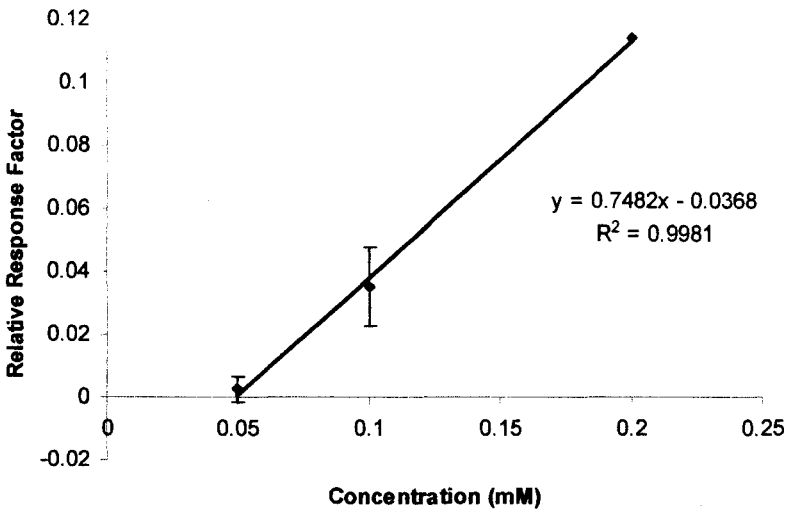


Figure A-2: Standard curve for fumarate



APPENDIX II

Verification of compound identity (putrescine and trehalose)

As previously discussed in Materials and Methods, the AMDIS program provides a mass spectral matching service operating in conjunction with the NIST library program. However, matches are not always correct as some compounds have very similar mass spectral patterns at low concentrations because not all of the molecular ions will be detected thus leading to discrepancies in compound identification. To overcome this problem and to validate the identity of a compound, co-injection with a known amount of a given compound may be performed (Materials and Methods).

In this work, co-injection was performed to confirm the identity of putrescine (Figure A-3). The identity of trehalose was confirmed through the matching of mass spectral patterns obtained in this study to those previously reported for trehalose.

Figure A-3: Gas chromatograms of putrescine co-injection

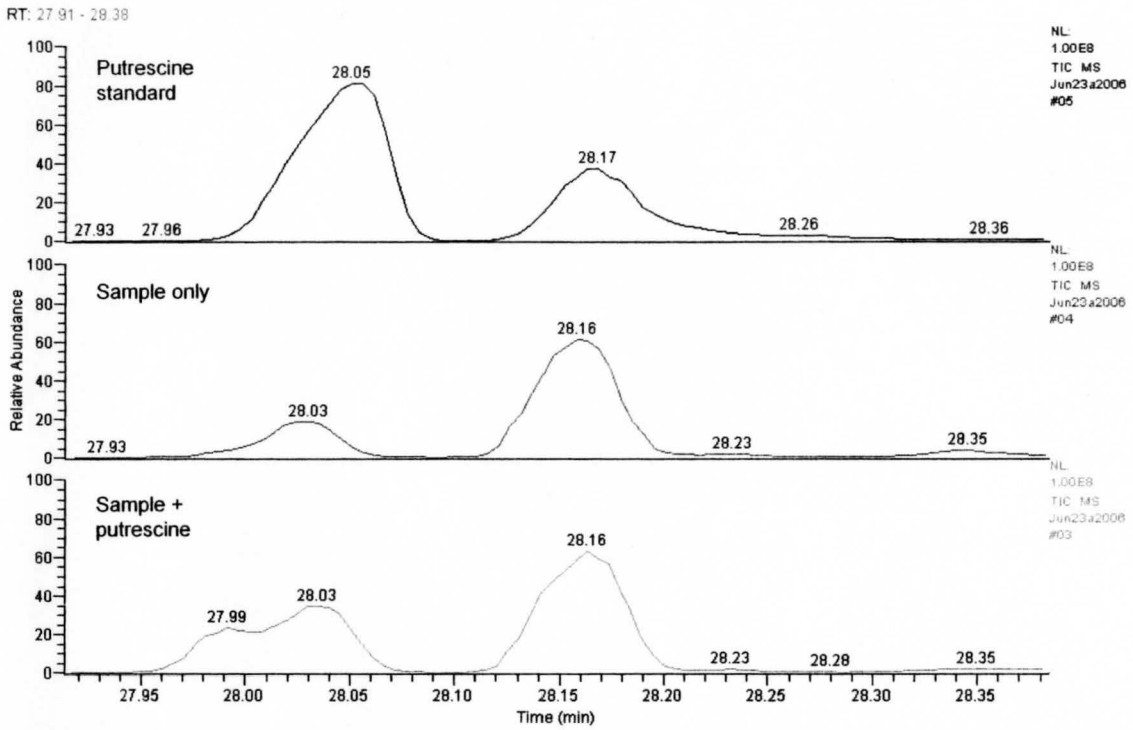


Figure A-3: Gas chromatograms for putrescine Co-injection experiment (into extracts derived from *tme* mutant cells grown in M9-succinate). Images were taken from the Xcalibur program's Qual Browser function.

Results from the co-injection experiment with putrescine were unusual as two peaks were formed at retention times (RT) 27.99 and 28.05 minutes when putrescine was co-injected into the sample derived from *tme* cells grown in M9-succinate. However, the peak at RT 27.99 minutes was not present on the chromatogram when the putrescine standard was derivatized and run. It is important to note that all of these samples were run on the same day. This co-injection experiment was repeated and the results were similar. The cause for this splitting is unclear, however it may be hypothesized that the peak at retention time 28.03 minutes is putrescine because the peak at this retention time

increased in height between blank and co-injected samples; the putrescine standard had one peak at retention time 28.03 minutes; and the mass spectral patterns for the standard putrescine sample and the putrescine contained in the *tme* mutant-derived sample were a good match (Figure A-4). Note that some of the molecular ions that are less abundant do not appear on both the patterns for the putrescine standard and the putrescine derived from the sample; these were seen through magnification of the spectra using AMDIS.

Figure A-4: Mass spectral patterns of putrescine co-injection

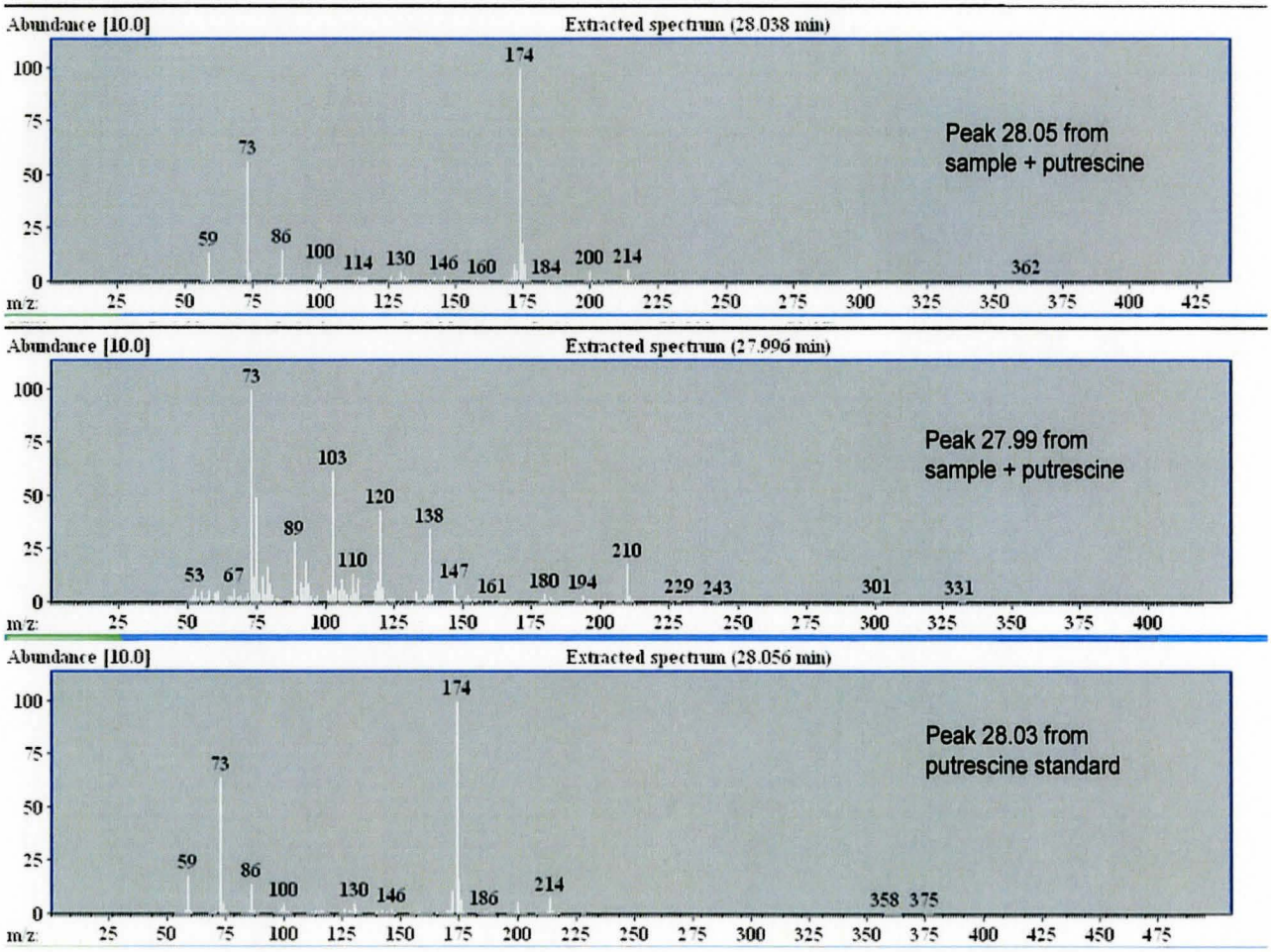


Figure A-4: Mass spectral patterns for putrescine Co-injection experiment (into extracts derived from *tme* mutant cells grown in M9-succinate). Images were taken from the AMDIS program.

Trehalose has a retention time which is identical to that of maltose, generally around 46 minutes. In addition, the mass spectral patterns for trehalose and maltose are nearly identical, which is not surprising as both of these molecules are disaccharides consisting of two glucose molecules. The difference between maltose and trehalose is that the maltose molecule consists of two glucose residues joined by an α(1→4) linkage

whereas in trehalose the two glucose molecules are joined by an α(1→1) linkage. The above-mentioned similarities make it difficult to differentiate between these two molecules using GC-MS. Vogel *et al* (2001) published a spectrum for trehalose which had characteristic molecular ions at high m/z ratios: 435, 451 and 463, which were not present in a maltose standard that was run in this study but were present in the peak at retention time 46 minutes in the *dme* mutant extract. Also, maltose is known to possess a molecular ion with an m/z ratio of 480 (Dave Guevara, unpublished data), which the peaks in the extracts of *dme* mutant cells and the trehalose standard did not possess. From this information it may be hypothesized that the peak at the retention time of approximately 46 minutes in the *dme* mutant extracts consists of trehalose and not maltose.

APPENDIX III

Tables A-1 and A-2 shown below list all metabolites that can be identified on the polar extracts of the wild type strain, Rm1021 grown in M9-succinate and M9-glucose, respectively. Note that the identity of these compounds has not been verified by co-injection. Figure A-5 shown below is a Venn diagram of the compounds in M9-glucose versus M9- succinate.

Table A-1: Metabolites present in polar extracts of Rm1021 cells grown in M9-succinate

Metabolite	Retention time	RRF
succinic acid	17.97	4.480
uridine	18.60	0.062
fumaric acid	18.77	0.906
L-aspartic acid	20.78	0.179
2-methyl malic acid	20.97	0.035
malic acid	22.58	3.808
L-pyroglutaric acid	23.34	3.729
L-glycine	23.50	3.670
A-ketoglutaric acid	24.64	0.453
L-glutamic acid	25.66	5.714
5-hydroxy-L-tryptophan	26.81	0.072
L-xylose	27.19	0.176
2-amino-adipinic acid	27.73	0.027
putrescine	28.07	0.114
3-phosphoglycerate	29.93	0.144
citric acid	30.10	0.279
adenine	30.89	0.223
glutamine	31.14	0.009
L-norvaline	31.50	0.072
L-lysine	31.96	0.138
D-ribose	34.98	0.040
ribose-5-phosphate	35.96	0.041
spermine	37.00	0.068
fructose-6-phosphate	39.49	0.139
mannose-6-phosphate	39.70	0.226
glucose-6-phosphate	39.98	0.024
D-glucuronic acid	43.64	0.010
andenosine	44.30	0.166
trehalose	26.40	0.035

Table A-2: Metabolites present in polar extracts of Rm1021 cells grown in M9-glucose

Metabolite	Retention time	RRF
succinic acid	17.60	0.151
fumaric acid	18.41	0.053
malic acid	22.15	0.282
L-pyroglutaric acid	23.02	2.048
L-glycine	23.14	2.643
L-glutamic acid	25.23	5.865
putrescine	27.58	0.212
2-phosphoglycerate	29.03	0.009
ornithine	29.44	0.221
3-phosphoglycerate	29.56	1.095
citric acid	29.68	0.047
D-glucose	31.22	0.992
D-mannose	31.44	20.16
threonic acid	32.70	0.607
talose	33.03	2.246
D-gluconic acid	33.40	7.920
spermine	37.55	0.081
fructose-6-phosphate	39.02	0.689
mannose-6-phosphate	39.23	1.460
glucose-6-phosphate	39.51	0.024
6-phosphogluconic acid	41.54	0.030
sorbitol-6-phosphate	41.78	1.444
L-leucine	42.57	0.018
D-glucuronic acid	43.64	0.152
andenosine	43.91	0.318

Figure A-5: Venn diagram of compounds in M9-glucose versus M9-succinate

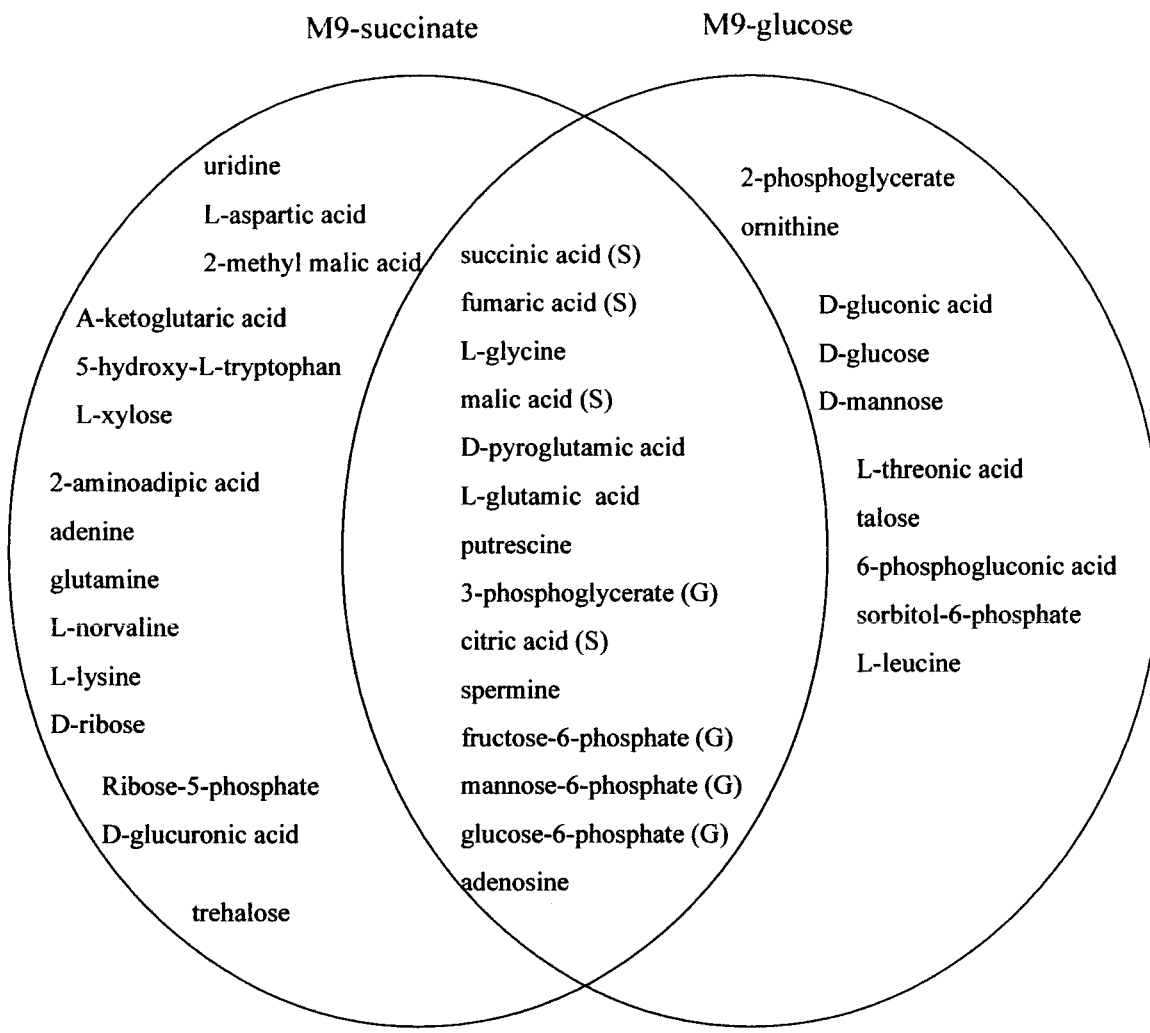


Figure A-5: Venn diagram of metabolites from Rm1021 cultures grown in M9-glucose versus those grown in M9-succinate. G and S represent metabolites in the common space which had higher RRF values in M9-glucose and M9-succinate respectively.

Note from Figure A-5 that many of the compounds found only (or at a higher RRF) in the extracts of cells grown in M9-succinate are TCA cycle intermediates and amino acids. Those found only (or at a higher RRF) in the extracts of cells grown in M-

glucose are mostly intermediates of the Entner Doudoroff pathway or derivatives of glucose. Table A-3 below lists all compounds that were found in cells grown in M9-glucose, succinate, glycerol or arabinose. Note that these metabolites were identified with the NIST database and not with purified standards.

Table A-3: Metabolites identified by GC-MS in the polar extracts of *S. meliloti* grown in M9-succinate, glucose, glycerol or arabinose

Amino acids	Sugars	Organic acids	N-containing compounds	Other
Aspartic acid	fructose	2PG	2-aminoadipic acid	adenine
Glutamic acid	F6P	3PG	putrecine	adenosine
Glutamine	glucose	citric acid	spermine	uracil
Glycine	G6P	fumaric acid		
Leucine	mannose	malic acid		
Lysine	M6P	succinic acid		
Ornithine	ribose	α-ketoglutaric acid		
Serine	Ribose5P	2-methylmalic acid		
Threonic acid	talose	gluconic acid		
Tyrosine	trehalose	glucuronic acid		
Pyroglutamic	ribose	6-phosphogluronic acid		
5-hydroxy tryptophan	Ribulose5P			
	Lyxose			
	Sorbitol-6-phosphate			

* Note that in Ribose5P and Ribulose5P, P represents phosphate

APPENDIX IV

Analysis of the effect of malic enzyme mutations on the expression of the genes of central carbon metabolism

We hypothesized that the accumulation of glucose-6-phosphate and fructose-6-phosphate in the *dme* mutant may be resulting from disturbances in the central carbon metabolism of the organism. To test this hypothesis, expression of central carbon metabolism genes was determined by analysis using either *lacZ* or *gusA* reporter fusion clones from an *S. meliloti* genomic DNA library (Cowie *et al,* 2006). The fusion plasmids (non-replicating) were mated into Rm1021 and the three mutant backgrounds (*dme3*, *tme4* and *dme3tme4*) and gene expression was tested in four media types: M9-succinate, M9-glucose, M9-glycerol and M9-glucose and succinate.

From this experiment, we observed that there was little difference in gene expression between the mutant backgrounds despite differences in media type (Figure 3-15). However, since there are still many unanswered questions in the area of *S. meliloti* central carbon metabolism, included here are graphs of *lacZ* or *gusA* expression (Figures A-6 and A-7) in an Rm1021 background.

Figure A-6: β-galactosidase specific activity of central carbon metabolism genes in M9-succinate, M9-glucose and M9-glycerol

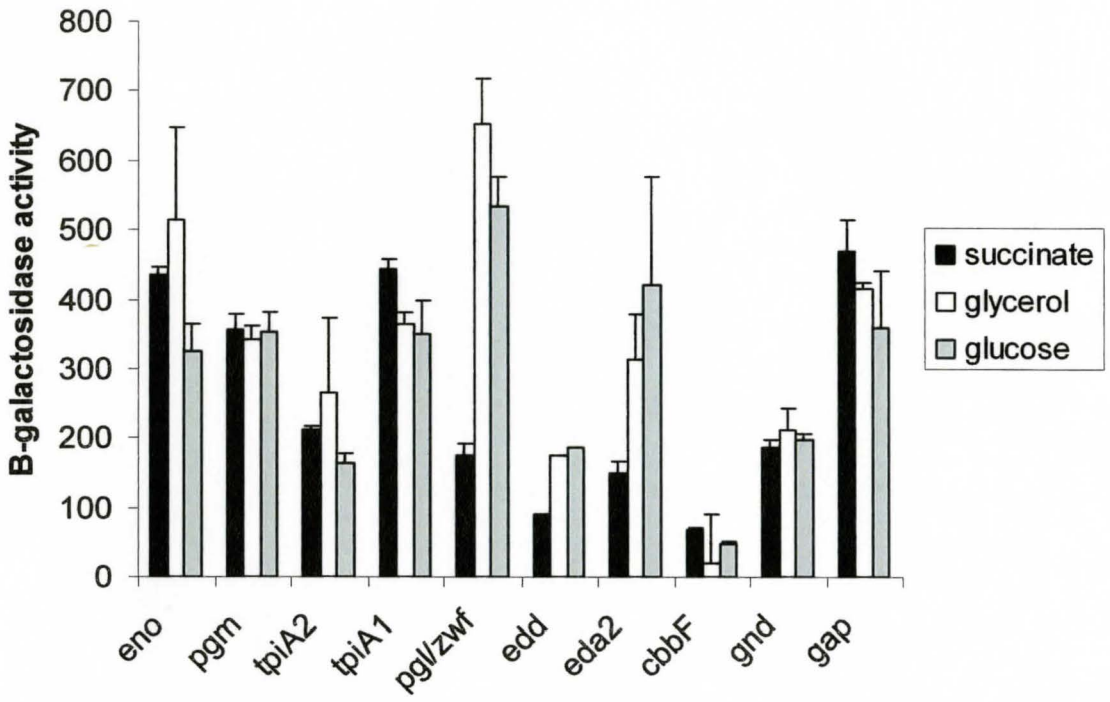


Figure A-6: Graph of *lacZ* gene expression data for selected carbon metabolism genes where strains were grown in M9-succinate, M9-glucose and M9-glycerol. Assays were conducted in triplicate and the error bars were calculated from the standard deviation between the three replicates. Note that *edd* does not have an error bar, this is because there was only one replicate that grew in all media types. Error bars were calculated using standard deviation where N=3.

Figure A-7: β-glucuronidase specific activity of central carbon metabolism genes in M9-succinate, M9-glucose and M9-glycerol

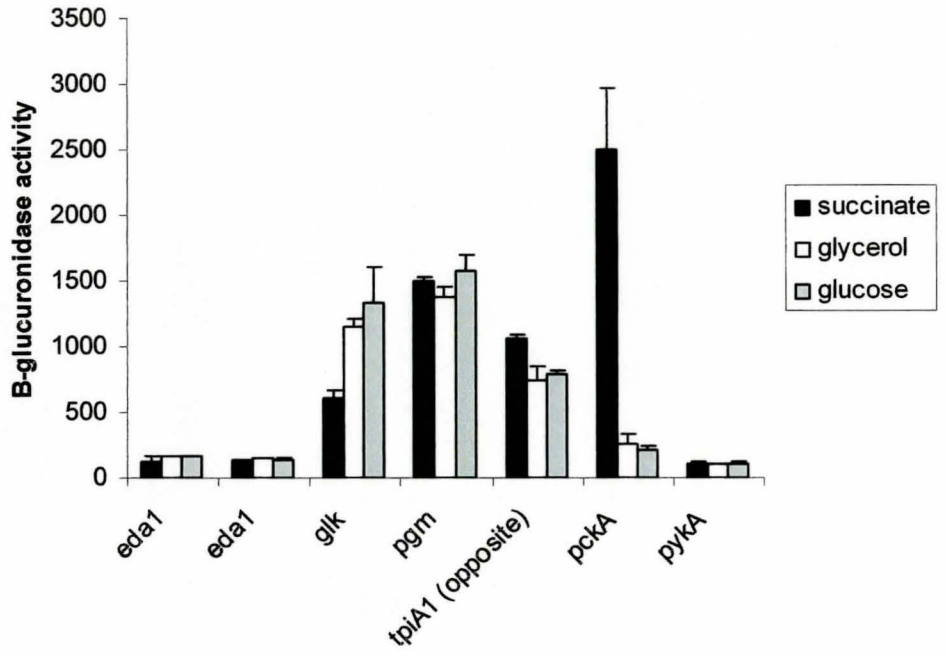


Figure A-7: Graph of *gusA* gene expression data for selected carbon metabolism genes where strains were grown in M9-succinate, M9-glucose and M9-glycerol. Assays were conducted in triplicate and the error bars were calculated from the standard deviation between the three replicates. Error bars were calculated using standard deviation where N=3.

The data in Figures A-6 and A-7 shows that there are few genes that could be re-examined in future work. These include the fusion to both genes encoding 6-phosphogluconolactonase (*pgl*) and glucose-6-phosphate 1-dehydrogenase (*zwf*) which showed 3-fold greater expression in glucose and glycerol versus succinate; *eda2* which encodes 2-dehydro-3-deoxyphosphogluconate aldolase exhibited 2-fold greater expression in glucose and glycerol versus succinate and *glk* which encodes a transmembrane glucokinase was expressed in glucose and glycerol to levels 3-fold greater

than in succinate. Finally, the fusion to the gene encoding phosphogluconate dehydratase, *edd,* should also be re-tested in triplicate as only one replicate grew in this study.

APPENDIX V

Cloning of new fusions using the library vector, pTH1703

Five central carbon metabolism genes which did not have a corresponding library clone, were PCR-amplified and cloned into the vector pTH1703, creating fusions to the *lacZ* and *gfp+* genes. Table A-4 below gives the cloned regions for each gene.

Table A-4: Regions cloned into library vector, pTH1703

Gene name	resulting plasmid	start nucleotide #	stop nucleotide #
edd (-)	pTH2068	767274	766636
eda2 (-)	pTH2069	3111686	3111278
cbbF (-)	pTH2070	212767	212976
gnd (-)	pTH2081	2091728	2091007
gap (+)	pTH2082	2975907	2976475

NB: -/+ next to gene name is indicative of bottom or top strand respectively. Nucleotide numbers were taken from the full genome sequence for Rm1021, NC_003047.

APPENDIX VI

Accession numbers for malic enzyme protein sequences used in Figure 7-2

In order to assess the diversity of malic enzymes that possess the modular structure, a phylogenetic tree was constructed (Figure 7-2). Due to lack of space on the tree diagram, accession numbers for each sequence included are listed below in Table A-5.

Table A-5: Accession numbers for sequences included in Figure 7-2

Species	Accession number
Halobacterium sp	NP_280405
Haloarcula marismortui	AAV46652
Pseudomonas aeruginosa	NP_253733
Rhodobacter sphaeroides	ZP_00004636
Silicibacter pomeroyi	AAV96173
Bradyrhizobium japonicum	NP_770785
Rhodopseudomonas palustris	NP_948381
Bartonella henselae	YP_033559
Bartonella Quintana	YP_032194
Sinorhizobium meliloti (DME)	CAC46416
Agrobacterium tumefaciens	NP_354642
Brucella melitensis (NAD)	AAL52148
Brucella abortus	YP_221733
Brucella suis	AAN29939
Azoarcus sp	YP_159319
Bordetella pertussis	NP_879849
Bordetella parapertussis	NP_885389
Bordetella bronchiseptica	NP_887475
Burkholderia pseudomallei	YP_109836
Burkholderia fungorum	ZP_00277183
Ralstonia eutropha	ZP_00168277
Ralstonia solanacearum	CAD16474
Azoarcus sp	YP_159517
Bordetella parapertussis	NP_883476
Polaromonas sp	ZP_00361351
Ralstonia eutropha	ZP_00168692
Ralstonia metallidurans	ZP_00275268
Burkholderia pseudomallei	YP_109553
Dechloromonas aromatica	ZP_00152505
Xanthomonas campestris	AAT88066
Haemophilus influenzae	NP_439401
Chromobacterium violaceum	AAQ58590
Photorhabdus luminescens	NP_929953
Erwinia carotovora	YP_048980
Yersinia pestis	NP_668769
Yersinia pseudotuberculosis	YP_071263
Salmonella enterica	NP_804256
Escherichia coli CFT073	NP_754870
Escherichia coli O157:H7	BAB36748
Escherichia coli K12	NP_416958
Shigella flexneri	NP_708302
Wolbachia	NP_966271
Rickettsia prowazekii	NP_220756
Sinorhizobium meliloti (TME)	AAB82460
Brucella melitensis (NADP)	AAL52983
Campylobacter coli	ZP_00367336
Bacillus cereus	AAP11511
Clostridium thermocellum	ZP_00312074
Clostridium tetani	NP_781056
Thermobifida fusca	ZP_00293990