TRANSPORT AND METABOLISM OF B-GLYCOSIDIC SUGARS IN *SINORHIZOBIUM MELILOTI*

TRANSPORT AND METABOLISM OF B-GLYCOSIDIC SUGARS IN *SINORHIZOBIUM MELILOTI*

By

Katerine Kibitkin

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 (Biology)

MCMASTER UNIVERSITY
Hamilton, Ontario

TITLE: Transport and Metabolism of β-Glycosidic Sugars in *Sinorhizobium meliloti*

AUTHOR: Katerine Kibitkin, B.Sc.

SUPERVISOR: Dr. T.M. Finan

NUMBER OF PAGES: xii, 121

ABSTRACT

The bacterium *Sinorhizobium meliloti* forms N2-fixing nodules on the roots of alfalfa and some other leguminous plants. A gene cluster *smc04260-smc04251* was previously found to be expressed when *S.meliloti* is grown in media containing cellobiose, gentiobiose, salicin and arbutin. These monosaccharides of these sugars are connected via glucosidic bonds of β (1-4) or β (1-6) configuration. The gene cluster *smc04260-smc04251* appears to include an ABC-type uptake system as well as genes for catabolism of these sugars. Data from growth curves with strains carrying plasmid co-integrants suggested that the *smc04260-smc04251* gene region were tested for their ability to grow on cellobiose, gentiobiose, salicin or arbutin as sole sources of carbon. In order to study the effects of specific gene deletions on the growth of *S.meliloti*, *smc04259*, *manB*, *smc04254* were deleted in-frame and growth curves with these strains revealed that *smc04259, manB and smc04254* were all important for growth of aryl-glycosides such as salicin and arbutin.

ManB protein, a putative β-mannosidase enzyme, was overexpressed, purified and then tested in protein assays to determine its activity with different substrates. In one assay, 4-nitrophenyl-β-D-glucopyranoside and 4-nitrophenyl-β-D-mannopyranoside were used as substrates and it was determined that ManB was only active on 4-nitrophenyl-β-D-mannopyranoside. Glucose oxidase assay was used with cellobiose, gentiobiose, mannobiose, salicin and arbutin to study the activity of ManB with each substrate, and it

was found that ManB only had significant activity when mannobiose was used as a substrate.

Transposon mutagenesis with transposon Tn5-B20 was used to select strains that could not grow or grew poorly when cellobiose and arbutin were used as sole sources of sugar. The mutants had transposon insertions within *smc04260-smc04251* gene cluster as well as the *smc03160-smc03165* gene cluster. Five mutants were mapped and tested for their ability to grow on cellobiose, gentiobiose, mannobiose, arbutin and salicin. These mutants were later complemented with pLAFR1 cosmid library. A complete understanding of metabolic capabilities of *S. meliloti* can be beneficial for studying similar systems in soil organisms. The ability to break down a variety of sugars increases the fitness of the organisms and widens the scope of understanding of such systems, which have some potential in biotechnology.

AKNOWLEDGEMENTS

I would like to thank Dr. Finan for giving me the opportunity to work in his lab. I will never forget his never ending patience, his guidance and his understanding. Dr. Finan not only challenged me to be a better student, but also made me grow as a person. I will always remember this experience. Thank you!

I would never be able to finish this without my family. They have been there for me from the start to the very end. Their encouragement put me through the toughest periods of my studies and my life. From the bottom of my heart, thank you for always believing in me.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

Amp – Ampicilin
Cm - Chloramphenicol
Gm – Gentamycin
Km – Kanamycin
Nm – Neomycin
Rif – Rifampicin
Sm – Streptomycin
Sp – Spectinomycin
Tc – Tetracycline
r – resistant
s – sensitive
CAZy – Carbohydrate Active Enzyme Database
GH – glycoside hydrolase
Cel – Cellobiose
Glu – Glucose
Suc – Succinate
LB – Luria Bertani broth
PCR – polymerase chain reaction
SDS-PAGE – sodium dodecyl sulphate polyacrylamide gel electrophoresis
ABC – ATP binding cassette
NM – p-nitrophenyl-β-D-mannopyranoside
NG - p-nitrophenyl-β-D-glucopyranoside
bp – base pair
FRT site – flippase recognition target site

CHAPTER 1

Introduction: general information, glycosides and glycoside hydrolases

Sinorhizobium meliloti genome

Sinorhizobium meliloti is a soil dwelling organism that is agriculturally important due to its ability to form nitrogen – fixing nodules with leguminous plants such as alfalfa. The nitrogen – fixing nodules provide a source of usable nitrogen to the plant, which in turn provides *S. meliloti* with a source of carbon and energy, thus forming a symbiotic relationship. The bacteria can also grow and survive as a saprophythic free-living organism in soil. The genome of *S. meliloti* is unusual as in addition to a chromosome that is 3.65 Mb in size, *S. meliloti* has two megaplasmids pSymA and pSymB that are 1.35 Mb and 1.68 Mb in size respectively (Galibert, et al., 2001). The genome sequence of *S. meliloti* has been determined and annotation revealed that it contains approximately 6,200 protein coding genes. Many (~40%) of these genes code for proteins of unknown function or proteins with only predicted function (Galibert, et al., 2001). It is of interest to determine the function of these various proteins, as among other properties this will reveal what carbon sources *S.meliloti* utilized in the endosymbiotic and saprophytic states. The ability to obtain fixed nitrogen via root nodules is a tremendous advantage for the plant, since on its own it is not able to acquire nitrogen from the atmosphere, and hence this association is very important. The ability of *S. meliloti* to form such relationship with the plant is a result of its ability to successfully survive in the soil on its own as well as in the association with the plant. Many different carbon sources are

available in the soil and in the rhizospehere surrounding the plant root system (Somers, Vanderleyden, & Srinivasan, 2004). The ability of this organism to use these different sources of carbon is important in understanding its metabolic capabilities and the role it plays in that environment. Due to the niche that *S. meliloti* occupies it is clear that the sources of carbon originate mostly from plants and decomposing matter found in the soil. The carbon sources could also be products of large polysaccharides and oligosaccharides that are excreted as products and byproducts of metabolism by other organisms such as fungi (Falip, et al., 2005).

Carbohydrate sources in the soil

Some commonly found sources of carbon in the soil and rhizosphere are small and large carbohydrate molecules. The glycosidic bond which connects two sugar units together may vary in its configuration, although it is the only bond used to form large sugar structures from smaller ones. Glycosidic bonds come in a variety of configurations and specific enzymes are needed to break the bond releasing either of the oligosaccharides, disaccharides as well as monosaccharide sugars such as glucose, mannose, xylose, etc. Simple monosaccharide molecules are readily used by *S.meliloti* for growth (Mauchline, et al., 2006), (Stowers, 1985).

β-Glycosidic bonds in sugars

Plant biomass varies in composition although the main constituents of plant cell walls are always cellulose, hemicelluloses, and others such as lignin or pectin, that depend on

the source of plant cell wall (Minic & Jouanin, 2006). Different kinds of plants possess all three of these constituents in different ratios with other sugars present in varying quantities. Glycosidic bonds connect two sugar units or a sugar to a different moiety in one of two possible configurations: α or β (see Figure 38 in the Appendix). The type of configuration differs from one source to the next although both are well represented throughout various polysaccharides available that result from breakdown of fungi and plants. Two different configurations of the same monosaccharides may result in completely different properties of a resulting disaccharide. β configuration tends to be more common due to its stability, as all the bulkier units in a molecule are in equatorial position, stabilizing the whole structure in a favorable chair conformation, as illustrated in Figure 38 in the Appendix.

Glycoside Hydrolases

Enzymes that hydrolyze glycosidic bonds are referred to as glycoside hydrolases (GH) or glycosidases. Many soil organisms rely on these enzymes in order to break down complex polymers into simple monosaccharide units which can be further used as a source of carbon. These enzymes are specific for the type of the bond they cleave and its configuration. They have been organized into different families and the enzymes in the same GH family tend to have very similar properties and functions (Henrissat & Bairoch, 1993), (Cantarel, Coutinho, Rancurel, Bernard, Lombard, & Henrissat, The Carbohydrate Active Enzymes Database (CAZy) - an Expert Resource for Glycogenomics, 2009). These families are listed in the Carbohydrate Active Enzyme Database (CAZy) (Cantarel,

Coutinho, Rancurel, Bernard, Lombard, & Henrissat, Glycoside Hydrolase Family 2, 1998-2010) and can be accessed on line (www.cazy.org). This site contains all the known and predicted GHs in the various families and allows one to explore the predicted function of each enzyme based on its amino acid sequence. The family organization is based on amino acid similarity with a sub categorization into folds, which corresponds to already known 3D structures.

It is common for bacteria producing glycosyl hydrolases to be equipped with a number of other enzymes responsible for breaking down of larger as well as smaller sugar units, proteins forming uptake systems for these carbon sources, as well as others that can be involved in metabolism. Generally, most research has focused on just the hydrolytic enzymes or just the uptake systems separately from each other. No operons in soil organisms similar to *S.meliloti* have yet been investigated for their ability to form a system for uptake and metabolism of β-glycosidic sugars. It is of major importance to be able to establish that relationship as these operons share some common features as well as differences with other soil organisms.

The gene cluster shown in Figure 40 in the Appendix is under investigation here and contains a gene, *smc04255* or *manB,* annotated as putative β-mannosidase (ManB). Based on this annotation, this enzyme is predicted to be in family 2 of GH enzymes based on the CAZy system of glycoside hydrolase enzyme designation. These kinds of enzymes have been found to be specific for disaccharides that consist of two mannose units connected by a β- (1-4) glycosidic bond. Similar β-mannosidases have been studied in such

organisms as *Bacteriodes thetaiotaomicron* (an intestinal inhabitant of many animals including humans) (Tailford, Money, Smith, Dumon, Davies, & Gilbert, 2007), *Cellulamonas fimi* (an actinomycete) (Stoll, Stalbrand, & Warren, 1999), *Thermobifida fusca* (Beki, et al., 2003), *Thermatoga maritima* (Parker, et al., 2001) and *Thermatoga neopolitana* (Duffaud, McCutchen, Leduc, Parker, & Kelly, 1997) (both thermophilic organisms). A multiple sequence alignment with the known β-mannosidases and ManB from *S.meliloti* showed a significant (almost 30%) amount of similarity especially at catalytically important sites (See Figure 43 in Appendix). Based on the function of most typical β-mannosidases, these enzymes should only be capable of cleaving a β (1-4)-glycosidic bond between two mannose units. This kind of a disaccharide is present in high concentrations in structural and storage plant polysaccharides in the soil, where glucomannan and hemicelluloses are common (Liepman, Nairn, Willats, Sorensen, Roberts, & Keegstra, 2007). Hemicelluloses come in a variety of structures, but mostly their backbones consist of xylose, mannose or mixed sugars (Ebringerova, Hromadkova, & Heinze, 2005). In this case, hemicelluloses with mannose backbones would be more relevant. These can be further subcategorized into glucomannans and galactomannans (Figure 41). Upon breakdown with appropriate enzymes, these sources will yield mannose, glucose and galactose in different ratios. Cellobiose, another product of polysaccharide degradation, is very common as it is the major product of cellulose hydrolysis.

Previous Data

Previously, it has been shown by using strains created earlier for transcription fusion studies (Cowie, et al., 2006) that *S.meliloti* is able to use sugars such as cellobiose, gentiobiose, salicin and arbutin (see Figure 41) as a sole source of carbon (Sartor, 2006). The reporter genes present in the transcription fusion strains are induced when *S.meliloti* is grown in the presence of these sugars. It has also been demonstrated that the promoter upstream from *smc04259* is strongly induced with these same sugars (Mauchline, et al., 2006). However, according to the annotations assigned for each open reading frame in this gene cluster, protein Smc04255 (also referred to as ManB) is predicted to be specific for mannobiose breakdown. As mentioned previously, the ManB protein is highly similar to family 2 of GH, which are strictly β-mannosidases. This is somewhat surprising, as based on the sugars inducing these genes one would expect to see a glucosidase enzyme, such as β-glucosidase, in the gene cluster, paired up with the ABC transporting system as opposed to β-mannosidase. On the other hand, a multiple alignment of known β-mannosidases and putative ManB reveals that there is some similarity with family 2 of β-mannosidases, especially at catalytically important sites (See Figure 43 in Appendix). The validity of this annotation will be studied here. The obvious question is why β-mannosidase is present in the *smc04260-smc04251* gene cluster, when it would be presumably unable to cleave cellobiose, gentiobiose and other glucose containing disaccharides. Also, the associated genes that are present may play some role in further metabolism of the products of sugar breakdown, especially in the case of aryl-glycosides salicin and arbutin. In aryl-glycosides, an aromatic group is linked to the glycone unit

(Figure 41), so it could be that following hydrolysis of glycosidic bond, the other proteins are present to either break down the aromatic group or detoxify it.

ABC transport system

As mentioned before, the *smc04260-smc04251* cluster also includes a putative ABC transporter system encoded by the four genes *smc04259* to *smc04256*. This transporter system is potentially involved in transport of sugars into the cell. ABC transporter systems are very common in both prokaryotes and eukaryotes and their function is to allow for the passage of compounds in and out of the cell (Davidson, Dassa, Orelle, & Chen, 2008). *Smc04259* is annotated as a periplasmic solute binding protein (SBP). In prokaryotes, SBPs bind ligand molecules (or other molecules to be transported) in the periplasmic space and transports these inside the cell (Tam & Saier, 1993). The ABC cassette in prokaryotes normally consists of the following components: two membrane spanning domains and two cytoplasmic ATPase proteins that expend of at least one ATP molecule, to provide energy for the whole system (Driessen, Rosen, & Konings, 2000), (Davidson, Dassa, Orelle, & Chen, 2008) as shown in Figure 42.

Other genes in the cluster

Other genes present in the *smc04259-smc04251* cluster (Figure 40) appear to be involved in regulation and metabolism of compounds transported via the putative ABC transporter. *smc04260* is a putative transcriptional regulator, of the *LacI* family. It may function to repress the system when the ligands are not present. This gene is transcribed

in the opposite direction to the rest of the genes of this cluster and therefore has its own promoter, as has been determined previously (Figure 51). Using library fusion strains it has been determined that in the presence of cellobiose, gentiobiose and salicin *smc04260* transcription is induced (Sartor, 2006). Whether *smc04260* in fact regulates transcription of *smc04260* has yet to be determined.

Gene *smc04254* is annotated as a hypothetical conserved protein. However, comparing the amino acid sequence of this protein with the database available (Nucleotide Basic Local Alignment Sequence Tool, 2009-2010), showed *smc04254* has high similarity to xylose isomerase proteins with a TIM barrel domain. This type of an enzyme is known to interconvert aldoses to ketoses in further metabolism of sugars (Fraenkel & Vinopal, 1973). The *smc04254* gene product may therefore be involved in further metabolism of mannose or glucose. This once again suggests that the genes in this gene cluster are somehow involved in carbohydrate metabolism. Overall it seems very probable that *smc04259* to *smc04256* would make up a transporter system for sugars, as the genes downstream from the putative ABC transporter genes seem to be involved in sugar metabolism. It is worth mentioning that gentiobiose, one of the inducers of the *smc04260-smc04251* gene cluster, is different from the rest of the sugars as it is connected in a β-(1-6) configuration as opposed to β-(1-4) configuration of cellobiose or mannobiose (Figure 41). Gentiobiose may be metabolized via a different set of enzymes, as has been previously established in *Aerobacter aerogenes* (Palmer & Anderson, 1972). Metabolism of gentiobiose in *A.aerogenes* requires synchronized action of two enzymes: β-glucoside kinase and phospho-β-glucosidase. This suggests that perhaps a glucosidase

enzyme alone is not sufficient enough for breakdown of certain sugars and therefore an organism will use an alternate set of enzymes in order to cleave the glycosidic bond resulting in a product of glucose-6-phosphate. This could explain the presence of *smc04254* in the gene cluster of *S.meliloti*. It would be very interesting to investigate the possible activity of *smc04254* gene product in the context of sugar metabolism.

Genes *smc04253* and *smc04252* are both annotated as encoding oxidoreductases. These gene products are likely responsible for different functions, as oxidoreductases are a large group of enzymes that can have a variety of different roles in the cell. *smc04253* is classified within the GFO/IDH/MocA family of oxidoreductases which are identified by a characteristic NAD – binding domains with a Rossman fold (Kingston, Scopes, & Baker, 1996). Proteins that have been previously studied with a similar annotation have been found to mostly be involved in a metabolism of sugars such as L-arabinose (Watanabe, Kodak, & Makinos, 2006), or glucose/fructose (Kingston, Scopes, & Baker, 1996). This implies a similar function in *S.meliloti*.

The conserved domain in *smc04252* appears to be an aldo/keto reductase. This domain family is widely spread within both eukaryotic and prokaryotic kingdoms and is present in proteins involved in metabolism of various compounds, including sugars (Jez, Bennett, Schlegel, Lewiss, & Penninng, 1997). It is not clear whether *smc04253* and *smc04252* would act specifically on certain sugars (mannose for instance, as suggested by putative ManB protein upstream from them) or have a more broad specificity.

Lastly, *smc04251* is annotated as a mannitol – solute binding protein. However, upon searching for a similar amino acid sequence, the major hits were similar to TRAP dicarboxylate transporter – DctP subunit (with closest hits in *Rhizobium sp.* NGR234 and *Sinorhizobium medicae*). This is curious, as TRAP dicarbodylate transporters are involved in transport of nutrients across membranes, sugars included (Rabus, Jack, Kelly, & Saier, 1999). Based on its hydrophobicity plot, Smc04251 protein is predicted to have a signal peptide located at its N-terminus, suggesting that it might be a periplasmic protein. It may be a DctP-like component of a TRAP transporter. It is curious that this is the only gene present on this cluster and other TRAP transporter components are lacking. It has been demonstrated previously that this gene was induced when salicin and arbutin were present in the media (Sartor, 2006).

As a working hypothesis we suggest that together genes *smc04260* to *smc04251* make up a system for transport, regulation and metabolism of specific sugars such as cellobiose, gentiobiose, mannobiose, arbutin and salicin in *S. meliloti*. The major objective in this study is to determine whether this hypothesis is true. An understanding of β-glycosidic sugar utilization in *S. meliloti* will also shed some light on other species of the order *Rhizobiales* as BLASTp searches of these proteins revealed homology with other members of this order. It seems that this gene cluster is conserved in several different species within the order *Rhizobiales*, most of which are soil organisms. Determination of functions of *S. meliloti* genes can suggest the similar system of carbohydrate metabolism in others or can reveal a unique metabolic capability of *S.meliloti*.

This work

In previous studies, reporter gene fusions to genes in the *smc04260-smc04251* gene cluster showed that these genes were induced upon growth with sugars such as cellobiose, gentiobiose, salicin and arbutin (Mauchline, et al., 2006), (Sartor, 2006). The annotations suggested that these genes encode an uptake system as well as genes that are involved in hydrolyzing the β (1-4) glycosidic bond, and perhaps are responsible for further metabolism of sugar monosaccharides.

To investigate the function of the *smc04260-smc04251* gene cluster, we determined the phenotype of each pTH1522 library cointegrate strain, as the gene fusion in each strain will sometimes result in a lack of transcription of that gene or the genes downstream of it. Plasmid pTH1522 (Figure 45) was initially used for creating transcriptional fusion strains and to create the library of these fusions in *S.meliloti* (Cowie, et al., 2006).

The function of the *manB* gene product, annotated as a β-mannoside hydrolase was examined to establish whether it had hydrolase activity on a range of sugars. So far, all the studies that have looked at β-mannoside hydrolases have only found activity on mannose – type glycosides and never on glucose – type glycosides.

Strains carrying in frame deletions of the *smc04259*, *smc04255* and *smc04254* genes were constructed. This experiment was carried out to avoid the ambiguity that library fusion strains can produce in operons, as they do have a polar effect. The polar

effect results from a disruption in a polycistronic mRNA causing loss of transcription due to the end of transcription signals present in a fusion strain. A strain carrying a deletion of the large part of the operon (from *smc04259* to *smc04252*) was studied in order to demonstrate the effect of absence of the whole operon on its ability to use each sugar.

To identify additional genes involved in metabolism of these sugars, transposon induced mutants deficient in the utilization of selected sugars were isolated. The insertion sites in these mutants were determined to identify their exact position within the *S.meliloti* genome.

CHAPTER 2

Materials and Methods

Bacterial Strains and Growth Conditions

All bacterial strains that were used in this work are listed in Table 1. The bacterial cultures were initially grown in Luria-Bertani (LB) media, which consists of 10 g of tryptone, 5 g of yeast extract, 5 g of NaCl in 1 L of double deionized water (ddH_2O). When *S .meliloti* is grown in this media, it is supplemented with 2.5 mM of $MgSO_4$ and 2.5 mM of $CaCl_2$. For solid LB medium, 15 g of agar (Difco) was added per liter of LB broth. All *E. coli* cultures were grown at 37 °C, and *S. meliloti* cultures were grown at 30 °C on a New Brunswick roller wheel at an angle of 30°. Liquid cultures were all grown in 5 mL of liquid media in 150 mm x 18 mm test tubes. When M9 minimal media was used for growth of *S.meliloti*, 5x M9 salts (Difco) were used: Na_2HPO_4 (33.9 g/litre), KH_2PO_4 (15 g/litre), NaCl (2.5 g/litre), NH_4Cl (5 g/litre). M9 salts were further supplemented with 1 mM of $MgSO_4$, 0.250 mM of $CaCl_2$, 1 µg/ml of biotin, 10 ng/ml of $CoCl_2$, and a source of carbon to either 5 mM or 10 mM as necessary. Where M9 media plates were used, 15 g of agar was added for each 1 liter of M9 minimal media. All media and all solutions were autoclaved before use at 15 lbs/inch^2 at 121 °C for 30 minutes. All compounds that were temperature sensitive, such as carbon sources and biotin, were filtered through a 25 mm syringe filter with 0.20 µm membrane (Pall Acrodisc). Carbon sources used for supplementation in minimal media were glucose, cellobiose, gentiobiose, salicin and arbutin (all from Sigma) and mannobiose (Megazyme).

The antibiotics that were used were obtained from Sigma or Boehringer Ingelheim. These were prepared as stock solutions prior to use in culture and kept at -20 °C. Streptomycin, gentamycin, kanamycin, neomycin, spectinomycin and ampicilin antibiotic stock solutions were prepared with water, then filter sterilized as described earlier. Tetracycline and chloramphenicol were prepared in 70% ethanol (high grade, Sigma), while rifampicin was prepared in methanol (HPLC grade, Caledon). The following concentrations of antibiotics were used with *S.meliloti* cultures: gentamycin at 30 μg/ml, streptomycin at 100 μg/ml, neomycin at 100 μg/ml, tetracycline at 5 μg/ml, spectinomycin at 100 μg/ml and rifampicin at 20 μg/ml. With *E.coli* the following concentrations of antibiotics were used: gentamycin at 5 μg/ml, kanamycin at 10 μg/ml, tetracycline at 5 μg/ml, ampicillin at 50 μg/ml, chloramphenicol at 10 μg/ml and spectinomycin at 50 μg/ml. When agar plates were used, these concentrations were doubled for both *S. meliloti* and *E. coli*. When 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) (made at a 12 mg/mL stock in dimethylformamide) was used for blue and white screening it was added to a final concentration of 80 μg/ml.

Table 1: Bacterial strains and plasmids used in this study.

Plasmids used in this work

Plasmid	General Characteristic	Reference
pRK600	Plasmid pRK2013 *npt*::Tn9; Cm^r	(Finan, Kunkel, DeVos, & Signer, 1986)
pTH3224	Plasmid pTH1522 with *S.meliloti* chromosome fragment from 2077096 to 2078151	(Cowie, et al., 2006)
ManB cloning		
pTH1931	pTRC-SC vector; expression vector with ptac promoter; used to create StrepII/CBP fusions at C terminus via *PacI* restriction site	Finan collection
pTH2493	pTH1931::*smc04255(manB)* via primers 3687(F) and 4131(R)	This work
pTH2442	pTH1227::*smc04255(manB)* via *Bam*HI and *Xba*I	This work
pTH1227	pTac-lacZ^q from pMal-c2x in pFUS1; Tc^r	(Cheng, Sibley, Zaheer, & Finan, 2007)
Deletion strains		
pKD4	derivative of pANTSγ containing FRT flanked Km^r gene from pCP15; Km^rAmp^r	(Datsenko & Wanner, 2000)
pKD13	derivative of pANTSγ containing FRT flanked Km^r gene from pCP15; Km^rAmp^r	(Datsenko & Wanner, 2000)
pCP46	derivative of pINT-ts containing araC-P_{araB} and γ, β, exo from pBAD18, Amp^r	(Datsenko & Wanner, 2000)
pUCP30T	ColE1, *oriT* from RK2 cloning vector, RK2; Gm^r	(Schweiser, Klassen, & Hoang, 1996)
pTH2507	*smc04259* flanked with 300 nt homology regions in pUCP30T via primers 210(F) and 211(R)	This work
pTH2518	pTH2507 where *smc04259* has been substituted with Km^r gene from pKD13	This work

pTH2543	*smc04255/manB* flanked with 300 nt homology regions in pUCP30T via primers 2208(F) and 2209(R)	This work
pTH2547	pTH2543 where *smc04255/manB* has been substituted with Km^r gene from pKD4	This work
pTH2636	pTH2615 where *smc04254* has been substituted with Km^r gene from pKD4	This work
pTH1937	pACYC177 with *oriT* from RK2, Nm^r from Tn5	B. Poduska, T.M. Finan
pTH2553	pTH1937 with *S.meliloti* chromosomal region from 2077096-2078151 via *Eco*RI and *Spe*I sites; Km^r	This work
pTH2505	pRK7813 with Flp cassette from pTH1944 via *Pst*I site; Tc^r	C.E. White, T.M. Finan
pTH2434	pTH1227::*smc04259* via *Bam*HI and *Xba*I; Tc^r	This work
pTH2442	pTH1227::*smc04255/manB* via *Bam*HI and *Xba*I; Tc^r	This work
pTH1944	ΔpBBR MCS carrying RK2 with 2 kb *pcaFlp* (via *PstI*); Tc^r	B.Poduska, T.M. Finan
pPH1JI	IncP plasmid with Gm^r, Sm^r, Sp^r from R1033	(Hirsch & Beringer, 1984)
Cosmid clone library		
pLAFR1	RK2 derivative cosmid cloning vector, *Eco*RI site; Tc^r	(Freidman, Long, Brown, Buikema, & Ausubel, 1982)
pTH2615	pLAFR1 with an insert from *S.meliloti* spanning from 2090518 (*smc04260*) to 2071324 (*smc04247*); Tc^r	This work
pTH2616	pLAFR1 with an insert from *S.meliloti* spanning from 3135357 (*smc03173*) to 3113885 (*smc03156*); Tc^r	This work
pTH2617	pLAFR1 with an insert from *S.meliloti* spanning from 3099213 (*smc02977*) to 3122950 (*smc03163*); Tc^r	This work

E.coli Strains used in this work

Strain	Relevant Characteristics	Reference
DH5α	F⁻, *endA1*, *hsdR17* (r_K^-, m_K^-), *supE44*, *thi-1*, *recA1*, *gyrA96*, *relA1*, Δ(*argF-lacZYA*), U169, Φ80d*lacZ*, ΔM15	Invitrogen
MT616	MT607 (pRK600), Cm^r	(Finan, Kunkel, DeVos, & Signer, 1986)
MM294	pRK600::Tn5-B20, pro-82, thiI, hsdR17, supE44, Cm^r, Nm^r	(Finan, Kunkel, DeVos, & Signer, 1986)
ManB Expression		
BL21 pLysS	F⁻, *ompT*, $hsdS_B$ (r_B^- m_B^-) *gal dcm* (λDE3), pLysS; Cm^r	Invitrogen
M1377	DH5α (pTH2442); Tc^r	This work
M1436	DH5α (pTH2493); Sp^r, Sm^r	This work
M833	DH5α (pTH1931); Sp^r	T.M. Finan collection
M65	DH5α (pTH1227); Tc^r	(Cheng, Sibley, Zaheer, & Finan, 2007)
Deletion Strains		
M1417	BW25141 (pKD4); Km^r, Amp^r	(Datsenko & Wanner, 2000)
M1418	BW25141 (pKD13); Km^r, Amp^r	(Datsenko & Wanner, 2000)
M1420	BW25113 (pCP46); Amp^r, Cm^r	(Datsenko & Wanner, 2000)
J252	DH5α (pUCP30T); Gm^r	(Schweiser, Klassen, & Hoang, 1996)
M1451	M1420 (pTH2507); Gm^r, Amp^r, temperature sensitive replicon	This work
M1464	DH5α (pTH2518); Gm^r, Km^r	This work
M1495	M1420 (pTH2543); Gm^r, Amp^r, temperature sensitive replicon	This work
M1502	DH5α (pTH2547); Gm^r, Km^r	This work

M1708	M1420 (pTH2615); Gm^r, Amp^r, temperature sensitive replicon	This work
M1745	DH5α (pTH2636); Tc^r, Km^r	This work
M835	DH5α (pTH1937); Km^r	B. Poduska
FL3224	DH5α (pTH5992); Gm^r	(Cowie, et al., 2006)
M1494	DH5α (pTH2542); Km^r	This work
M1509	DH5α (pTH2553); Km^r	This work
M1449	DH5α (pTH2505); Tc^r	C.E. White
M842	DH5α (pTH1944); Tc^r	B. Poduska
M1896	DH5α (pPH1JI); Gm^r	This work
Transposon mutagenesis		
HB101	F^-, Pro^-, Gal^-, Str^r, Rec^-, ramC1, (r_B^-, m_B^-); contains pLAFR1	(Boyer & Roulland-Dssouix, 1969)
M1709	DH5α (pTH2615); Tc^r	This work
M1710	DH5α (pTH2616); Tc^r	This work
M1711	DH5α (pTH2616); Tc^r	This work

Sinorhizobium meliloti Strains

Strain	Genotype	Reference
RmP110	Rm1021 *ptsC* $^{+}$; Sm^{r}	(Yuan, Zaheer, Morton, & Finan, 2006)
RmK763	Rm5000, isolate of SU47, Rif^{r}	(Finan, Hartweig, LeMieux, Bergman, Walker, & Signer, 1984)
SmFL1580	RmP110 *smc04260*::*gfp/lacZ*, *smc04259*:*tdimer2/gusA* – pTH1522, Gm^{r}	(Cowie, et al., 2006)
SmFL2237	RmP110 smc04254::*gusA/tdimer2*-pTH1522, Gm^{r}	(Cowie, et al., 2006)
SmFL3049	RmP110 smc04255::*gfp/lacZ*-pTH1522, Gm^{r}	(Cowie, et al., 2006)
SmFL4430	RmP110 smc04253::*gusA/tdimer2*-pTH1522, Gm^{r}	(Cowie, et al., 2006)
SmFL6588	RmP110 smc04257::*gusA/tdimer2*-pTH1522, Gm^{r}	(Cowie, et al., 2006)
SmFL4262	RmP110 *smc04251*::*gfp/lacZ*, Gm^{r}	(Cowie, et al., 2006)
RmP1934	Rm5000 (pTH2493); Rif^{r}, Sp^{r}	This work
RmP1752	RmP110 (pTH2442); Sm^{r}, Tc^{r}	This work
RmP2085	SmFL1580 (pTH2553); Sm^{r}, Nm^{r}, Gm^{r}	This work
RmP2086	RmP110 with deletion spanning from 2088466 to 2077096 (from *smc04259* to *smc04252* inclusive) on a chromosome via FRT sites; Sm^{r}, Nm^{r}, Gm^{r}	This work
RmP2202	RmP110 φRmP2086; Sm^{r}, Nm^{r}, Gm^{r}	This work
RmP2201	RmP110 with *smc04255/manB* replaced by Km cassette from pKD4; Sm^{r}, Nm^{r}	This work
RmP2207	RmP110 with *smc04259* replaced by Km cassette from pKD13; Sm^{r}, Nm^{r}	This work
RmP2227	RmP2201 with Km cassette removed via Flp recombinase; Sm^{r} (no pTH2505)	This work
RmP2228	RmP2207 with Km cassette removed via Flp recombinase; Sm^{r} (no pTH2505)	This work
RmP2645	RmP110 (pTH2636); Sm^{r}, Nm^{r}, Tc^{r}	This work
RmP2647	RmP2645 with Km cassette removed via Flp recombinase; Sm^{r} (no pTH2636	This work

	or pTH2505)	
Tn5 Mutagenesis		
RmG212	Rm1021 lac$^-$, Sm^r	Finan collection
RmP2260	RmG212 with Tn5-B20 inserted at 2088084 nt (within *smc04259*) on a chromosome; Arb$^-$, Cel$^-$, Glu$^+$, Suc$^+$; Sm^{r}Nm^r	This work
RmP2261	RmG212 with Tn5-B20 inserted at 2087829 nt (within *smc04259*) on a chromosome; Arb$^-$, Cel$^-$, Glu$^+$, Suc$^+$; Sm^{r}Nm^r	This work
RmP2262	RmG212 with Tn5-B20 inserted at 2083292 nt (within *smc04255/manB*) on a chromosome; Arb$^-$, Cel$^-$, Glu$^+$, Suc$^+$; Sm^{r}Nm^r	This work
RmP2263	RmG212 with Tn5-B20 inserted at 3121280 nt (within *smc03162*) on a chromosome; Arb$^-$, Cel$^-$, Glu$^+$, Suc$^+$; Sm^{r}Nm^r	This work
RmP2264	RmG212 with Tn5-B20 inserted at 3121042 nt (within *smc03162*) on a chromosome; Arb$^-$, Cel$^-$, Glu$^+$, Suc$^+$; Sm^{r}Nm^r	This work
RmP2265	RmP2260 (pTH2615); Arb$^+$, Cel$^+$, Glu$^+$, Suc$^+$; Sm^r, Nm^r, Tc^r	This work
RmP2266	RmP2262 (pTH2615); Arb$^+$, Cel$^+$, Glu$^+$, Suc$^+$; Sm^r, Nm^r, Tc^r	This work
RmP2267	RmP2263 (pTH2616); Arb$^+$, Cel$^+$, Glu$^+$, Suc$^+$; Sm^r, Nm^r, Tc^r	This work
RmP2268	RmP2263 (pTH2617); Arb$^+$, Cel$^+$, Glu$^+$, Suc$^+$; Sm^r, Nm^r, Tc^r	This work

Table 2: Table of primers used in this study

Amplified region	Forward primer	Reverse Primer	Features
ManB	ML-07-3687: 5'GTC ACC TTA ATT AAC CAC GAA TGG ACC ATG GAA ATG C 3'	ML-07-4131: 5' TCA AGG ATA TCG CTC ATT TTT CGA ACT GCG GGT GGC TCC AAG CGC TCC CTT CCT GGC AGG ATT CGA G 3'	Amplifies *manB* via restriction sites *Pac*I and *Hind*III from genomic DNA. StrepII tag is included in reverse primer
ManB check up	ML-07-2297: 5' AAC GGC CCG CAT CGA ACA TG 3' ML-07-2298: 5'GCC GGA TTT CCT GGA GAA TG 3' ML-07-2299: 5'TGC ATG GGC ACG CTC TAT TG 3'		Verify the integrity of amplified and cloned *manB* sequence. Primers 07-3687 and 07-4131 can be used to sequence the very start and end around *manB*
Deletion of *smc04259*	ML-08-210: 5' AGA AGC TTA GGG CGC AAC CCG CAG GTT GAG 3'	ML-08-211: 5' CGT CTA GAC ACG AAG ATG ACC ATG GCC GTG 3'	Amplifies *smc04259* together with 300 nt flanking regions
	ML-08-412: 5' GTT TTT ACG GCA ACG TTA CAG TGA GGG AGG ACA CTC GTG TAG GCT GGA GCT GCT TC 3'	ML-08-413: 5' CAGCCC CGT CAA AGT GCG GGG CTT GGA GCG CCG CTC ATT TCG GGG ATC CGT CGA CC 3'	Amplifies 36 nucleotides around *smc04259* with priming site 1 (forward) and priming site 2 (reverse) from **pKD13**
	ML-08-414: 5' GCC AGC TGG CGA AAG GGG GAT GT 3'	ML-08-415: 5' GGA ATT GTG AGC GGA TAA CAA TTT CAC AC 3'	Amplifies from either side of cloning site in pUCP30T to establish the integrity of the insert
	ML-09-0582: 5' CGA CGC CGA TGT TCA GCC GGA AAT C 3'	ML-09-0583: 5' GCC GCG GCA GGA GCT TGG AAT TGG T 3'	Amplifies 350 nucleotides around *smc04259* in order to check the integrity of the sequence and FRT scar after deletion
Deletion of *manB*	ML-08-2208: 5' CAA ATG GAT CCG CAC GTG AAG GTC GAC GAG GC 3'	ML-08-2209: 5' ATT CAT CTA GAC GAT CTC AAG CAC GCG TGG AG 3'	Amplifies *manB* together with 300 nt flanking regions
	ML-08-2293: 5' CCG CAT CAA TGA AAC GGC GCC GGT GCA GGC CGC CCG GTG TAG GCT GGA GCT GCT TCG 3'	ML-08-2294: 5' GTT GGA ACG GTT GCT TGG GAG TGC GCG TAA AAG GGG ATT CCG GGG ATC CGT CGA CC 3'	Amplifies 36 nucleotides around *manB* and includes priming sites 1 (forward) and 2 (reverse) from **pKD13**

	ML-09-0580: 5' GCT GCA GCC GGG CCA GAA AGT CGT T 3'	ML-09-0581: 5' GCG TCG CTC TCC AGC ATA TCC AGG C 3'	Amplifies 350 nucleotides around *manB* in order to check the integrity of the sequence and FRT scar after deletion
Deletion of *smc04254*	ML-09-2919: 5' AAG CAA CCG TTC CAA CGA AAC CAC CAG GAG ATT TCG TGT AGG CTG GAG CTG CTT CG 3'	ML-09-2995: 5' ACA CCA AGT TCT TTT GTC ATG GGA CTT ATC CGT CGA ATG GGA ATT AGC CAT GGT CC 3'	Amplifies 36 nucleotides around *smc04254* and includes priming sites 3(forward) and 2 (reverse) from pKD4
	ML-09-3093: 5' CGT GAC GGT CGC CGC GAA GGA CAT TC 3'	5' TCC ACC CAG GAA AGT GTC CGG TGC GC 3'	Amplifies 430 nucleotides around *smc04254* to check the integrity of the sequence and FRT scar
RmP2086 construction	ML-4875: 5' TTA GGA CAA CTC CAG TGA AAA GTT C 3'	ML-4876: 5' ATA AGG GAC TCC TCA TTA AGA TAA C 3'	Amplify from borders of GFP gene (forward) and RFP gene (reverse)
	ML-08-4932: 5' TTT CGC AGC CAC TGG TTC GG 3'	ML-08-4933: 5' TTC GAC GTG CTG ATG GTC GC 3'	Amplifies a region inside the borders of insert 1580, from 2087505 to 2088403
	ML-08-4934: 5' TGA TCG CAG GGA TCG AAG CG 3'	ML-08-4935: 5' GCA TTA TCC GCT CGC AGT GC 3'	Amplifies a region outside the borders of insert 1580, from 2090281 to 2091335
	ML-08-4928: 5' CAA CGA AGC GGT AGA GAC CG 3'	ML-08-4929: 5' TTT CAG GCC AAG GAC CTG GC 3'	Amplifies a region inside the borders of 3224, from 2076047 to 2077061
	ML-08-4930: 5' TCG ACA GAA CTT CCG CAC CG 3'	ML-08-4931: 5' CTT CGA TCT CGG CAT CAC CC 3'	Amplifies a region outside the borders of 3224, from 2078298 to 2079637
Tn5-B20 LacZ		ML-1160: 5' CGC CAG GGT TTT CCC AGT CAC GAC GTT GTA 3'	Primer for sequencing the region off of *lacZ* gene off of Tn5-B20
pLAFR1 cosmid	P129: 5' GCA GGT GCT GGC ATC GAC ATT CAG C 3'	P128: 5' CCT CGA TCA GCT CTT GCA CTC G 3'	For sequencing the borders around the insert in pLAFR1

PCR Amplification

PCR amplifications were completed using either HIFI Taq or regular Taq polymerase (Invitrogen) using buffers supplied by the manufacturer. All the primers used in this work are listed in Table 2. The PCR reaction volume was 50 µL for amplification and 20 µL when PCR was used to check the integrity of the amplified DNA fragment. All the reactions had the following final concentrations: 1x buffer, 1.5 mM or 2.0 mM of $MgSO_4$, 0.1 mM dNTPs (Invitrogen), 1 µM of each primer (Mobix), about 50 ng of genomic DNA and 1-2.5 units of polymerase. Where the plasmids were used as a template for amplification, they were diluted with water and used at a final concentration of 0.1 ng per 50 µL reaction.

For colony PCR the template DNA was obtained directly by picking the colony up off the agar plate and resuspending it in the PCR mixture. All the other PCR components were used as described above.

The PCR program differed depending on the size of the product to be amplified as well as the melting temperature of the primers. Generally, the extension time was 30 s per 1 kb, and the primer annealing temperature was 3 °C less than the lowest melting temperature of either primer. Two step PCR was used when amplifying, where the first 5 cycles were run at lower annealing temperature, and the subsequent 25 cycles were carried out with the annealing temperature increased by several degrees to take into account the increasing number of amplified fragments that were then available for templates.

Cloning and DNA Manipulations

All plasmids were obtained from *E.coli* strains using QIAprep Spin Miniprep Kit from Qiagen or Wizard Plus SV Miniprep Kit from Promega. All PCR products and restriction enzyme digestions were purified using QIAquick PCR Purification Kit from Qiagen. Where necessary, gel extraction was carried out to purify DNA from agarose gel using Qiagen QIAquick Gel Extraction Kit. All DNA was quantified with Eppendorf BioPhotometer 6131 using proper dilutions, usually 1 μL of sample with 49 μL of diluent (water). DNA agarose gel was prepared using agarose (BioShop, biotechnology grade) in 0.5x Tris-Borate-EDTA (TBE) buffer and adjusted to make 0.75% to 1% agarose gel (depending on the size of the DNA fragment).

Restriction enzymes were used in 0.5 units to 1 unit amounts in each reaction, depending on the amount of DNA present. For plasmid DNA digestion, the reaction was incubated for at least 4 hours at 37 °C, while PCR products were incubated for at least 6 hours at 37 °C. All DNA was stored at -20 °C.

Ligations were set up in 20 μL reactions. The fresh buffer provided by manufacturer for ligation was used with 2 units of T4 DNA ligase (New England Biolabs), an insert as well as plasmid in 3:1 to 5:1 ratio were used. The volume was adjusted with sterilized ddH_2O. The reaction was incubated overnight at 16 °C and kept in the fridge for storage at -4 °C.

Cloning of the *manB* gene into pTrc-SC vector

Gene *manB/smc04255* is 2,496 bp in length. Due to its size, some difficulty was encountered when cloning this gene into the pTrc-SC vector. pTrc-SC vector is designed to be used in cloning via *Pac*I site, which would result in non-directional cloning. When this site was used to clone PCR amplified *manB*, the experiment resulted in self ligation of the vector with no *manB* present. In order to avoid that, dephosphorylation was used to remove phosphate from the sticky ends on the vector thus preventing self ligation. When this method also failed to give a recombinant plasmid, we employed directional cloning, using a *Hind*III restriction site located downstream from the StrepII tag and *Pac*I site. Since *manB* has a *Hind*III restriction site within its sequence, PCR amplified *manB* was first digested with *Eco*RV to generate a blunt end. *Hind*III digested end on pTrc-SC vector was filled in to make it blunt using 1 unit of T4 polymerase in a reaction mixture containing buffer #2 (New England BioLabs), 1 unit of T4 Polymerase, 1.5 μg of pTrcSC DNA and 100 μM of dNTPs. The reaction was incubated at 12 ˚C for 15 minutes and the vector was then digested with *Pac*I. Because *Hind*III site is located downstream from StrepII tag on pTrc-SC vector, and therefore was cut out in the process, *manB* gene had to be amplified again in order to include the necessary sequence for the Strep tag, along with the proper restriction sites. During amplification, HIFI Taq polymerase was used in order to avoid point mutations and one nucleotide deletions or insertions. Despite this precaution, it has been found that the cloned *manB* gene had several mutations in its sequence resulting in a change of two different amino acids. Due to the difficult nature of the cloning, it was much easier to use already made plasmid pTH2442 (previously made,

pTH1227 with *manB* cloned in it). The cloned region of this plasmid had been sequenced and found to be free of errors. The piece of DNA fragments from *manB* containing the two errors was substituted with the identical fragment from pTH2442 via restriction sites *Pst*I and *Hind*III (see Figure 44). Ligation and digestion were all carried out as described previously. The resulting plasmid, pTH2493, had no mutations, and contains StrepII tag downstream from *manB* gene.

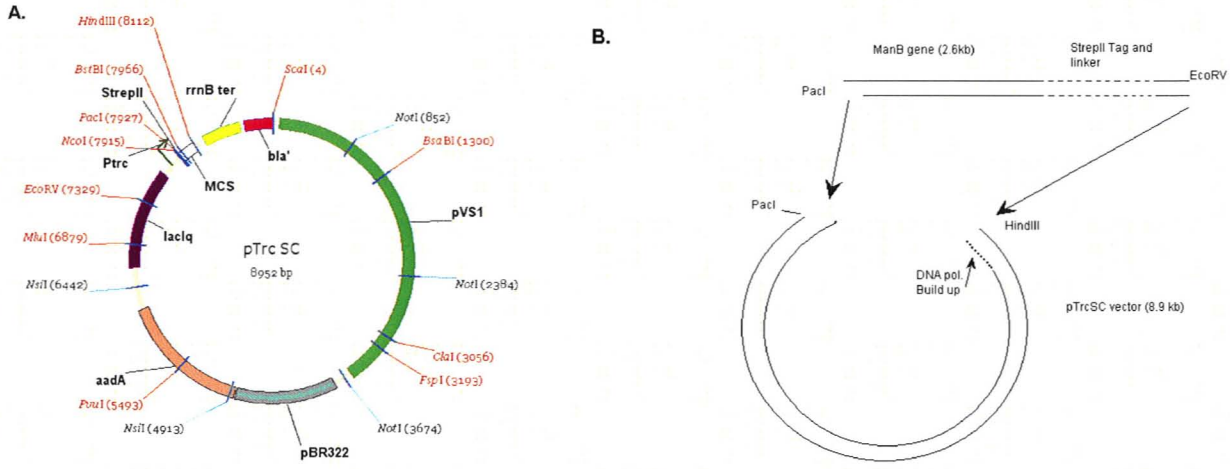


Figure 1: **(A)** pTrcSC vector. This vector is an expression vector used for overexpression of ManB protein. It contains C-terminal StrepII tag and calmodulin – binding peptide. Ptac promoter is IPTG inducible. **(B)** The cloning of *manB* into pTrcSC vector is shown. Since the restriction sites used for pTrcSC vector also cut within *manB* gene, a different approach had to be used. The sites that were used also cut out the tags included in the vector. The Strep tag was therefore included in the reverse primer for *manB* amplification. The resultant vector is pTH2493.

Transformation of DNA into Bacteria

Transformation was carried out as described in Molecular Cloning Laboratory Manual (Sambrook Joseph, Russel David W., 2001). DH5α *E.coli* cells were grown to an OD_{600} of 0.35-0.4 and the culture was transferred to a clean Falcon tube and chilled on ice at 4 °C for 10 minutes. The culture tube was centrifuged on a Beckham centrifuge at 4500 rpm at 4 °C to get rid of the media. The pellet was resuspended in 30 mL of ice cold $MgCl_2$-$CaCl_2$ solution (80 mM $MgCl_2$, 20 mM $CaCl_2$). The pellet was recovered by centrifugation at the same speed for 10 minutes, then resuspended in 1 mL of ice cold $CaCl_2$ (0.1M). In order to increase efficiency, these were stored on ice for exactly 16 hours, after which the cells were aliquoted in 50 μL volumes into eppendorff tubes to be frozen at -80 °C. The cells were thawed on ice when needed. Ligation mix was added to each tube in either 2 μL or 5 μL volumes.

Electroporation competent cells were prepared by growing 2 mL of culture to an OD_{600} of 0.6. No antibiotics were used. The cells were then spun down in Beckman tabletop centrifuge at 13,000 rpm, and the cells were washed twice in cold 10% glycerol solution, and resuspended again in 500 μL- 1 mL of ice cold 10% glycerol. These cells were then flash frozen in 100 μL aliquots. The cells were thawed on ice when needed. For electroporation, competent cells and 1-2 μL of DNA were mixed in cold electrode gap cuvettes. The electroporation was carried out using BioRad Gene Pulser 2 with 0.4 cm electrode gap cuvettes. The adjustments on the apparatus were as follows: current was adjusted to low setting, 200 Ohms resistance, capacitance was 25 μFaradays. The voltage

was adjusted to 2.2 kiloVolts for *E.coli* and 2.5 kiloVolts for *S.meliloti*. After the electroporation was complete, the cells were incubated for 1-2 hours and spread plated using the same procedure as described above for transformation.

Conjugal Transfer of DNA

Recombinant plasmids were transferred to recipient *S.meliloti* via conjugation with the donor *E.coli*. Conjugation was carried out in the presence of MT616, the *E.coli* helper strain containing plasmid pRK600, which contains *tra* genes necessary for the transfer of DNA (Finan, Kunkel, DeVos, & Signer, 1986). The necessary donor, recipient and MT616 cultures were grown to OD_{600} of about 0.8-0.9. 1 mL of each culture was spun down in Eppendorf tabletop centrifuge at 13,000 rpm, and the pellet was washed twice in 0.85% NaCl solution to get rid of antibiotics. Following the wash, each cell pellet was resuspended in 100 µL of 0.85 % NaCl and 30 µL was then spotted onto LB. As a control, each culture was also spotted onto LB separately, as well as a mix of recipient and helper, and donor and helper. Once the spots were dry, they were incubated at 30 °C over night.

The next day, the mating mix was gently collected from LB agar plate and resuspended in 1 mL of saline solution. A dilution series was prepared from 10^0 to 10^{-7}. 0.1 mL of selected dilution tubes was spread plated on selective media to select for each of the donor, recipient and transconjugant. Likewise, the control mixes of donor and helper as well as recipient and helper were resuspended in 1 mL saline and 0.1 mL of each was spread plated onto the same selective media used for transconjugant. The

dilution tubes that were spread plated were chosen based on the number of cells that were expected to grow. All the cultures used in the mating were streaked on all the media used to make sure no contamination or errors had occurred. Transconjugant strains were streak purified 3 times prior to saving those into the frozen perm library at -80 ˚C.

Library Strains

The reporter vector pTH1522 contains the genes encoding green fluorescent protein (*gfp*) and β-galactosidase (*lacZ*) encoded in one direction and genes specifying red fluorescent protein (*tdimer2*) and β-glucuronidase (*gusA*) encoded in the opposite direction (Cowie A. , et al., 2006), (see Figure 45 in Appendix). pTH1522 is a suicide vector in *S. meliloti* that must undergo a homologous recombination event resulting in a fusion library strain. The strains in Figure 2 have been used in this study, since they contain fusions at the specific place in *S.meliloti* genome. Since these are integrated within the operon, they can result in polar mutations. These mutations can be used to study phenotypes when these strains are grown with cellobiose, gentiobiose, salicin or arbutin as a sole source of carbon.

Growth Curves

The ability of strains to grow in minimal media with defined carbon sources was determined in M9 minimal media and supplemented with 10 mM of glucose, 5 mM of gentiobiose, mannobiose, cellobiose and arbutin. Salicin was used at 10 mM concentration, to make it comparable to the rest of the sugars used as it only has one

glucose unit. Arbutin was only used at 5 mM because it was inhibitory to the bacterial growth at higher concentrations.

For growth curves, strains were grown in 5 mL of media in 150 mm x 18 mm test tubes aerated on a rotating roller drum at 30 °C. Strains were grown overnight in LB, washed three times with 0.85% NaCl and inoculated to a final OD_{600} of 0.01. Readings were taken every 6 hours using Bench Lab Spectronic 20. All samples were done in triplicate and the final number is the average of the three. At the termination of each curve, each tube was tested for contamination by streaking its contents onto LB agar and examining the morphology of the resulting colonies.

In-Frame Deletion Strains

To delete specific genes in frame, and hence reduce the polar effects, λ-red recombinase was used (Datsenko & Wanner, 2000). 36-nucleotide homology regions, spanning the gene that was to be deleted were used with the following templates for PCR amplification: pKD13 (Acc #AY048744) for deletion of *smc04259* and *manB*, pKD4 (Acc#AY048743) for deletion of *smc04254*. The DNA fragment of the gene and 300 nucleotides directly outside that gene were amplified using *S.meliloti* genome. That fragment was then inserted into the suicide cloning vector pUCP30T, which replicated in *E.coli* but not in *S.meliloti*. The recombinant plasmid was transformed into *E.coli* M1420 cells and grown at 30 °C. The transformants were selected on LB containing gentamycin (10 μg/mL) to select pUCP30T. This strain was made electrocompetent and Fragment 1 was electroporated into it (Datsenko & Wanner, 2000). The strain was cured of the λ-red

recombinase plasmid by growing it at 37 °C. This resulted in a strain containing pUCP30T vector with 300 nucleotide regions spanning each gene that needed to be deleted, where that gene was substituited with Nm/Km cassette. This recombinant vector was transformed into *E.coli*. The deletion was introduced into *S.meliloti* via conjugation with the *E.coli* strain, as described above. The transconjugant was selected with Sm and Nm, and then screened for Gm sensitivity to find the strain that underwent double homologous recombination via 300 nucleotide regions included to take up the vector and then excise it out of the genome, only leaving Nm cassette where the gene of interest has once been. All Gm sensitive strains were sequenced with primers ML-09-0582 and ML-09-0583 for *smc04259* deletion and ML-09-0580 and ML-9-0581 for *manB* deletion to ensure a successful recombination event. Further, to delete the Nm cassette, and hence end up with in-frame deletion, FRT sites, spanning the cassette were used. This was done via strain M1449, which contains a plasmid with Flp recombinase being expressed. The strains were selected with Tc resistance and Nm sensitivity, since pTH2505 is Tc resistant. Because pTH2505 is an unstable vector, it is lost in subsequent generations. After streak purification of the desired strain, only Tc sensitive colonies were selected. The final strain was always checked up by sequencing using same primers as before. Genomic prep was performed for each strain, and this template was used in PCR to amplify the necessary fragment and send it for sequencing. See Figure 46 for more details.

To delete *smc04254* in frame, the procedure used was slightly different. DNA fragment with 36 nucleotides around *smc04254* was amplified as described above.

However pTH2615, a pLAFR1 cosmid vector with an insert of a region spanning from *smc04260* to *smc0247*, was transformed into M1420, where DNA fragment with 36 nucleotides around *smc04254* was then electroporated. The resultant strain, M1745 that contained the recombinant plasmid pTH2636 was verified by sequencing. This plasmid was introduced via conjugation into *S.meliloti* and the resulting strain was once again conjugated with *E.coli* M1896 that contained plasmid pPH1JI which is incompatible with pTH2636. In order to select for Nm, the plasmid pTH3636 would have to go through homologous recombination, deleting *smc04254* and substituting it with Nm cassette. The resultant strain was treated as previously, following homologous recombination. It was screened for Nm resistance and Tc sensitivity to ensure the plasmid was not taken up in recombination event, but was taken out via a second recombination event, leaving only Nm cassette where *smc04254* has been. Nm cassette was removed via Flp recombinase as described earlier. See Figure 47 in the Appendix for more details.

Construction of Strain RmP2202

To delete the 2088466-2077096 region spanning from *smc04260* to *smc04251* the pTH1522 library strain SmFL1580 was used, as it has a fusion present at *smc04260* and *smc04259*. *E.coli* library strain EcFL3224 contains pTH1522 with *S.meliloti* gene fragment spanning from 2077096 nucleotides to 2078151 nucleotides, which is around *smc04251*. This fragment was excised with *Eco*RI and *Spe*I and inserted into pTH1937 to give plasmid pTH2553 which was then transferred to SmFL1580 with selection on LB with Nm and Sm to generate strain RmP2085. The Flp recombinase plasmid pTH1937

was transferred into Rmp2085 (in *E.coli* M842) to take out the whole region spanning from *smc04260* to *smc04251* via Frt sites present (refer to Figure 48 in the Appendix). Because the region that came out also had *lacZ* gene present, the resultant strain was selected for with Sm, Tc and Xgal; only white colonies (where no *LacZ* enzyme was present and that were unable to cleave X-gal) were selected from a mix of blue and white colonies to ensure the proper fragment was gone. The final strain was entered into the library as RmP2086.

To examine this deletion in a wild type background, it was transduced from strain RmP2086 into RmP110 via phage φM12 (Finan, Hartweig, LeMieux, Bergman, Walker, & Signer, 1984) with selection on LB agar containing Gm (60 μg/mL). The final strain is entered into the library as RmP2202.

SDS-PAGE and Western Blot

SDS Page gel was prepared as described in the Appendix in Molecular Cloning: A Laboratory Manual (Sambrook Joseph, Russel David W., 2001). The gel was prepared at 7.5% and 10% acrylamide, where necessary. To prepare a sample for SDS –PAGE, a culture was grown up to OD_{600} of about 0.9, after which the cells were resuspended in SDS sample buffer (50 mM Tris-HCl, pH 6.8, 2% SDS, 1% β-mercaptoethanol, 12.5 mM EDTA, 10% glycerol, 0.01% bromophenol blue) and boiled for 5 minutes. Following boiling, each sample was centrifuged for 2 minutes at 13,000 rpm. The soluble portion of each sample was loaded onto a gel in 5 μL, 10 μL or 20 μL volumes, as needed. The gel was run in a running buffer (25 mM Tris, 200 mM Glycine and 1% SDS) for 60-70

minutes at 150 mV. After the gel was ready to be stained, it was removed from the gel-running apparatus and placed into SDS staining solution (30% methanol, 20% acetic acid, Comassie blue dye) for 2-3 hours or longer and destained until the protein bands were clearly visible. The gel was kept in de-ionized water once it was destained.

When Western blots were prepared, the same procedure for running a SDS gel was carried out. The Bio-Rad Trans – blot SD Semi dry transfer cell was used to transfer the protein from the gel to the PVDF membrane. The PVDF membrane was cut in order to fit the blotting paper. It was soaked in 80% methanol for 1 minute. After washing the membrane in de-ionized water twice, it was placed in a separate container. The membrane and the blotting paper were then both soaked in transfer buffer (24 mM Tris, 192 mM Glycine and 20% methanol). The gel obtained from SDS-PAGE was also rinsed in transfer buffer. To transfer the protein onto a membrane Bio-Rad Trans – blot SD Semi dry transfer cell was used. On the cathode side of the transfer cell, the following was arranged: blotting paper, membrane, gel, blotting paper. The cell was run at 10 V for 60-75 minutes. Once the membrane was ready, it was incubated for 1-2 hours at room temperature in a solution of TBST buffer (20 mM Tris-HCl, 150 mM NaCl, 0.05% of Tween 80) with 3% skim milk. This solution was then poured off and the primary antibody was used in 1:1000 dilution in TBST buffer with 1.5% skim milk. The primary antibody was anti-StrepII tag, as the recombinant ManB protein contains this tag. The membrane was incubated at room temperature for 1-2 hours with gentle shaking. Once the membrane was ready for secondary antibody, it was washed 3 times for 10 minutes each time in TBST buffer. In a clean staining dish, the membrane was incubated again in

TBST with 1.5% skim milk and goat anti-mouse antibody diluted at 1:5000 for 1 hour. When the incubation was done, the membrane was washed three times for 10 minutes each in TBST buffer. The colour development was done in a new clean dish. The membrane was equilibrated for 5 minutes in AP buffer (100 mM Tris, pH 9.5, 100 mM NaCl and 5 mM $MgCl_2$). The AP buffer was removed from the membrane and the membrane was then incubated in AP colour developing solution (AP buffer with 0.33 mg/mL of nitro blue tetrasolium (NBT, Sigma) and 0.165 mg/mL of 5-bromo-4-chloro-3-indoyl phosphate (BCIP, Sigma)). The membrane was stored in dark place in this solution, until the colour developed (10 min).

Bradford assay for protein quantification

Total protein in cell extracts as well as the quantity of pure ManB were all determined using Bradford assay (Bradford, 1976). Bradford assay dye reagent was obtained from BioRad and used according to manufacturer's instructions. The reactions were measured at 595 nm using Varian UV-Vis Carry Spectrophotometer in 1.5 mL plastic cuvettes, using the total reaction volume of 1 mL. The unknown concentration samples were tested using two different amounts. Then the average between the two results was taken.

French Press and ManB Protein Purification

To overexpress ManB protein in BL21 *E.coli* cells, cultures originally frozen and kept at -80 °C were made to be used for that purpose, as BL21 transformant cells are too

unstable (as the plasmid can be lost rapidly) to be used otherwise. pTH2493 was transformed into BL21, and 500 mL of these cells were grown up in LB (Amp 100 μg/mL) to an OD_{600} of about 0.9. The cell culture was resuspended in 0.5 mL of 10% glycerol in 0.85% saline. 50 μL aliquots were flash frozen in liquid nitrogen. For overexpression of ManB, 200 mL of culture was prepared and inoculated with 50 μL of BL21 cells containing pTH2493. The cells were grown at 30 ˚C to an OD_{600} of 0.4. At this point, isopropyl β-D-thiogalactopyranoside (IPTG) was added to a flask to a final concentration of 0.25 mM to induce the expression of ManB protein. 4 flasks were prepared in total: 1-Uninduced culture grown at 25 ˚C, 2-Induced culture grown at 25 ˚C, 3- uninduced culture grown at 30 ˚C, 4- Induced culture grown at 30 ˚C. The cultures were grown until OD_{600} reached 0.9. The culture was then transferred to a centrifuge tube and spun for 20 minutes at 7,000 rpm at 4 ˚C. The pellet was resuspended in 1 mL of Buffer W (20 mM Tris, 150 mM NaCl, 5% glycerol, 10 mM EDTA, 1 mM DTT) and stored at -80 ˚C. For overexpression of ManB protein in *S.meliloti* RmP1934, same procedure was used for growth and induction, except only one culture flask at a total volume of 1 litre was used. The resulting pellet was resuspended in Buffer W (20 mM Tris, 150 mM NaCl, 10 mM EDTA, 1 mM DTT, 5% glycerol) and stored at -80 ˚C.

A French press apparatus was used to effectively lyse the cells. Because the total volume for the small French press cell was 4 mL, all the samples were adjusted to that volume. The buffer used (20 mM Tris, 150 mM NaCl, 10 mM EDTA and 1 mM DTT) was previously chilled. All the French press components were also kept at 4 ˚C until used. The French press was performed at 1000 psi, and each sample was put through twice

(until the suspension was clear). Once lysed, 10 μL of proteinase inhibitor (Sigma) was added and all the samples directly out of the stock bottle and were kept on ice. Following lysis, each sample was centrifuged for 1 minute in an Eppendorf tabletop centrifuge at maximum speed to remove unlysed cells. The supernatant was further centrifuged in Beckman Ultracentrifuge at 40,000 rpm. Pre-chilled Beckman rotor SW4171 was used along with the proper buckets. Because the weight of each bucket against the one opposite to it is crucial, they were all adjusted with chilled BufferW to be the same weight (to 2 decimal points). Each cap and bucket was greased creating a seal to prevent sample leakage. The centrifugation was performed at 4 ˚C for 20 minutes. Once complete, the soluble portion was transferred into 1.5 mL tubes in 1.0 mL aliquots (5 tubes in total). The insoluble section was resuspended in 1% SDS in Buffer W+ (20 mM Tris, 150 mM NaCl, 10 mM EDTA and 1 mM DTT, 5% glycerol) to 5 mL and also saved in 1.0 mL aliquots. All tubes are stored at -80˚C.

To purify ManB protein from the soluble portion obtained, Strep-Tactin column was used with the protocol available (Schmidt & Skerra, 2007). BufferW was made as described in previous section, not as described in that protocol. The other buffers were made as in the protocol: Buffer E (100 mM Tris, 150 mM NaCl, 1 mM EDTA, 2.5 mM D-desthiobiotin (Sigma)) and Buffer R (mM Tris, 150 mM NaCl, 1 mM EDTA, 1 mM 2-(4-hydroxyphenylazo)benzoic acid (HABA)). A 5 mL column was used for purification, and 4 mL of resin was used. The resin has a capacity of 50-100 nmol/mL (Schmidt & Skerra, 2007) and there is approximately 226 nmol (g of protein /209,592 g/mol ManB) of ManB present in the supernatant (as determined by Bradford assay), therefore the

volume of 4 mL will assume the lowest possible capacity. The purification was performed at 4 °C, as suggested by (Schmidt & Skerra, 2007). The elutions were collected in 0.2 mL fractions, 7 fractions in total.

ManB Assay with p-nitrophenyl compounds

Purified ManB was assayed for activity using chromogenic substrates p-nitrophenyl –β-glucopyranoside (Sigma) and p-nitrophenyl –mannopyranoside (Sigma). The activity of ManB was tested at different pH values in 50 mM of each buffer. The following buffers were used: citrate buffer at pH5, phosphate buffer at pH 6, 7 and 8, and carbonate buffer at pH 9. Each substrate was used at a concentration of 1 mM. The volume of the total reaction was 1 mL. Each reaction was carried out by adding the following components: 500 µL of 0.1 mM buffer (to 50 mM final concentration), 30 µL of 32.5 mM substrate (to 1 mM final concentration), 20 µL of 0.1 µg/ µL ManB protein (to final concentration of $2.0x10^{-3}$ µg/ µL or 2 µg in total per 1 mL) and 450 µL of dd H_2O. The reaction was initiated upon addition of ManB protein. Absorbance was measured using Carry UV-Vis Spectrophotometer. The Carry Kinetics software was used and the conditions were adjusted to the following: 350 nm (isosbestic point of 4-nitrophenol), readings taken every 0.5 seconds for 15 minutes. The slope was calculated automatically by the software as ΔA/min with t_i=0 min and t_f=10 min. The line of best fit was built by the Kinetics software from Carry, Varian. Assays were carried out in triplicate. The activity was calculated with the molar absorption coefficient ε=5,270 L/mol cm at 350 nm (Iacazio, Perissol, & Faure, 2000), or 5.270 µmol/mL. The activity

was calculated by dividing the slope by the molar absorption coefficient and the specific activity by dividing the amount of protein present in the 1 mL or reaction, 0.002 mg. The final result was the specific activity of mmol/min mg protein.

Glucose Oxidase Assay for ManB Activity

Glucose oxidase assays were carried out using Glucose oxidase kit (Sigma, GAGO20). The procedure was carried out as suggested by the manufacturer, with these following changes. 50 mM sodium phosphate buffer was used in each reaction at pH 7, as that is the optimal pH for glucose oxidase enzyme (Bergmeyer & Bernt, 1974). The standard curve for glucose used amounts from 0 µg – 80 µg, and for mannose from 0 µg to 800 µg (higher amounts of mannose were used due to the inability of glucose oxidase to oxidize mannose as well as glucose) (Bergmeyer & Bernt, 1974). Each reaction was performed by adding enough of each sugar from a 1 mg/mL stock for glucose and a 10 mg/mL stock for mannose to the buffer to the total volume of 200 µL. Assays were performed in triplicate.

Each reaction was measured at 0 min, 40 min and 60 min. Each reaction was measured against a blank, which contained only the buffer and the substrate with no enzyme present. Assay mixture of 200 µL contained 5 mM: cellobiose, mannobiose, gentiobiose, salicin or arbutin, and 1.5 µg of pure ManB and buffer. Reactions were performed in triplicate. Reactions were stopped by boiling the tube for 5 minutes in order to inactivate ManB.

All the glucose oxidase kit components were added in the same ratio as described in the kit, except in 5 times less volume. This ensured that the total volume of reaction mix and glucose oxidase test components were 1 mL. To 200 µL of reaction, once it was stopped, 400 µL of glucose oxidase solution with o-dianisidine were added and incubated at 37 ˚C for 30 minutes. To develop the coloured complex at the end of the reaction, 400 µL of 12 N sulfuric acid was used. This volume was then transferred to plastic cuvettes and measured at 540 nm in Carry UV Spectrometer, against each blank.

Transposon Mutagenesis with Tn5-B20

Tn5-B20 mutagenesis was performed using *S.meliloti* strain RmG212 (recipient) and *E.coli* strain M294 (donor) (self transmissible plasmid pRK600 with Tn5-B20). Over 1000 colonies were obtained on LB/SmNm media and each colony was screened for growth by patching on M9-succinate, M9-glucose, M9-arbutin, M9-cellobiose (all M9 media plates also contain SmNm). Patches that grew well on LB/SmNm, M9-succinase, M9-glucose but failed to grow well or to grow at all on M9-arbutin or M9-cellobiose were selected. Insertion sites in strains that grew either slowly or didn't grow at all on M9-cellobiose, M9-arbutin but grew on all the other plates were sequenced to determide the location of the transposon. Genomic DNA was prepared from each strain and a *lacZ* primer (ML1160:5'CGCCAGGGTTTTCCCAGTCACGACGTTGTA 3') was used for sequencing. Mutants RmP2260, RmP2261, RmP2262, RmP2263 and RmP2264 were isolated (Table 1).

Complementation of transposon mutants with pLAFR1 Library

A previously prepared *S.meliloti* 1021 cosmid library (Freidman, Long, Brown, Buikema, & Ausubel, 1982) was used to isolate cosmids that complement the transposon insertion mutants. This library was propagated in *E.coli* HB101. A frozen vial containing the culture in 7% dimethyl sulfoxide (DMSO) was thawed and transferred 1:10 dilution into LB containing 2.5 µg/mL Tc. The cells were allowed to grow at 37 °C to an OD_{600} of approximately 1.0. A regular overnight conjugation procedure was then carried out using the library, the mobilizing strain *E.coli* MT616 and the *S.meliloti* recipient strains. Transconjugants were selected on M9-Salicin with 200 µg/mL Nm, and transconjugant colonies were picked up from M9-Salicin/Nm plates and streaked out again on LB/Tc. The insert DNA from several complementing clones was identified by DNA sequencing with primers that flank the insert cloning site, primer ML-128 (5'CCTCGATCAGCTCTTGCACTCG 3') and primer ML-129 (5'GCAGGTGCTGGCATCGACATTCAGC 3') (Table 2). All sequencing was completed in the MOBIX Lab at McMaster University.

CHAPTER 3

Growth Curve Results

S.meliloti Fusion Library Strains

Introduction

Vector pTH1522 is a suicide vector in *S.meliloti* where it is not able to replicate (Figure 45). Recombination between the insert in pTH1522 and the homologous region in the *S.meliloti* genome results in the integration of pTH1522 together with its reporter genes (see Figure 45). Therefore if growth of the resulting strain results in the induction of the promoter upstream from the reporter genes, it can be measured as fluorescence or enzyme activity. Because the vector pTH1522 includes transcription termination sites, any of the genes downstream of this insertion will not be transcribed if the insert does not contain a promoter. This results in polar mutations and using this property, certain library strains previously constructed were selected and used for growth studies of these in defined M9 minimal media with an appropriate carbon source. If these mutations result in specific phenotypes when grown in defined media, these can be useful in finding out the function of the genes in the *smc04259 – smc04252* operon (see Figure 40). Figure 2 shows the strains employed for growth studies when only glucose, cellobiose, gentiobiose, salicin or arbutin are used in the media. Therefore, strain SmFL6588 will only express *smc04259* and *smc04258*; strain SmFL3049 will express *smc04259*, *smc04258*, *smc04257* and *smc04256*; strain SmFL2237 will express *smc04259* to

smc04256 (inclusive); strain SmFL4430 will express *smc04259* to *smc04253* (inclusive); strain SmFL4262 will express *smc04259* to *smc04252* (inclusive).

When using the library fusion strains in a growth curve, it is important to note that the library strains can revert to the wild type phenotype. Under normal circumstances, Gm is added to the media, however for growth studies in M9 minimal media no antibiotic was added as it can interfere with the growth rate of strains and would inhibit growth of the wild type. To analyze for revertants at the termination of the growth curve, each tube was streaked out on LB and 100 colonies from each library strain were then patched on LB/Sm and LB/SmGm. The patches should all be able to grow on LB/Sm and LB/SmGm, however, upon reversion colonies will be Gm sensitive and therefore unable to grow on LB/SmGm.

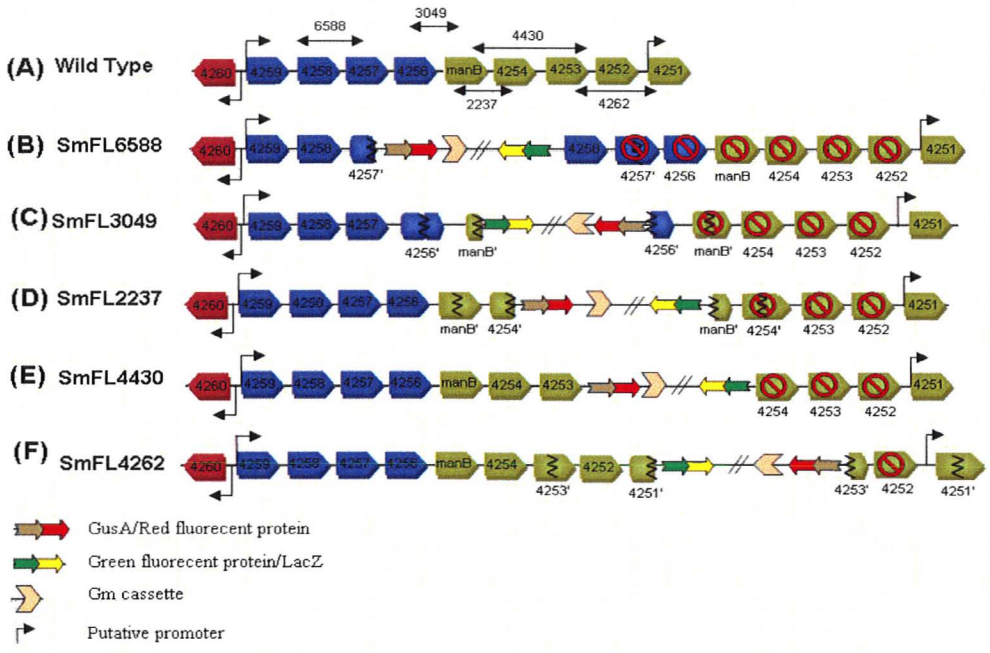


Figure 2: Library fusion strains that were used in this study. The direction of an arrow denotes the direction of transcription. The black zigzag line around some genes denotes the interruption of that gene due to the fusion made. The two promoters upstream from *smc04260* and upstream from *smc04259* have been mapped. Promoter upstream from *smc04251* is putative. Wild type strain is RmP110. The lines above that strain denote the locations of each fusion and the number above the line is the number assigned to each fusion. Around the reporter genes, the two parallel lines denote other genes present of pTH1522. For more detail, see Figure 45 in the Appendix.

Results

The library strains chosen for this experiment are all shown in Figure 2. They were all grown at the same time, under the same conditions. The measurement was taken every 6 hours until the strains were in stationary phase.

Growth in M9 minimal media with each different sugar was found to proceed at a different rate, and should be taken into consideration when comparing each growth curve. As shown in Figure 3, aryl-glycosides when used as sources of carbon tend to result in longer doubling times (Table 3). This slower growth can be explained by the structure of aryl-glycosides, as these sugars consist of a glucose unit bound to an aromatic molecule via glycosidic bond. These sugars may result in a toxic by-product upon the cleavage of the β-glycosidic bond and release of an aromatic compound into the cell.

Table 3: Doubling time of wild type *S.meliloti* strain RmP110 in M9 minimal media with each sugar as a sole carbon source. The data for this table is obtained from Figure 3. The mean growth rate constant k was calculated per hour using the equation $k=(\log N_t-\log N_o)/0.301$, where N is the number of organisms measured at OD600 at the start and at the end of measurement. Mean generation time g was calculated per hour using equation $g= 1/k$. g is the doubling time, which is the time it takes for the cells to increase to twice the number of the cells prior to measuring.

Carbon Source	Doubling time
10 mM Glucose	3.5
5 mM Cellobiose	4.0
5 mM Gentiobiose	3.5
5 mM Salicin	10.2
5 mM Arbutin	9.9

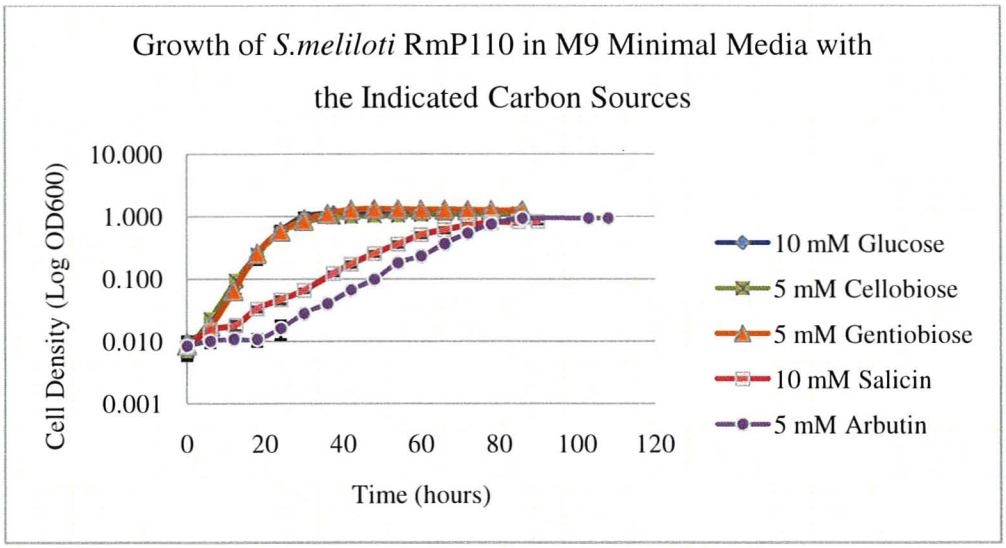


Figure 3: Growth rate of wild type *S. meliloti* strain RmP110 in M9 minimal media with an indicated sugar as a sole source of carbon. Growth with arbutin and salicin is measured over a longer time period as the doubling time of cells when salicin and arbutin are used is much longer.

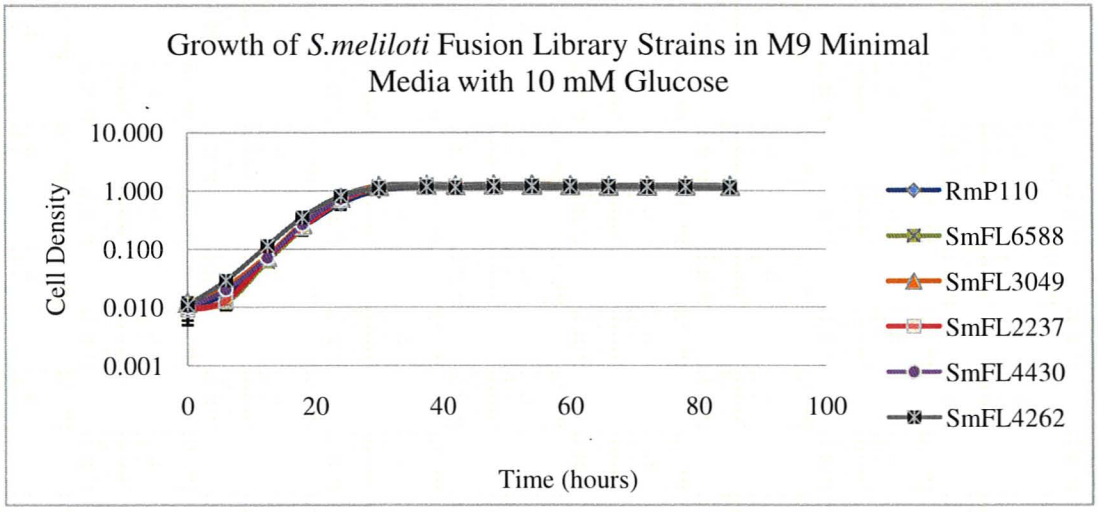


Figure 4: Growth of *S.meliloti* library fusion strains in M9 minimal media with 10 mM of glucose as a sole source of carbon. Cell density is log of OD600.

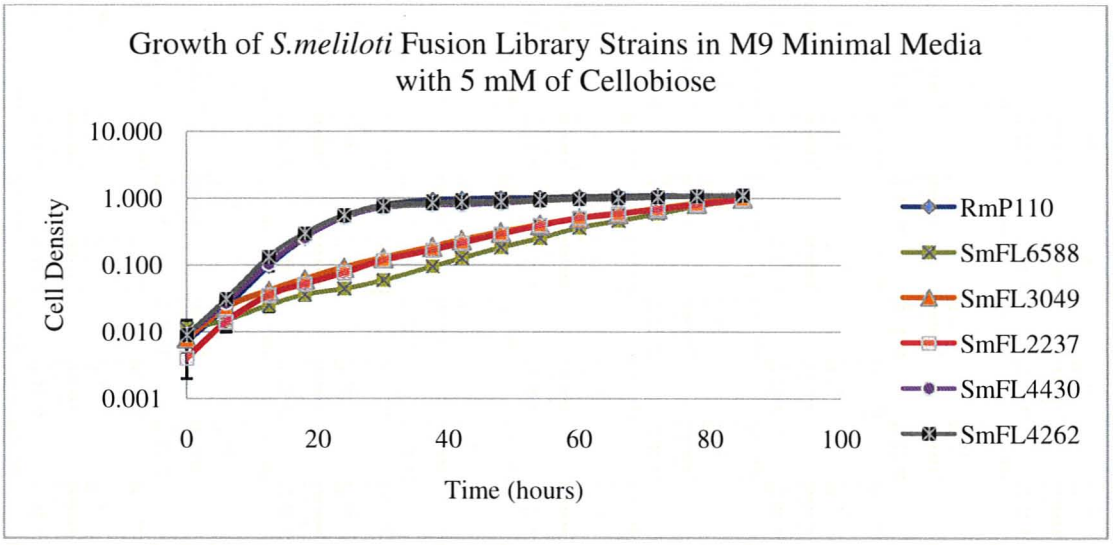


Figure 5: Growth of *S.meliloti* library fusion strains in M9 minimal media with 5 mM of cellobiose as a sole source of carbon. Cell density is log of OD600.

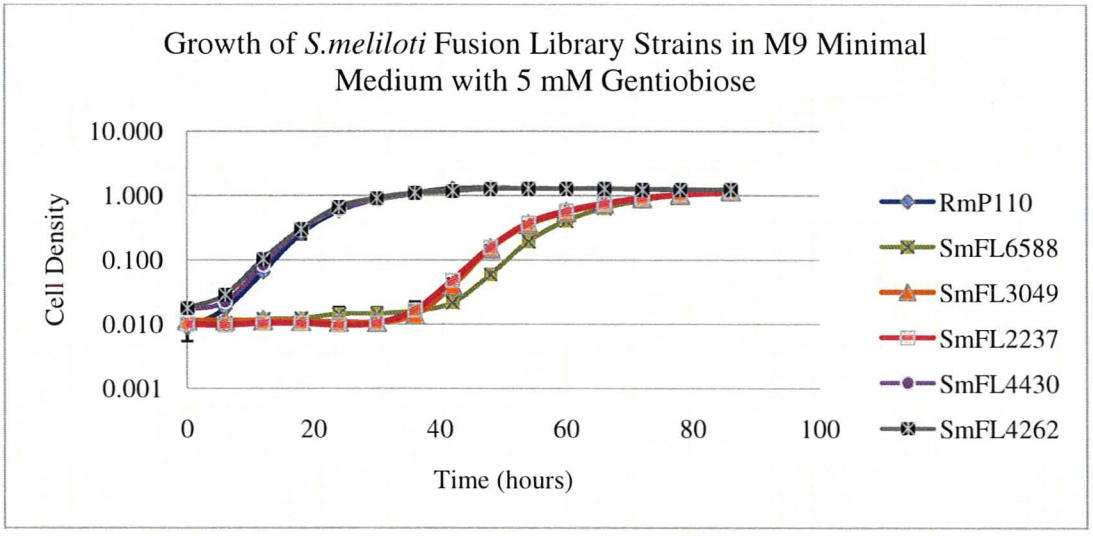


Figure 6: Growth of *S.meliloti* library fusion strains in M9 minimal medium with 5 mM of gentiobiose as a sole source of carbon. Cell density is log of OD600.

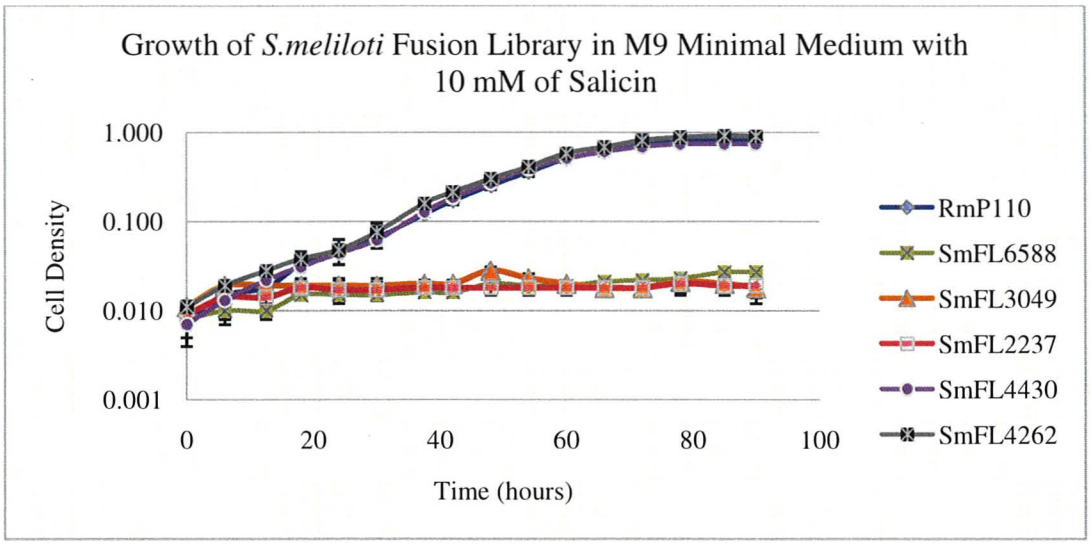


Figure 7: Growth of *S.meliloti* library fusion strains in M9 minimal medium with 10 mM of salicin as a sole source of carbon. Cell density is log of OD600.

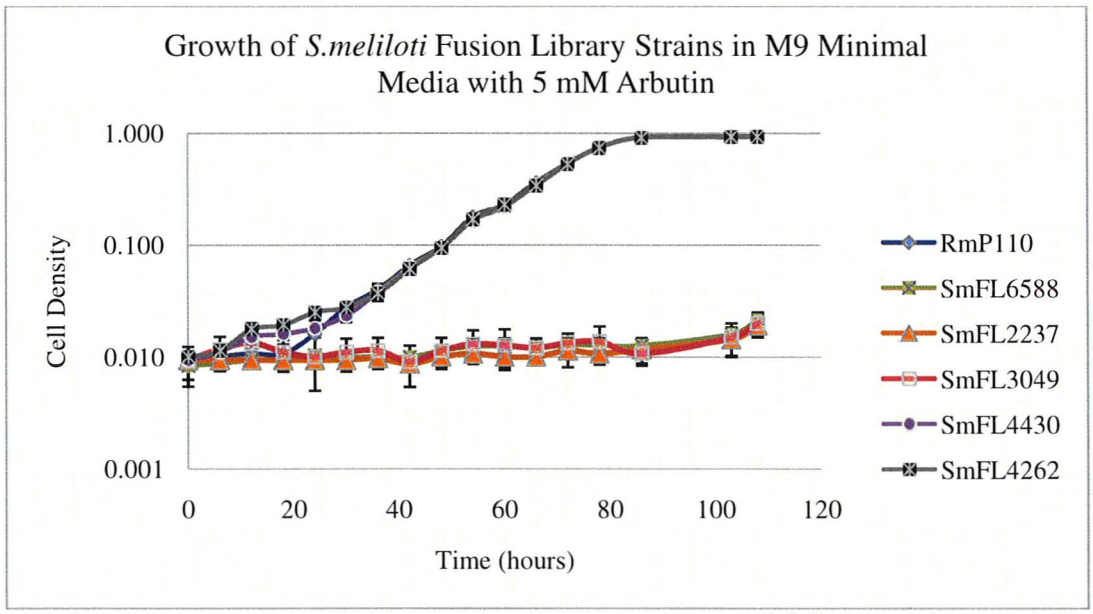
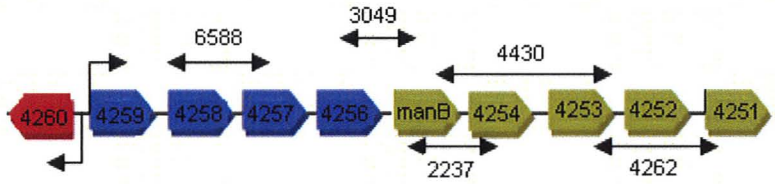


Figure 8: Growth of *S.meliloti* library fusion strains in M9 minimal medium with 5 mM of arbutin as a sole source of carbon. Cell density is log of OD600.

Table 4: Summary of growth of the library fusion strains with the selected sources of carbon. Data obtained from Figure 4, Figure 5, Figure 6, Figure 7 and Figure 8. +++ = growth comparable to wild type RmP110, ++ =growth much slower than the wild type RmP110, + = growth is very minimal compared to wild type RmP110, - = no growth. The figure shows the *smc04260-smc04251* gene cluster where the location of each fusion is indicated. The location at which the fusion has been inserted will be transcribed at the gene where it is inserted, however all the genes downstream from it will not be transcribed.



Carbon source	RmP110	SmFL6588	SmFL3049	SmFL2237	SmFL4430	SmFL4262
Glucose	+++	+++	+++	+++	+++	+++
Cellobiose	+++	++	++	++	+++	+++
Gentiobiose	+++	+	+	+	+++	+++
Arbutin	+++	-	-	-	+++	+++
Salicin	+++	-	-	-	+++	+++

Discussion

As demonstrated in Figure 4, all the gene fusion strains grew as well as the wild type with glucose as a sole carbon source. However, when cellobiose was used as the sole source of carbon, only SmFL4262 and SmFL4430 grew as well as the wild type. Since these strains have a fusion to the genes at the very end of the proposed *smc04259* –

smc04252 operon, it is evident that those genes are not essential for growth on cellobiose. The rest of the strains, SmFL3049, SmFL2237 and SmFL6588 all grow at a similar rate, which was much slower than the wild type. The wild type reached the stationary phase after growth for 30 hours, while the strains SmFL3049, SmFL2237 and SmFL6588 only reach it after growth for about 80 hours. This eventual increase in cell density may also be attributed to a reversion of strains back to the wild type phenotype. Strain SmFL3049 has a fusion to *manB* and therefore all the genes downstream of *manB* were not transcribed. Given the wild type growth of strains SmFL4430 and SmFL4262, it appears that some or all of the *manB* to *smc04254* genes are important in the metabolism of cellobiose. Strain SmFL6588 grew slightly slower than all the other strains, which makes sense since the fusion of pTH1522 was made to *smc04257*, which is annotated as a component of an ABC transporter system (Figure 42). Strain SmFL2237 had a pTH1522 fusion to *smc04254* and grew just as well as strain SmFL3049, which had the fusion to *manB*. This may suggest that the lack of growth in these two strains would be attributed to absence of *smc04254* as opposed to *manB* alone, since SmFL2237 grew just as poorly as SmFL3049, even though *manB* was transcribed in SmFL2237.

When gentiobiose was used as the sole source of carbon, the results were similar to those with cellobiose as the carbon source, although the lack of growth was much more evident. Again, SmFL4430 and SmFL4262 grew as well as the wild type, suggesting that *smc04253* and genes downstream of *smc04253* are not essential for the growth on gentiobiose. Much slower growth of SmFL2237, SmFL3049 and SmFL6588 can be attributed to a fusion of pTH1522 to genes that are important in gentiobiose breakdown

(see Figure 2). The growth on gentiobiose of these three strains was much more affected than growth on cellobiose. As shown in Figure 41, gentiobiose and cellobiose molecules only differ by the type of the bond between two glucose units; in case of cellobiose, the glycosidic bond connects two glucoses in β-(1-4) configuration, while gentiobiose is connected in β-(1-6) configuration. This suggests that these two sugars are perhaps broken down via two different enzymes; hence the difference in growth with these two sugar sources perhaps suggests different metabolic pathways.

When salicin and arbutin were used as the source of carbon again both SmFL4430 and SmFL4262 grew as well as the wild type, whereas strains SmFL6588, SmFL3049 or SmFL2237 failed to grow. The fact that growth was completely absent for these strains suggests the importance of the *smc04259-smc04252* operon in metabolism of aryl-glycosides. Both salicin and arbutin are aryl-glycosides, which means that the glucose unit (glycone) is connected to an aromatic unit (aglycone) via β-glycosidic bond (Figure 41). It is probable that the operon contains a gene product that is essential for metabolism of these sugars and therefore the absence of one or more genes renders the whole system unable to metabolize these sugars.

Since reversion of library strains back to the wild type phenotype was perhaps affecting the growth of these strains, 100 colonies from each strain were checked for the presence of reversion. The library fusion strains were Gm resistant; the loss of that resistance can be attributed to the reversion back into wild type phenotype. The reversion of library fusion strains to the wild type phenotype would result from excision of

pTH1522 from the genome. The rate of reversion can be used to understand how reliable the results are. To screen these strains single colonies from each LB plate were patched on LB/Sm and LB/SmGm. The strains that grew very well, SmFL4262 and SmFL4430, had very small reversion rates of zero to 25%. Strains SmFL3049, SmFL2237 and SmFL6588 showed reversion rates from 30%-50%, which shows that probably the later increase in growth of these strains can be attributed to these reversions. On the other hand, when reversion rates were taken from Arbutin and Salicin, only 4% have reverted. This is probably due to the fact that growth with aryl-glycosides is very slow. If it was allowed to proceed further, the reversion rate would no doubt increase, therefore allowing for growth in that sugar source.

In-frame Gene Deletions

Introduction

As the *smc04259-smc04252* appear to be transcribed as a single transcript from a promoter upstream of *smc04259* (Figure 51), insertion mutations will be polar on the downstream genes. To investigate the function of individual genes it was necessary to generate deletion mutations that were not polar on downstream genes. Accordingly we employed a method of Datsenko and Wanner, 2000 to delete the *smc04259, smc04255/manB* and *smc04254* genes. Upon the deletion of each gene, an FRT scar will form of about 85 nucleotides. This scar should not interfere with the transcription of the genes downstream from it.

The *smc04259* gene is annotated as putative periplasmic binding protein, a component of an ABC transporter system. The periplasmic binding protein binds specific solutes in the periplasm of Gram negative bacteria and delivers it to the ABC transporter (Davidson, Dassa, Orelle, & Chen, 2008). The strain RmP2228 has *smc04259* deleted in frame, and its scar sequence is shown in Figure 9. The preparation of this strain is described in Materials and Methods. See Figure 46 in the Appendix for a detailed procedure for deletion of this gene. When double homologous recombination was carried out, the frequency of recombination was 5.70×10^{-6} per donor and 8.26×10^{-6} per recipient. Out of 300 colonies screened for the absence of pUCP30T vector, only 4 colonies underwent double homologous recombination, which rendered them Gm sensitive.

Another gene that was deleted in frame was *smc04255/manB*, which is annotated as a putative β-mannosidase protein. As it was mentioned previously, these enzymes are characterized by their ability to hydrolyze β-glycosidic bond between two mannose units, in sugars such as mannobiose. When double homologous recombination was carried out, the frequency of recombination was 1.99×10^{-6} per donor and 7.78×10^{-7} per recipient. Out of 167 colonies tested for the absence of pUCP30T plasmid that was integrated into *S.meliloti* chromosome, only 3 demonstrated that. The strain RmP2227 has a deletion of *manB* gene, and its final scar sequence is shown in Figure 10. See Figure 46 for detailed procedure for deletion of this gene.

The *smc04254* gene which is annotated as hypothetical conserved protein with similarity to xylose-isomerases, was also deleted. Strain RmP2647 has an in-frame

deletion of *smc04254* and its scar sequence is shown in Figure 11. For detailed procedure describing deletion of this gene see Figure 47 in the Appendix. When double homologues recombination was carried out, the frequency of recombination was 3.40×10^{-4} per donor and 1.89×10^{-5} per recipient. Out of 17 colonies that were patched to check for the absence of pTH2636 from *S.meliloti* chromosome, only 2 had that feature.

Another strain that was constructed and used in the growth curves as a negative control along with the other strains is RmP2202. This strain was constructed using library fusion strains and FRT sites in order to delete a region from 2088466 to 2077096 in *S.meliloti* chromosome. This region includes the genes *smc04259* to *smc04252*. This strain was designated as Rmp2202 and is used alongside the deletion strains. See Figure 48 in the Appendix for detailed procedure for making of strain RmP2202.

To ensure the growth phenotypes of deletion mutants are due to the deletion of that gene and not any other mutation, the DNA sequence of the deletion mutation was studied. The sequences of the strains that were prepared were checked at each step via PCR and DNA sequencing of the genomic DNA of each strain. The final sequences were amplified from genomic DNA prepared from each strain.

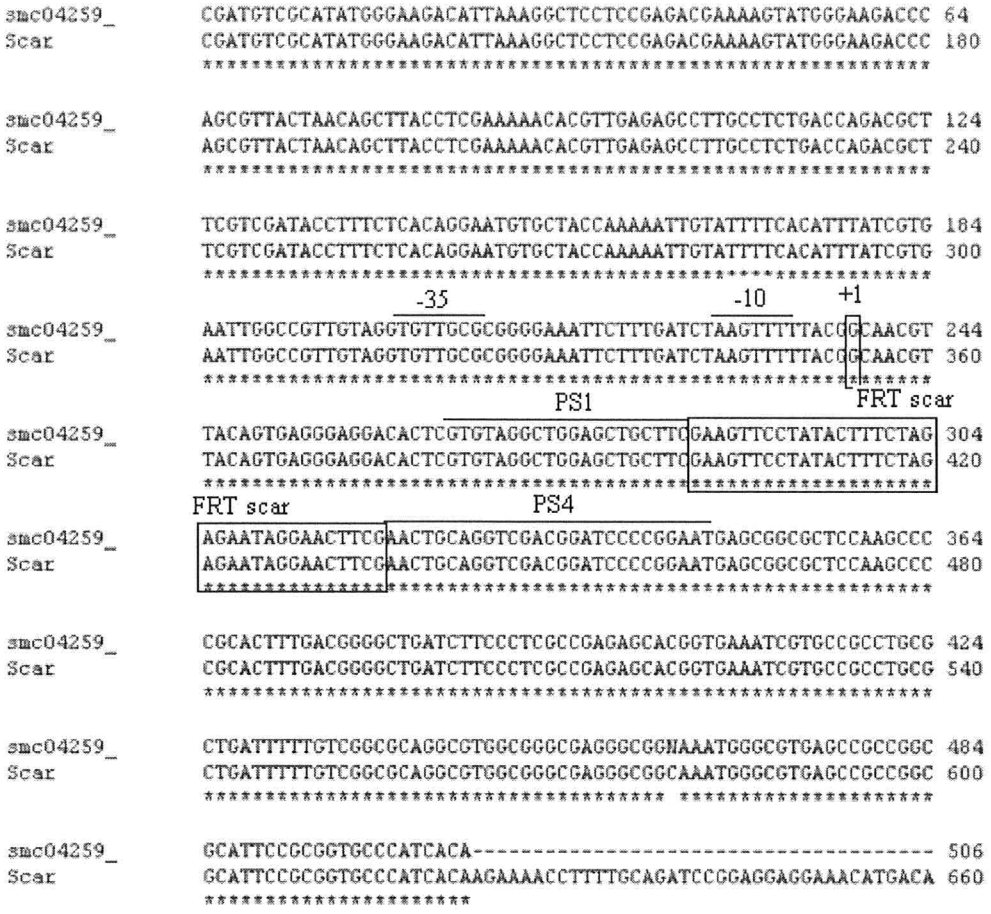


Figure 9: Scar sequence of genomic DNA from strain RmP2228, deletion of *smc04259*. The sequence was obtained by sequencing the genomic DNA of the final strain RmP2228. The promoter was previously mapped by Dr. C. White (Figure 51), and is indicated here. PS1 is priming site 1, PS4 is priming site 4. These priming sites were used for amplification of FRT sites off of template plasmid pKD13 (Datsenko & Wanner, 2000). FRT scar is indicated in a box.

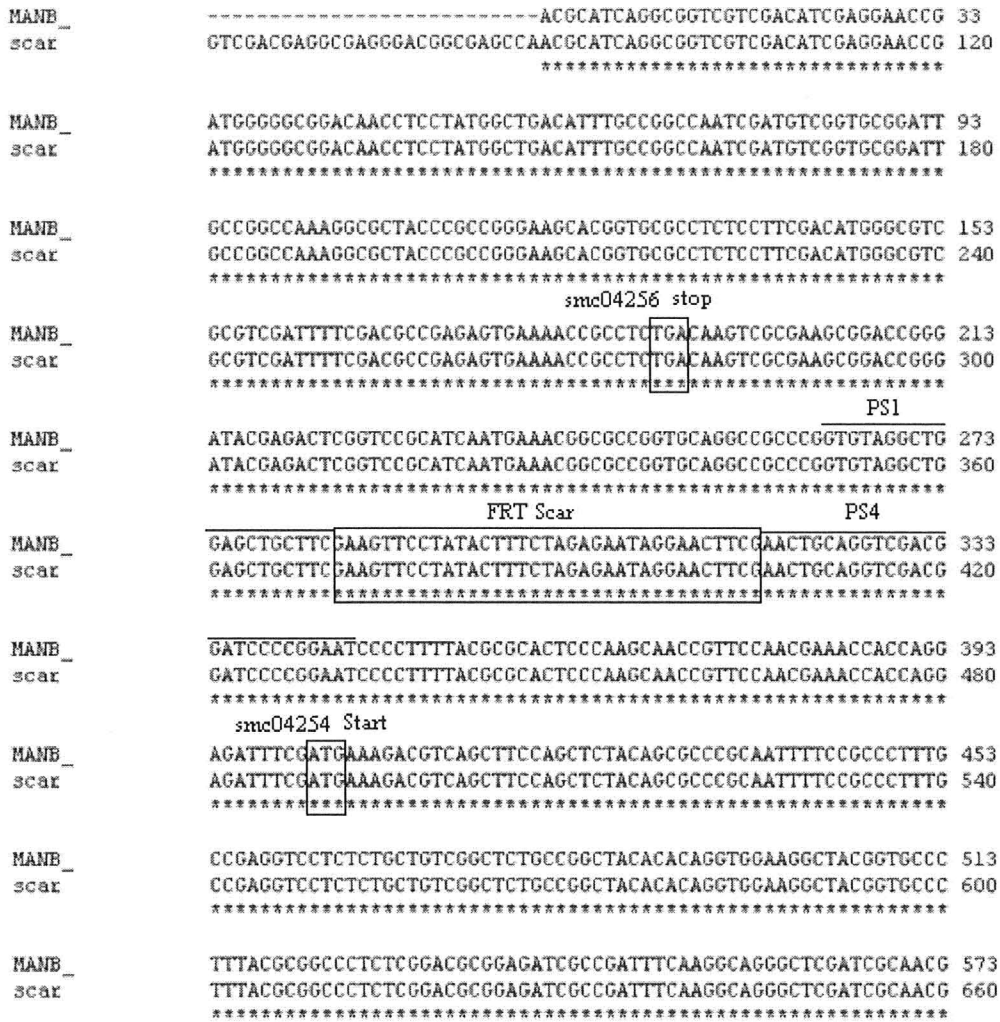


Figure 10: Scar sequence of genomic DNA from strain RmP2227, deletion of *manB*. This sequence was obtained by sequencing genomic DNA of RmP2227. PS1 is priming site 1, PS4 is priming site 4. These priming sites were used to initially amplify the FRT sites off of template plasmid pKD13 (Datsenko & Wanner, 2000). The stop codon for *smc04256* and start codon for *smc04254* are also shown. FRT scar is shown in a box.

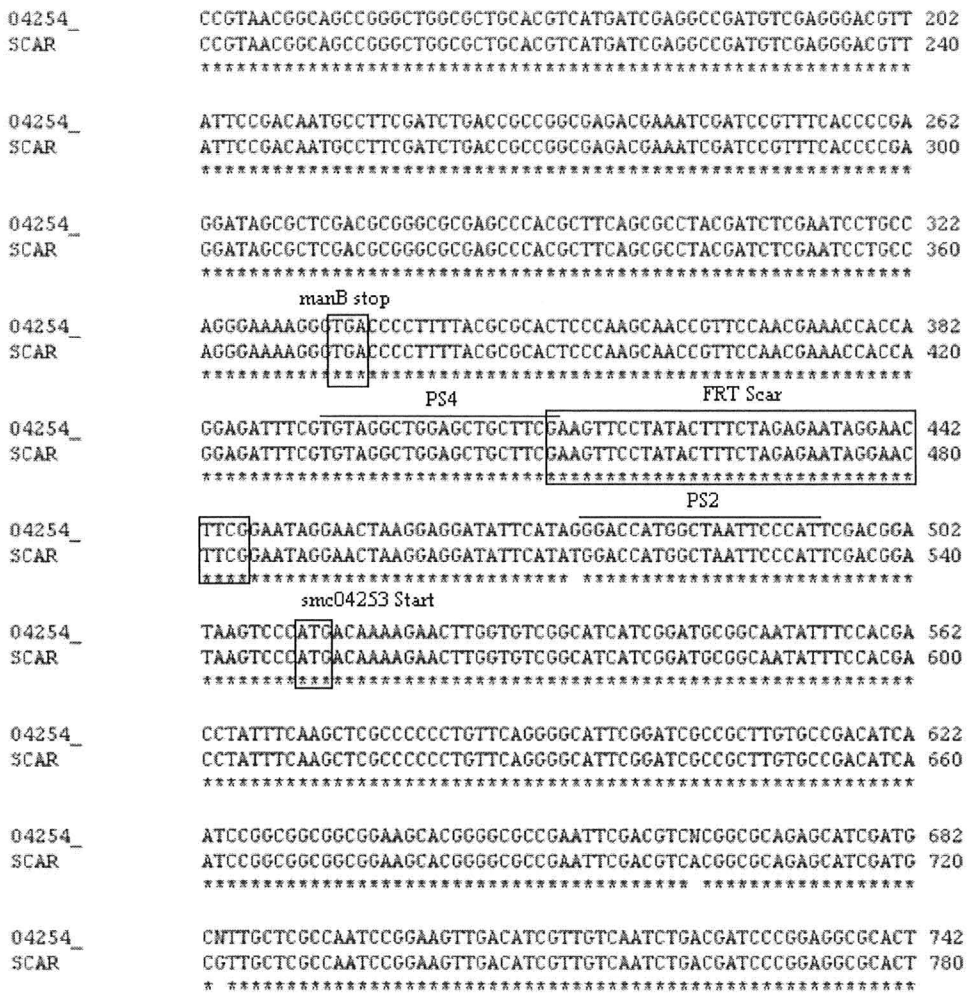


Figure 11: Sequence of genomic DNA from strain RmP2647, deletion of *smc04254*. This sequence was obtained by sequencing genomic DNA of RmP2647. PS2 is priming site 2, PS4 is priming site 4. These priming sites were used to initially amplify the FRT sites off of template plasmid pKD4 (Datsenko & Wanner, 2000). The stop codon for *manB* and start codon for *smc04253* are also shown. FRT scar is shown in a box.

Results

The strains RmP2228, RmP2227 and RmP2647 were tested for their growth phonotypes along with negative control RmP2202 and wild type strain RmP110. The growth of these strains in M9 minimal media supplemented with glucose, cellobiose, gentiobiose, mannobiose, salicin and arbutin would reveal the effect the deletion of each gene has on metabolism and growth of *S. meliloti*.

The rate of *S. meliloti* growth in M9 minimal medium with each carbon source is different and this difference must be taken into consideration while comparing each growth curve. Table 5 and Figure 12 compare the difference in growth of the wild type strain with all the different sugars used in this study.

Table 5: Doubling time of wild type *S.meliloti* strain RmP110 in M9 minimal media with each sugar as a sole source of carbon. The data was obtained from Figure 12. The mean growth rate constant k was calculated using the equation $k=(\log N_t - \log N_o)/0.301$, where N is the number of cells measured at OD600 at the start and at the end of measurement. Mean generation time g was calculated per hour using equation $g= 1/k$. g is the doubling time, which is the time it takes for the cells to reach double the number of the initial number.

Carbon Source	Doubling time
10 mM Glucose	3.7
5 mM Cellobiose	3.9
5 mM Gentiobiose	3.7
5 mM Mannobiose	4.3
10 mM Salicin	8.9
5 mM Arbutin	10.3

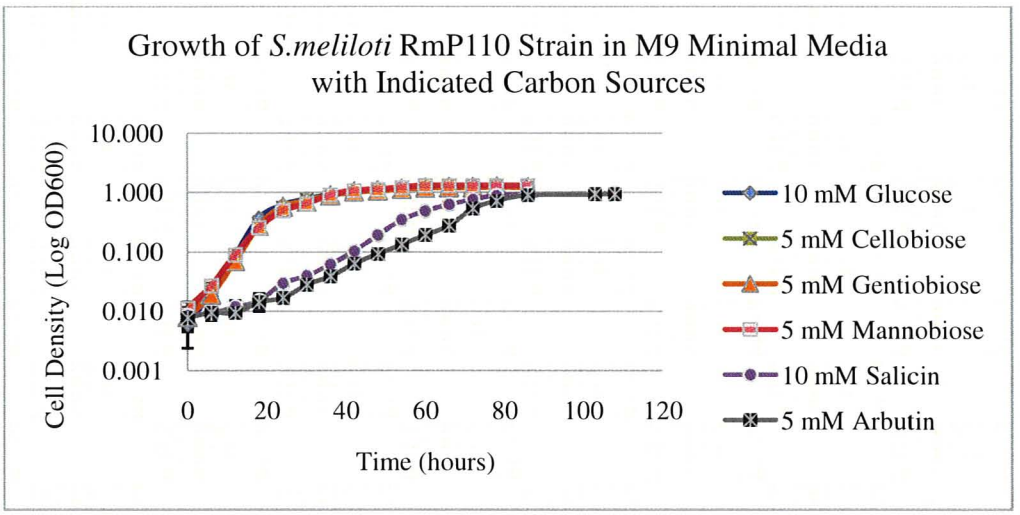


Figure 12: Growth rate of wild type *S. meliloti* strain RmP110 in M9 minimal media with an indicated sugar as a sole source of carbon. Growth with arbutin and salicin is measured over a longer time period as the cell doubling time is longer.

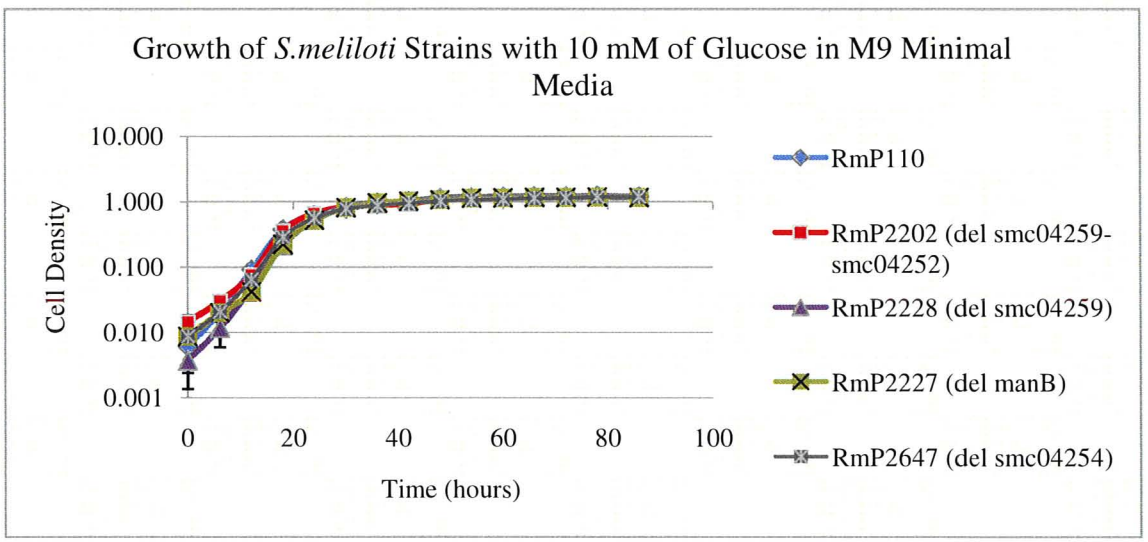


Figure 13: Growth of *S.meliloti* deletion strains in M9 minimal media with 10 mM glucose as a sole source of carbon. RmP2202 = deletion of genes *smc04259* to *smc04252*, inclusive; RmP2228 = deletion of *smc04259*; RmP2227 = deletion of *smc04255/manB*; RmP2647 = deletion of *smc04254*. Cell density is log of OD600.

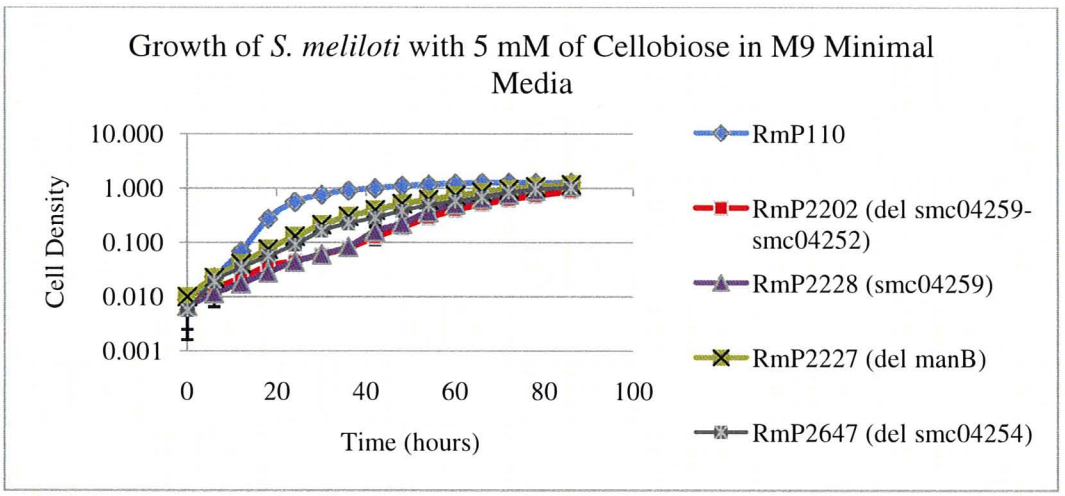


Figure 14: Growth of *S.meliloti* deletion strains in M9 minimal media with 5 mM of cellobiose as a sole source of carbon. RmP2202 = deletion of genes *smc04259* to *smc04252*, inclusive; RmP2228 = deletion of *smc04259*; RmP2227 = deletion of *smc04255/manB*; RmP2647 = deletion of *smc04254*. Cell density is log of OD600.

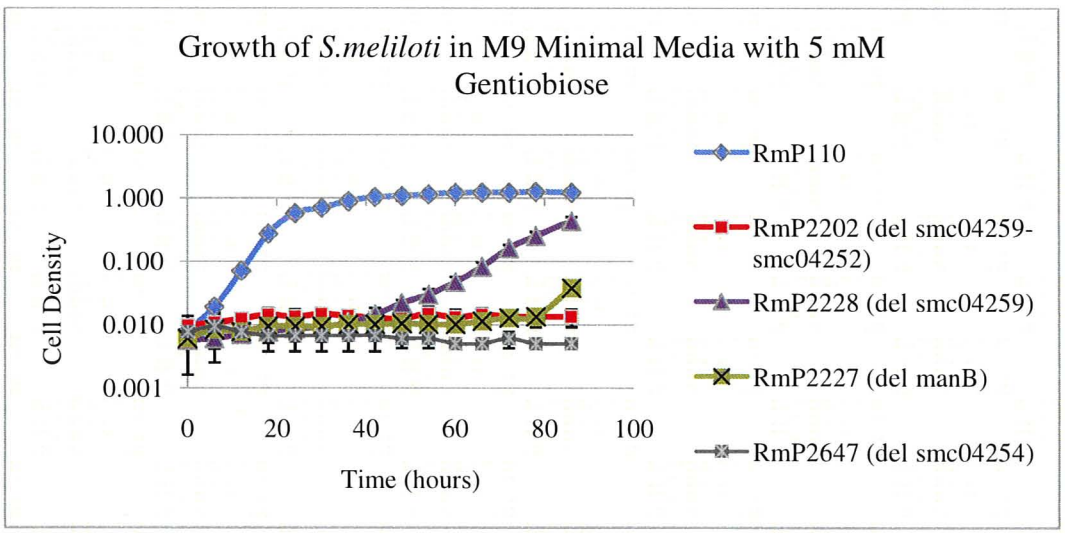


Figure 15: Growth of *S.meliloti* deletion strains in M9 minimal media with 5 mM of gentiobiose as a sole source of carbon. RmP2202 = deletion of genes *smc04259* to *smc04252*, inclusive; RmP2228 = deletion of *smc04259*; RmP2227 = deletion of *smc04255/manB*; RmP2647 = deletion of *smc04254*. Cell density is log of OD600.

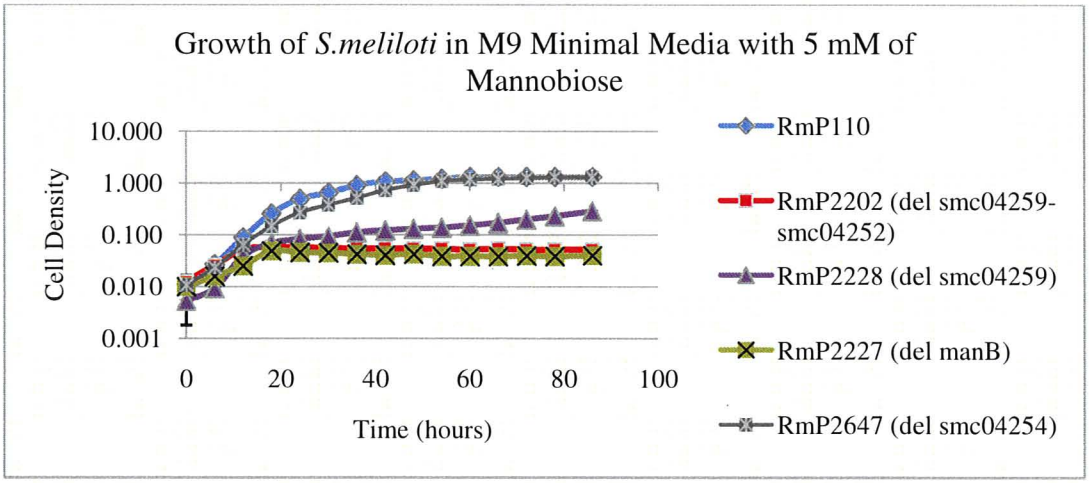


Figure 16: Growth of *S.meliloti* deletion strains in M9 minimal media with 5 mM of mannobiose as a sole source of carbon. RmP2202 = deletion of genes *smc04259* to *smc04252*, inclusive; RmP2228 = deletion of *smc04259*; RmP2227 = deletion of *smc04255/manB*; RmP2647 = deletion of *smc04254*. Cell density is log of OD600.

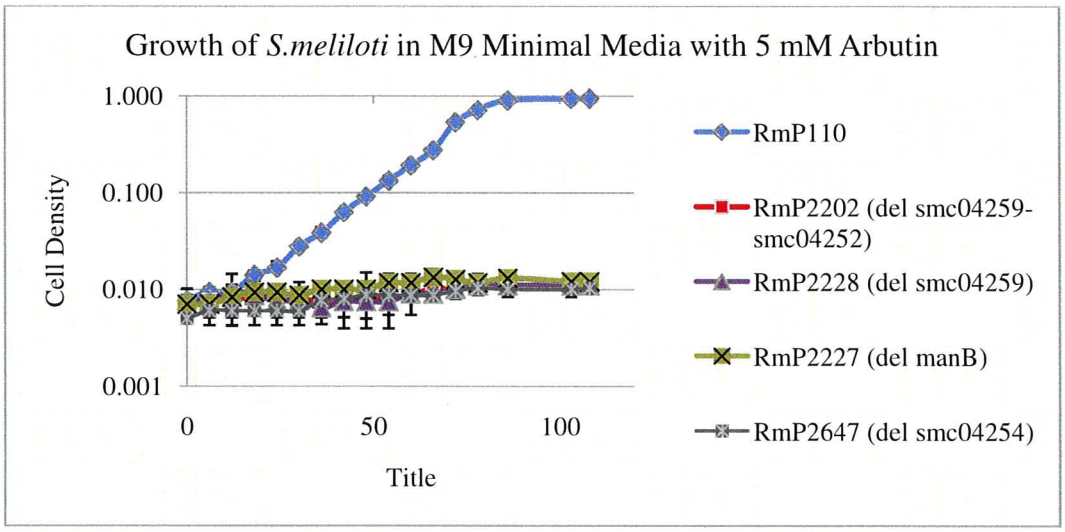


Figure 17: Growth of *S.meliloti* deletion strains in M9 minimal media with 5 mM of arbutin as a sole source of carbon. RmP2202 = deletion of genes *smc04259* to *smc04252*, inclusive; RmP2228 = deletion of *smc04259*; RmP2227 = deletion of *smc04255/manB*; RmP2647 = deletion of *smc04254*. Cell density is log of OD600.

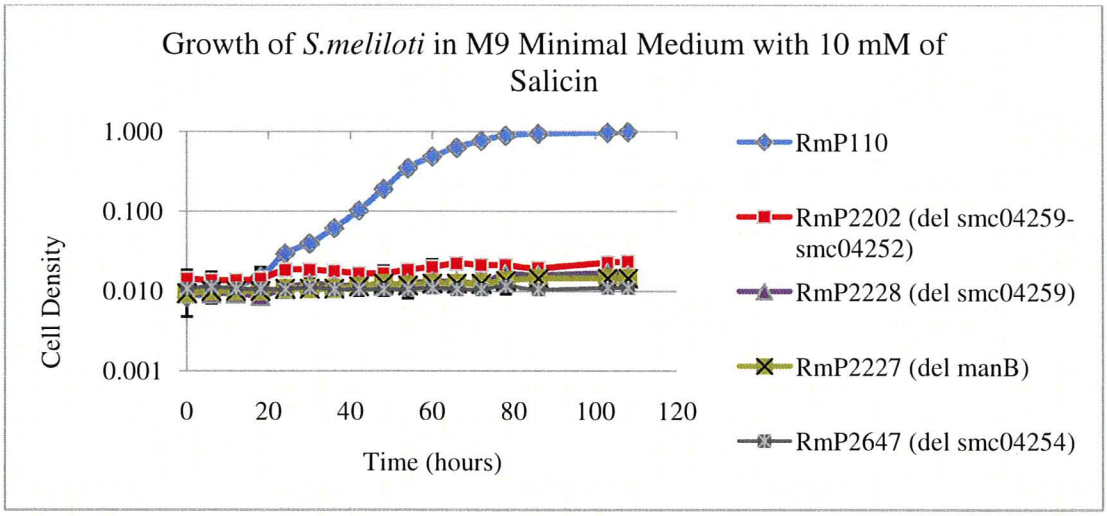
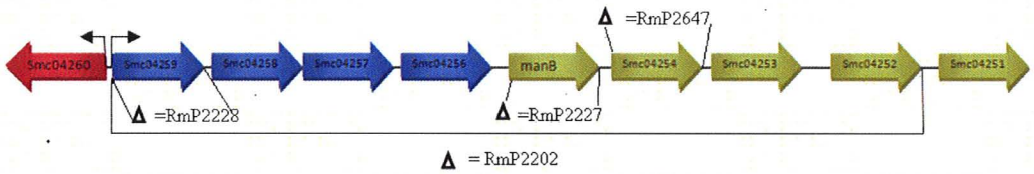


Figure 18: Growth of *S.meliloti* deletion strains in M9 minimal media with 10 mM of salicin as a sole source of carbon. RmP2202 = deletion of genes *smc04259* to *smc04252*, inclusive; RmP2228 = deletion of *smc04259*; RmP2227 = deletion of *smc04255/manB*; RmP2647 = deletion of *smc04254*. Cell density is log of OD600.

Table 6: Summary of growth phenotypes of the polar deletion strains with the selected sources of carbon. The data was obtained from Figure 13, Figure 14, Figure 15, Figure 16, Figure 17 and Figure 18. +++ =growth comparable to wild type strain RmP110, ++ = growth slower compared to wild type strain RmP110, +=growth is minimal compared to wild type strain RmP110, - = no growth. The figure shows the *smc04260-smc04251* gene cluster where each gene that was deleted is indicated. RmP2202 is a deletion of *smc04259-smc04252*, therefore it shows a range that is deleted. RmP2228 is deletion strain of *smc04259,* RmP2227 is deletion strain of *manB,* RmP2647 is a deletion strain of *smc04254.*



Carbon Source	RmP110	RmP2202	RmP2228	RmP2227	RmP2647
Glucose	+++	+++	+++	+++	+++
Cellobiose	+++	+	+	++	++
Gentiobiose	+++	-	++	-	-
Mannobiose	+++	-	++	-	+++
Salicin	+++	-	-	-	-
Arbutin	+++	-	-	-	-

Discussion

Glucose is an end product of the breakdown of cellobiose, gentiobiose and aryl-glycosides, while mannose is the product of mannobiose metabolism. All the strains grow as well as the wild type in M9 minimal media with 10 mM of glucose, suggesting that

failure to grow on these disaccharides resulted from a defect in uptake or hydrolysis to monosaccharide units.

When cellobiose is used as a sole source of carbon, all the strains still grow, although deletion strains grow much slower than the wild type *S. meliloti*. Figure 14 shows that even the deletion of the whole operon in strain RmP2202 (*Δsmc04259-smc04252*) allows for growth with this sugar. It is interesting that strain RmP2228 (Δ*smc04259*, an SBP) grows as well as RmP2202, which suggests that *smc04259* is necessary for growth with cellobiose as a carbon source. On the other hand, RmP2647 (Δ*smc04254*) and RmP2227 (Δ*manB*) grow on cellobiose at a very similar rate, although still slower than the wild type RmP110. It may be that either *manB* or *smc04254* have some minor function in metabolism of cellobiose, although it seems improbable that they would be directly involved in its metabolism. It is possible that cellobiose is broken down via some other system or operon, and may be only transported into the cell via the putative ABC transporter system.

When gentiobiose is used as a sole source of carbon (Figure 15), the growth of deletion strains seem a lot more affected, since the growth proceeds at a much slower rate compared to cellobiose. It is clear that both RmP2647 (Δ*smc04254*) and RmP2227 (Δ*manB*) grow just as poorly as RmP2202, the negative control. It is interesting to see that *smc04254* seems to be involved in metabolism of gentiobiose, and therefore metabolism of β (1-6) glycosidic bond, as opposed to β (1-4) glycosidic bond that is present in cellobiose. What is surprising is that RmP2227 (*ΔmanB*) also appears to grow

very slowly. It would be very unlikely for ManB to be responsible for gentiobiose breakdown, as it is thought to be specific for mannobiose only. Since the results suggest the contrary, it would be important to study this strain further for lack of transcription in the operon following *manB* gene. Lastly, the absence of putative periplasmic binding protein (*smc04259*) results in a slower growth when gentiobiose is used, although that strain does grow eventually. If the periplasmic binding protein is not transcribed due to its deletion, then gentiobiose will not be able to be transported into the cell. Because the strain does grow slowly, it suggests that perhaps some gentiobiose is transported into the cell at a slower rate, perhaps via another transporter system.

When mannobiose is used as a sole source of carbon (Figure 16), it is very clear that RmP2647 (Δ*smc04254*) grows almost as well as the wild type strain RmP110. This suggests that the gene product of *smc04254* is not required for proper metabolism of mannobiose. This again suggests that perhaps *smc04254* is only essential for metabolism involving β (1-6) glycosidic bond. On the other hand, RmP2227 (Δ*manB*) grows at a rate similar to the deletion of the whole putative operon RmP2202. This is not surprising, as ManB is supposed to be only capable of breaking down mannobiose. It is also noteworthy that both RmP2227 and RmP2202 still do grow to some extent, and reach a maximum cell density of 0.050, at which they both level off. This may suggest that mannobiose is perhaps broken down by another enzyme, although not to the same extent and the metabolic capacity is limited. This growth is most likely not due to contaminating simple sugars present in mannobiose, as this sugar has a very high degree of purity. The growth of strain RmP2228 (Δ*smc04259*) is comparable to that of gentiobiose and

cellobiose, suggesting once again that the growth is slowed down by the absence of putative periplasmic binding protein.

The growth of each strain used is very similar in both arbutin and salicin. This is most likely due to the fact that both are aryl-glycosides and have very similar structure, the only difference being an aromatic group attached to the glycone unit (Figure 41). It seems that this operon is necessary for proper growth with these aryl-glycosides. Previously, the absence of *smc04259* resulted in slower growth when cellobiose, gentiobiose and mannobiose were used. It seems that perhaps there is another operon that may contribute to the uptake of these sugars, while aryl-glycosides seem to be completely dependent on the *smc04259 – smc04252* operon. It is again surprising that RmP2227 shows no growth as it would be expected that it will grow just as well as the wild type, since ManB is not thought to break down β (1-4) glucosidic bond (bond between two glucose units). But consistent with the results obtained from growth with gentiobiose, it seems to be just as affected as growth of RmP2202.

Conclusion

The in-frame deletions demonstrated the distinct growth phenotype when grown with each different sugar. It seems that the results are in agreement with some of the findings from the library fusions strains. It is clear that cellobiose is not at all dependant on the putative *smc04259 – smc04252* operon that is studied here, although its absence does cause a slower growth phenotype. On the other hand, when grown with gentiobiose, the absence of either *smc04254* or *manB* results in a strain that fails to grow with this sugar.

The results for growth in mannobiose are interesting in that it seems that when this sugar is used, absence of *smc04254* has no effect on the growth, like it does when cellobiose, gentiobiose, and aryl-glycosides are used. Aryl-glycosides are dependent on each gene that was deleted for proper growth. This suggests that *smc04259* may be a periplasmic binding protein that is specific for certain sugars, especially aryl-glycosides. *manB* may be somehow involved in metabolism of all the sugars used, especially mannobiose. *smc04254* seems to be necessary for metabolism of every sugar except mannobiose.

CHAPTER 4

ManB Protein Purification and Activity Assays

Introduction

Previous expression studies showed that growth of *S.meliloti* with sugars such as cellobiose, gentiobiose and aryl-glycosides resulted in expression of the *smc04260-smc04251* genes annotated as putative ABC transport genes and other genes including ManB. This suggested that perhaps ManB was either a β-glucosidase, and was able to break down cellobiose, gentiobiose, etc., or that enzyme had a wider substrate range than was originally anticipated, and therefore would be able to break down sugars such as mannobiose as well. On the other hand, the sequence of ManB (Figure 43) showed that this enzyme had about 30% similarity to already known enzymes of family 2 of the glycosyl hydrolases, which were strictly mannosidases, and therefore would have no activity on sugars such as cellobiose, gentiobiose, salicin or arbutin. Therefore it is unclear what the function of ManB really is. Since it is associated with the other genes that are thought to be present in a gene cluster responsible for metabolism of β-glucosides, it would be easy to assume that ManB also has a similar function. Its sequence however, suggests otherwise. If ManB is not capable of breaking down cellobiose, gentiobiose and aryl-glycosides, then this activity must be present in another protein in the *S.meliloti* genome. None of the other gene annotations in the *smc04260-smc04251* cluster suggest that it is one of these genes. Following ManB purification the function of ManB protein was examined using two assays. One assay used p-nitrophenyl

linked substrates where the ability of ManB to distinguish between glucosides and mannosides was examined. In another assay, glucose oxidase was used to measure the activity of purified ManB with different sugars.

Purification of ManB

Following lysis of cells in a French press, ManB protein was found to be present in the soluble protein fraction. It was known from crude cell extracts obtained earlier that ManB protein is soluble and therefore should be present in the supernatant. Both supernatant and pellet were loaded onto SDS-PAGE gel to find out whether the protein of the right size will be present. ManB is 839 amino acids (with the Strep Tag included) and should be about 94,160 Daltons or 94 kDa (Gasteiger, Gattiker, Hoogland, Ivanyl, Appel, & Bairoch, 2003). ManB was overexpressed from plasmid constructs both in *E.coli* BL21 (M1436) and *S.meliloti* RmP1934. Both M1436 and RmP1934 strains have a plasmid construct pTH2493, which overexpresses ManB protein. ManB also contains a StrepII fusion tag at the C-terminus. Figure 19 shows an SDS - PAGE gel where both *E.coli* and *S.meliloti* strains that are able to overexpress ManB protein have been loaded side by side. *S.meliloti* strain RmP1934 is able to express the protein in a high amount, and so it was chosen for purification. Initially, the purification was carried out with only 2 mL of supernatant, in order to test the method and ensure everything was working well. Each fraction from the purification was loaded onto the SDS-PAGE gel to make sure Man B was present after the elution step and not in the flow through or any of the wash steps (Figure 20 and Figure 21).

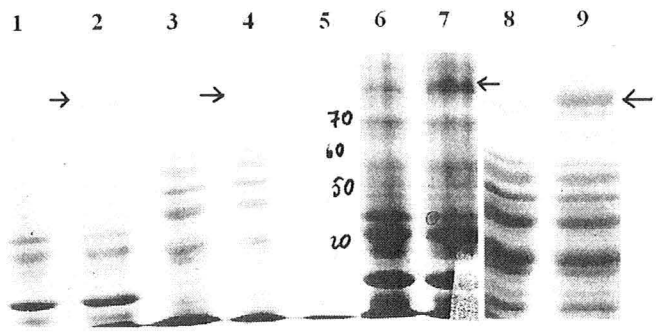


Figure 19: SDS-PAGE gel showing ManB protein overexpressed in both *E.coli* and *S.meliloti* whole cell pellets. Lane 1: Uninduced whole cells from M1436, 10 µL loaded. Lane 2: Induced whole cell pellet from M1436, 10 µL loaded Lane. 3: Uninduced whole cell pellet from M1436, 20 µL loaded. Lane 4: induced whole cell pellet from M1436, 20 µL loaded. Lane 5: Protein marker. Lane 6: Uninduced whole cell pellet from RmP1934, 10 µL loaded. Lane 7: Induced whole cell pellet from RmP1934, 10 µL loaded. Lane 8: Uninduced whole cell pellet from RmP1934, 20 µL loaded. Lane 9: Induced whole cell pellet from RmP1934, 20 µL loaded.

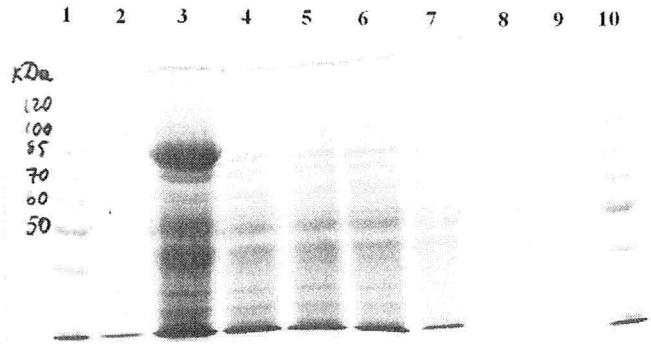


Figure 20: SDS-PAGE gel showing test purification of ManB protein from *S.meliloti* strain RmP1934. Each purification fraction was collected in 200 µL aliquots, 10 µL of which was loaded on a gel. Lanes 1 and 10: protein ladder. Lane 2: Cell pellet following French press passage. Lane 3: Supernatant following French press passage. Lanes 4, 5, 6: flow through collected, Lanes 7, 8, 9: washing fractions.

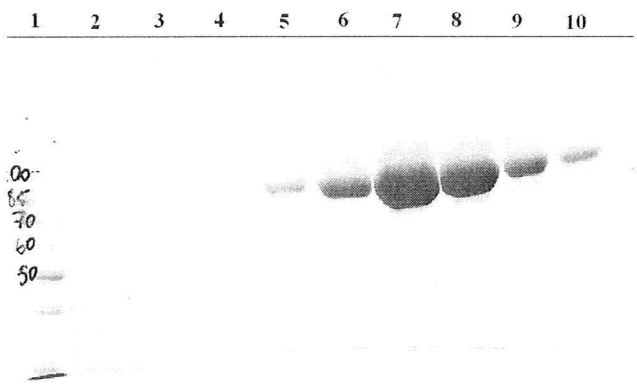


Figure 21: SDS-PAGE gel showing test purification of ManB – continued from Figure 20. ManB was eluted from the Strep Tacin column (1 mL of bed volume with 2.5 mM of desthiobiotin in Buffer W). Lane 1: protein ladder, Lanes 2 though Lane 10 are all elution steps. Each elution fraction was collected in 200 µL aliquots and 10 µL of each fraction was loaded on a gel. ManB is eluted towards the middle of the elution fractions. It is present in high concentration in aliquots 5, 6 and 7 (Lanes 6, 7, 8 respectively).

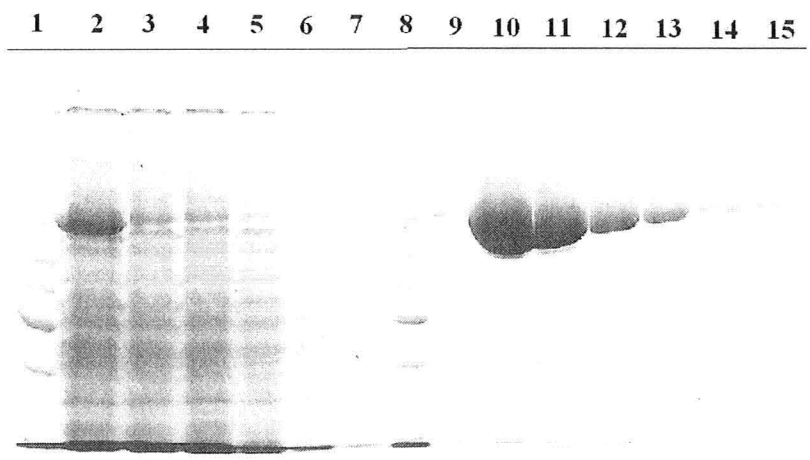


Figure 22: The purification profile of ManB from *S.meliloti* strain RmP1934. 7 µL of each aliquot is loaded. Since the test purification worked well, the rest of the supernatant was used to purify all of ManB. Lane 1, 8: Protein Ladder. Lane 2: Supernatant portion after French press passage. Lanes 3 and 4: flow through. Lane 5, 6, 7: washing fractions. Lane 9 through 15: Fractions eluted upon the addition of 2.5 mM desthiobiotin in 200 µL portions. ManB is visible at high concentrations in most elution fractions.

As shown in Figure 21 for the test purification, the protein was eluted only when the elution buffer (100 mM Tris, 150 mM NaCl, 10 mM EDTA, 1 mM DTT, 2.5 mM desthiobiotin) was passed through the column, and there was none present in the washing buffer or the flow through. Elution fractions 5 to 8, of 200 μL each, contained the highest amount of protein. The Bradford assay revealed fractions 4 to 8 contained 0.4, 1.7, 3.3, 0.8 and 0.7 mg protein/mL respectively. It is clear that ManB is purified in a good yield and therefore the purification on a larger scale was pursued. A large scale purification of ManB was done on an IPTG-induced 500 mL culture of *S.meliloti* (Figure 22).

Eluted ManB was present at 2.6, 2.5 and 0.8 mg/mL (Figure 22, lanes 10, 11 and 12) in Buffer W containing 5% glycerol. Fraction 2 was dialized in a buffer containing 50% glycerol in buffer with 100 mM Tris, 150 mM NaCl and 10 mM EDTA. DTT was not added to the buffer, as reducing agents are known to inhibit glucose oxidase (Bergmeyer & Bernt, 1974).

Activity of ManB with P-Nitrophenyl compounds

The ability of ManB to hydrolyze the 4-nitrophenol-sugar derivatives 4-nitrophenyl-β-D-mannopyranoside and 4-nitrophenyl-β-D-glucopyranoside was determined (Figure 49). 4-Nitrophenol absorbs at 405 nm, and the rate of its production is proportional to the rate of breakdown of either 4-nitrophenyl-β-D-mannopyranoside (NM) or 4-nitrophenyl-β-D-glucopyranoside (NG). These products can also be reliably measured at 350 nm at different pH values, as indicated below. The assays distinguish β-mannosidase from β-glucosidase activities.

The assays were carried out at pH values ranging from pH5 to pH 9. The absorbtion of p-nitrophenol is influenced by the pH, as it absorbs differently in its acidic form at lower pH, and anionic form in alkaline solution (Biggs, 1954). Since the suggested wavelength for measurement of absorbtion is 405 nm, this wavelength will be more enhanced in solutions of pH 7 and higher, as pKa of p-nitrophenol is 7 (Bergmann, Rimon, & Sigal, 1958). In order to avoid this discrepancy, the rates of reaction at different pH were measured at isosbestic point of 350 nm, which is constant at every pH value and should produce an accurate absorbance value throughout the whole range of pH.

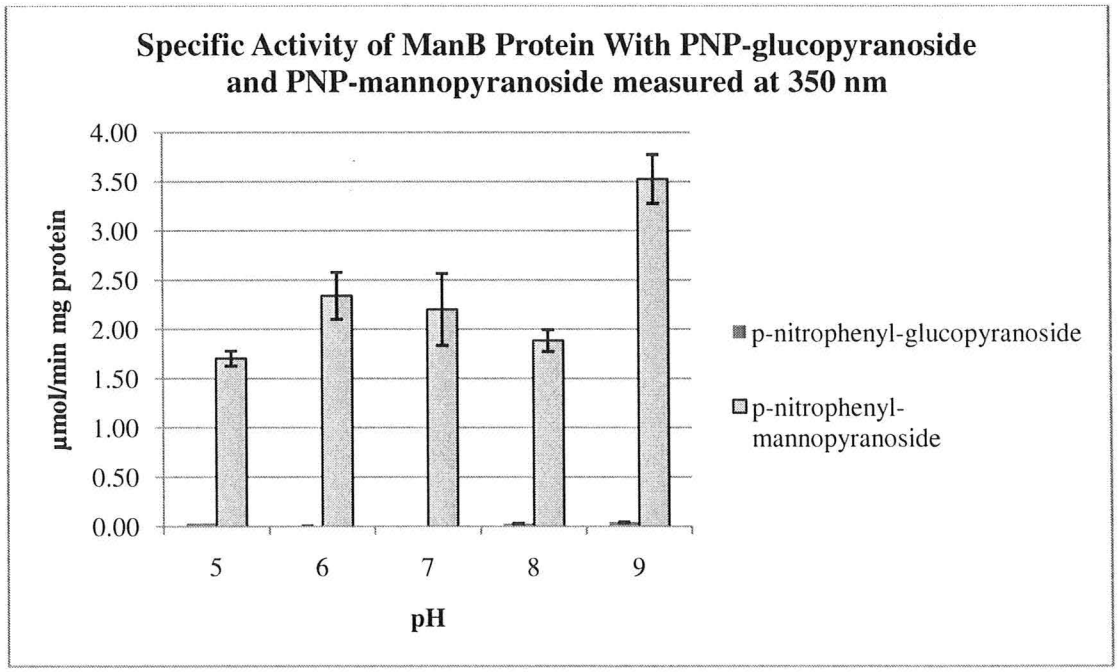


Figure 23: Specific activity of ManB when tested with two different substrates at 1 mM, p-nitrophenyl – glucopyranoside and p-nitrophenyl – mannopyranoside. ManB shows very good activity with PNP-mannoside and has almost no activity on PNP-glucopyranoside.

It is interesting to note, that as the pH increases ManB activity also increases. It appears that the highest pH for its activity is 6, 7 and 9 with specific activity of 2.34 µmol/min mg, 2.20 µmol/min mg and 3.53 µmol/min mg respectively. The different pH values at which ManB has high activity could be a result of two different pKa values for it. At pH 5 ManB had an activity of 1.70 µmol/min mg and at pH 8 its activity was 1.88 µmol/min mg.

When p-nitrophenyl-β-D-glucopyraoside was used as a substrate, the activity was not present. From this data it is clear that ManB is not able to break down glucopyranosides, which suggests that ManB is strictly a mannosidase and is only able to cleave sugars that have mannose at the reducing end of the molecule. The inability to metabolize p-nitrophenyl-β-D-glucopyraoside also shows that in order to break down sugar such as cellobiose, gentiobiose, salicin and arbutin, another enzyme must be present. Since there is no enzyme that is suspected to carry out this function in gene cluster from *smc04259* to *smc04251*, there must be another gene present elsewhere in the *S. meliloti* genome that is able to perform this action. It is also possible that ManB may be able to hydrolyze these substrates along with another enzyme.

Glucose Oxidase Assay for ManB Activity

To determine if ManB has any activity with cellobiose, gentiobiose, mannobiose, salicin or arbutin as substrates, it was important to determine whether ManB could release glucose or mannose (where appropriate) for these sugars. For this purpose, glucose oxidase – peroxidase enzyme assay was carried out (Bergmeyer & Bernt, 1974)

as described in Materials and Methods section. The assay is carried out at pH 7, as glucose oxidase is most active at that pH (Bergmeyer & Bernt, 1974). Even though glucose oxidase is recorded to have negligible activity with sugars other than glucose, it was important to assay its activity with mannose as a substrate. Mannose was used for construction of a standard curve along with glucose, except it was used in 10 times larger quantity than glucose to make up for lower detection levels by glucose oxidase.

The assay was measured using o-dianisidine, which is colourless when it is reduced, but forms a coloured complex in its oxidized form. To stabilize this complex 12 N sulfuric acid was added, and the absorbance was measured at 540 nm (Figure 50).

The standard curves for glucose and mannose were done in triplicate, and the average of the three samples was taken in order to calculate the equation of the line. Glucose was used in 0 – 80 µg and mannose was used in 0 – 800 µg.

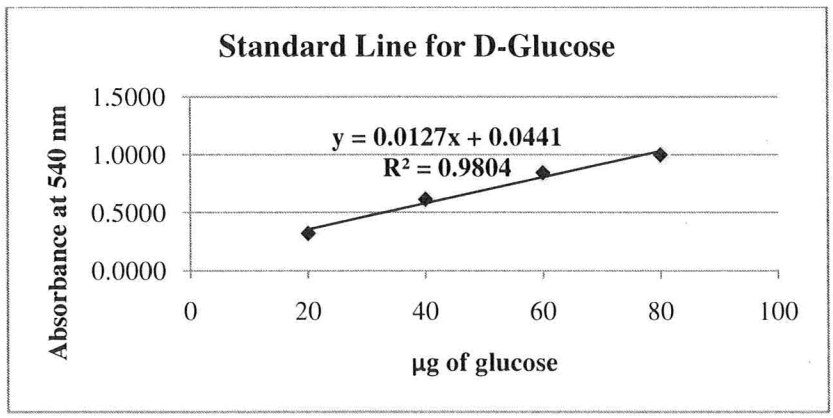


Figure 24: Standard line for measurement and quantification of glucose using glucose oxidase assay. Each point is the average of three samples. The best fit line is shown here, as it was used to calculate the amount of glucose after each reaction.

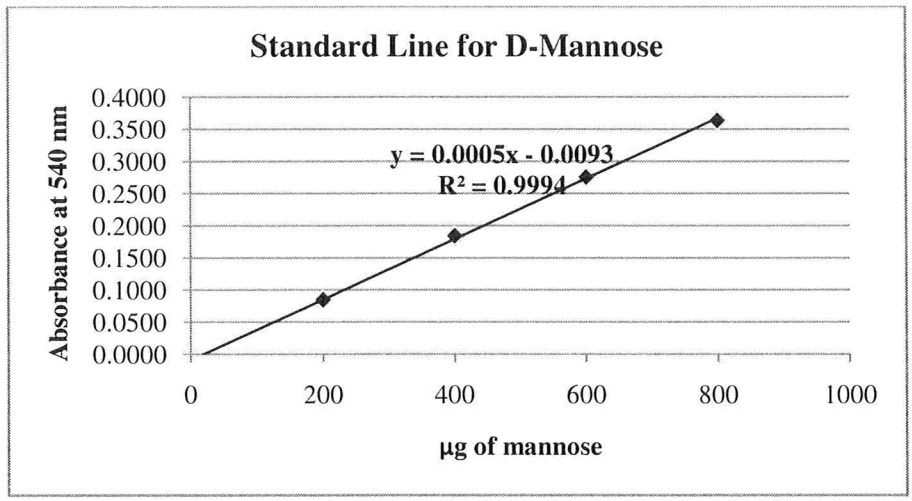


Figure 25: Standard line for measurement and quantification of D-mannose. The numbers are the average of the three samples and the best fit line is shown.

The equations obtained from each standard line were used to calculate the amount of glucose or mannose present in ManB protein reaction mixtures. Only a few time points were used as it was only important to establish that there was some activity and to be able to compare them to each other. In total, the reaction was allowed to proceed for 60 minutes. At the end of 60 minutes, the samples were checked for the presence of either glucose or mannose. From Figure 26 it is clear that there was only mannose detected when mannobiose was used as a substrate in a reaction of pure ManB with that substrate. When cellobiose, gentiobiose, salicin and arbutin were used, there was very minimal activity since a small amount of glucose was detected. All together, it was found that when mannobiose was used, the specific activity was about 133 µmol/min mg ManB, while with cellobiose the activity was 0.0007, with gentiobiose 0.0002, with salicin 0.00008 and arbutin 0.00004. Because the reaction of ManB with these substrates was only allowed to continue for 60 minutes, it shows that ManB was able to cleave

mannobiose at that time, although it is not clear whether other substrates were not hydrolyzed due to insufficient time given for the reaction or due to ManB not having any activity with these. In future studies, it would be of benefit to let the reaction proceed for longer period of time to determine whether ManB has slower activity with other substrates. Also, since mannobiose had to be used at much higher concentrations, it was hard to compare the amount of glucose with the amount of mannose obtained at the end of reaction, since there is no way to measure precisely the extent to which mannose is being detected in this assay. Previously, from using this kit with glucose and mannose side by side, it was found that only 30% of mannose is detected compared to 100% of glucose, which was used as a standard most of the time. Therefore, this discrepancy in detection levels of mannose and glucose cannot be measured accurately using this method alone, and these values should be taken as an indication of either presence or absence of mannose or glucose and not to compare the activity between these substrates. It can still be deduced from these results that ManB tends to have properties of a true mannosidase enzyme, and has no activity or perhaps much slower activity on substrates that are glucosides.

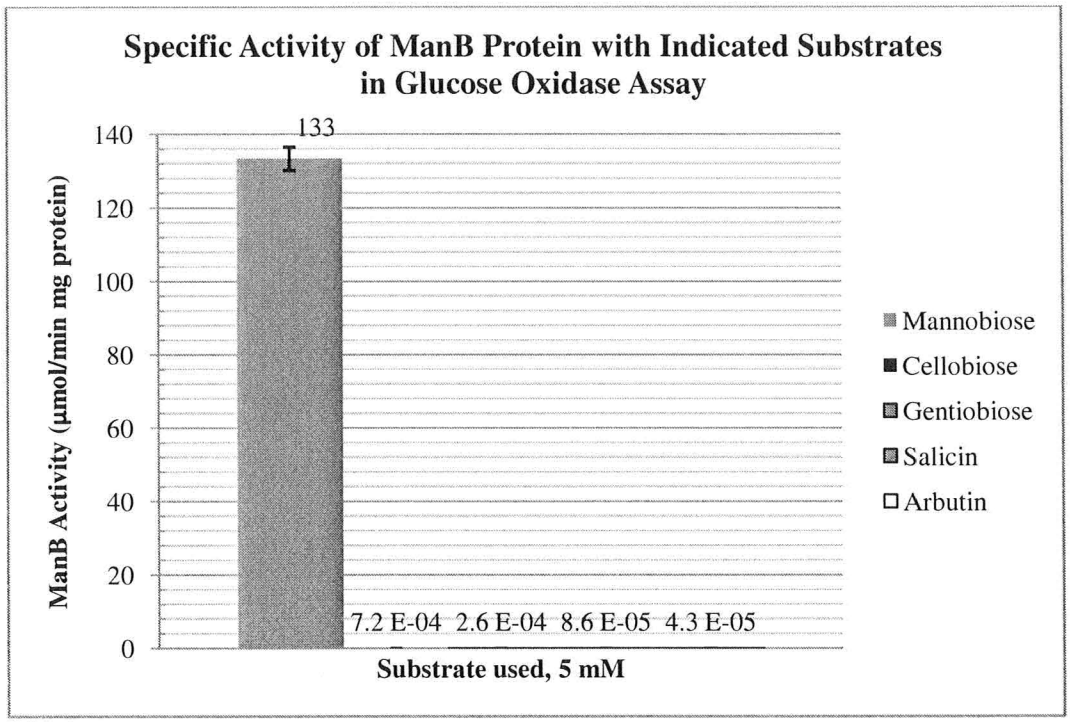


Figure 26: Glucose oxidase assay showing specific activity of ManB protein used in a reaction with each of these sugars (in the order listed on the graph): 5 mM mannobiose, 5 mM cellobiose, 5 mM gentiobiose, 5 mM salicin, 5 mM arbutin. Standard lines shown in figures Figure 24 and Figure 25 were used to quantify the amount of each product and then to determine further activity. The assay was allowed to proceed for 60 minutes.

These results also suggest that there is another enzyme present in *S. meliloti* genome that is capable of metabolizing these sugars. Whether ManB is the only enzyme that is capable of metabolism of mannosides is not clear, as there may be another enzyme present with a very similar activity in a genome. This question was addressed by construction a non-polar *manB* deletion mutant (Chapter 2). ManB is most likely a family

2 glycosyl hydrolase enzyme that is capable of cleaving mannobiose. It definitely has activity on mannosides and no activity on glucosides.

CHAPTER 5

Transposon Mutagenesis of *S.meliloti*. Screening for the Presence of Strains Unable to Utilize Cellobiose and Arbutin as Sole Carbon Sources

Transposon Mutagenesis

Transposon mutagenesis was carried out using transposon Tn5-B20 shown in Figure 27(Simon, Quandt, & Klipp, 1989). This Tn5 derivative has a neomycin cassette which can be used as a means for selection and a promoterless *lacZ* gene which can be used to screen for insertions downstream of an active promoter. The sequence of genomic DNA upstream from *lacZ* can also be determined using a primer going in the reverse direction of *lacZ* (P) (Figure 27). *S. meliloti* strain RmG212 (*lac*$^-$ strain) was used as a recipient for Tn5-B20 mutagenesis.

Based on the results from the protein assays, it is clear that sugars such as cellobiose, gentiobiose, salicin and arbutin are metabolized by enzymes other than ManB. It was interesting to determine which mutations would allow for inability of *S. meliloti* to use either one of these sugars, as these can suggest the possible genes involved in their metabolism.

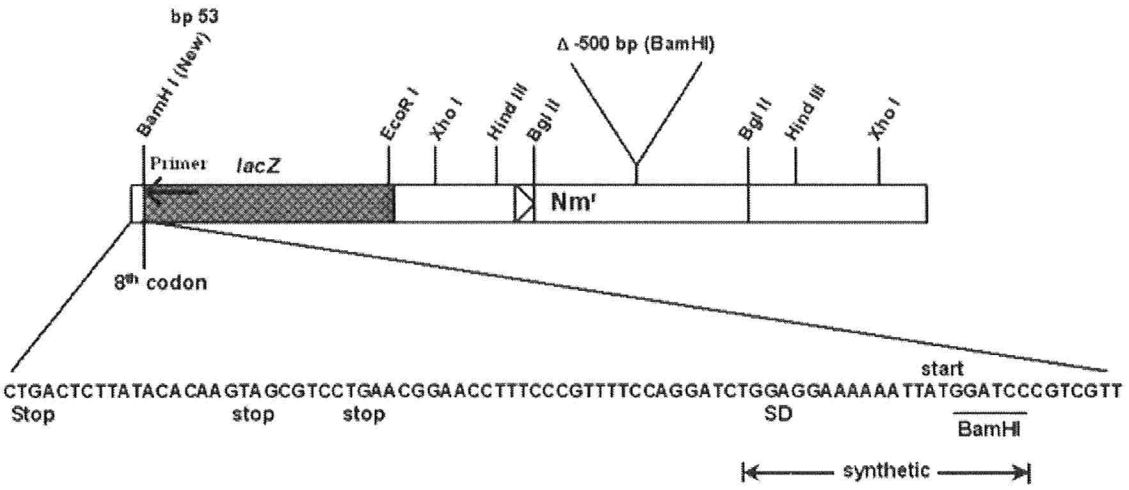


Figure 27: Map of Tn5-B20 transposon. This transposon was derived from Tn5 transposon. A primer (P) that is complementary to *lacZ* region can be used in a reverse direction in order to sequence a small fragment around the transposon. Using BLAST program from NCBI it is possible to search for the matches to the obtained sequence and find out a precise location of the transposon. (Figure obtained from Dr. R. Zaheer; and is a modified figure based on Simon, Quandt and Klipp,1989).

Tn5-B20 was introduced into *S.meliloti* strain RmG212 and colonies were selected on LB with Sm (200 µg/mL), Nm (200 µg/mL) and were screened for growth in M9 minimal media with either 5 mM cellobiose or 5 mM arbutin. It was expected that the strains that failed to grow on cellobiose would also fail to grow on gentiobiose and mannobiose as the lack of growth may result from the same mutation. Likewise, strains unable to use arbutin should also not be able to use salicin.

One thousand Tn5-B20 insertions were screened and two mutants, C1 and C2, were isolated that had a slow growth phenotype on cellobiose and could not grow on arbutin. Six more mutants were isolated that either grew with cellobiose but not arbutin,

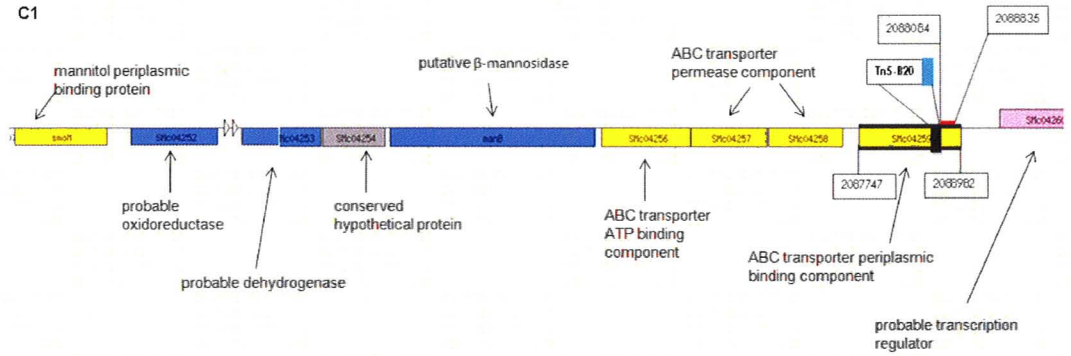
or grew slowly on cellobiose: A2, A5, A6, A8, A12, and A13. Sequencing genomic DNA using *lacZ* reverse primer (ML-1160: 5' CGC CAG GGT TTT CCC AGT CAC GAC GTT GTA 3') from seven Tn5-B20 insertion mutants were determined. Searching *S.meliloti* genome with these sequences identified the insertion sites as outlined in Table 7. Since mutants A2 and A6 were determined to grow very slowly in liquid minimal media with glucose and succinate as sole carbon sources these were not investigated further.

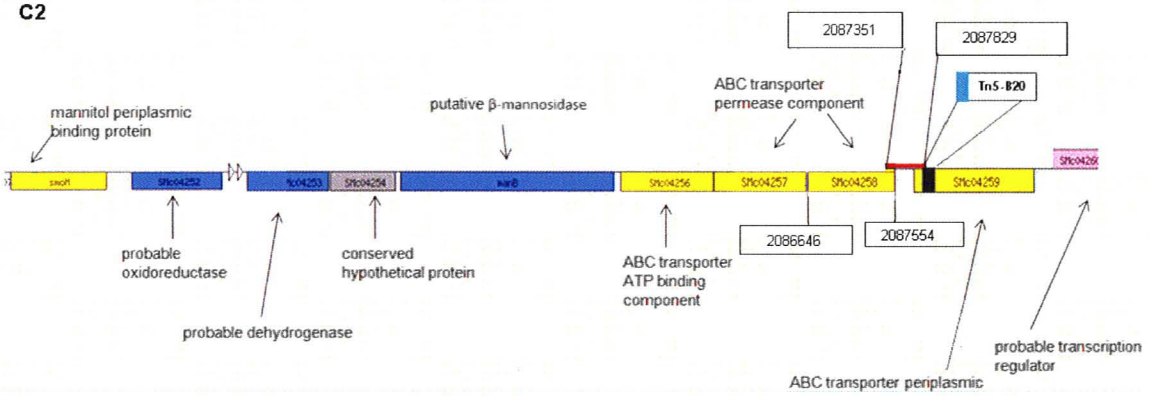
Table 7: Sequences from Tn5-B20 mutants deficient in growth on 5 mM cellobiose and 5 mM arbutin. Abbreviations: nts=nucleotides, N/A= not applicable

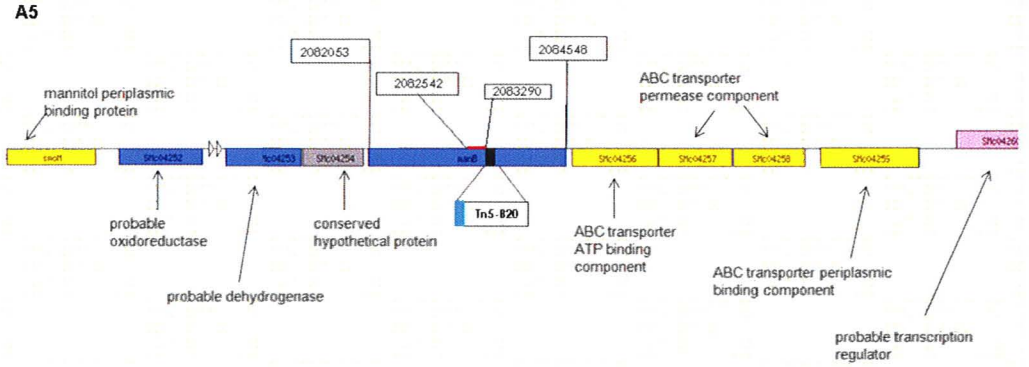
Mutant number	Library Reference	Sequenced region location in the genome	Length of sequence obtained	Location of insertion
C1	RmP2260	2088084-2088835	751 nts	*smc04259*
C2	RmP2261	2087829-2087351	478 nts	*smc04259 – smc04258*
A2	N/A			
A5	RmP2262	2083290-2082542	748 nts	*smc04255/manB*
A6	N/A	2958342-2959140	798 nts	*smc03964-smc03965*
A8	RmP2263	3121280-3120601	679 nts	*smc03162*
A13	RmP2264	3121042-3120645	379 nts	*smc03162*

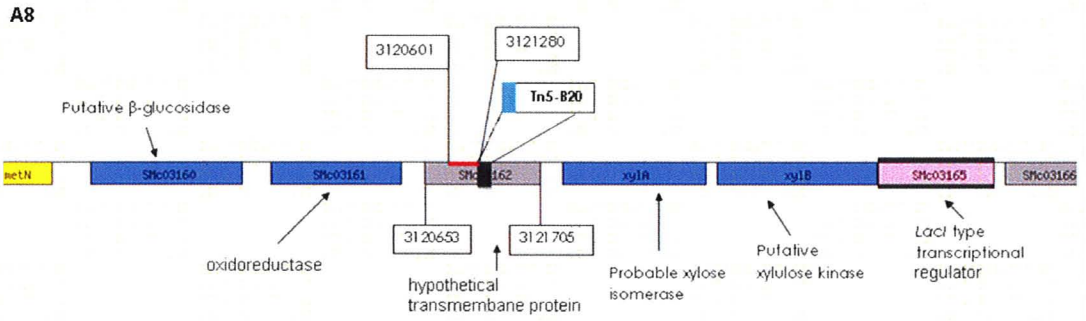
The locations of the Tn5-B20 insertions relative to genes near the insertion sites are shown in Figure 28. The two mutants that grew slowly with cellobiose (C1 and C2) both have a Tn5-B20 insertion within the putative ABC transporter system. Furthermore, strain A5 that showed no growth with arbutin as a sole carbon source was determined to

have a Tn5-B20 insertion within *manB* gene. This means that either ManB or any of the genes following it would be responsible for metabolism of aryl-glycosides.


C1
mannitol periplasmic binding protein
putative β-mannosidase
ABC transporter permease component
2088084
2088835
Tn5-B20
probable oxidoreductase
conserved hypothetical protein
ABC transporter ATP binding component
2087747
2088982
ABC transporter periplasmic binding component
probable dehydrogenase
probable transcription regulator


C2
2087351
2087829
Tn5-B20
mannitol periplasmic binding protein
putative β-mannosidase
ABC transporter permease component
probable oxidoreductase
conserved hypothetical protein
2086646
2087554
ABC transporter ATP binding component
probable dehydrogenase
probable transcription regulator
ABC transporter periplasmic


A5
2082053
2084548
2082542
2083290
mannitol periplasmic binding protein
ABC transporter permease component
Tn5-B20
probable oxidoreductase
conserved hypothetical protein
ABC transporter ATP binding component
ABC transporter periplasmic binding component
probable dehydrogenase
probable transcription regulator


A8
3120601
3121280
Tn5-B20
Putative β-glucosidase
SMc03160
SMc03161
xylA
xylB
SMc03165
SMc03166
oxidoreductase
3120653
3121705
hypothetical transmembane protein
Probable xylose isomerase
Putative xylulose kinase
LacI type transcriptional regulator

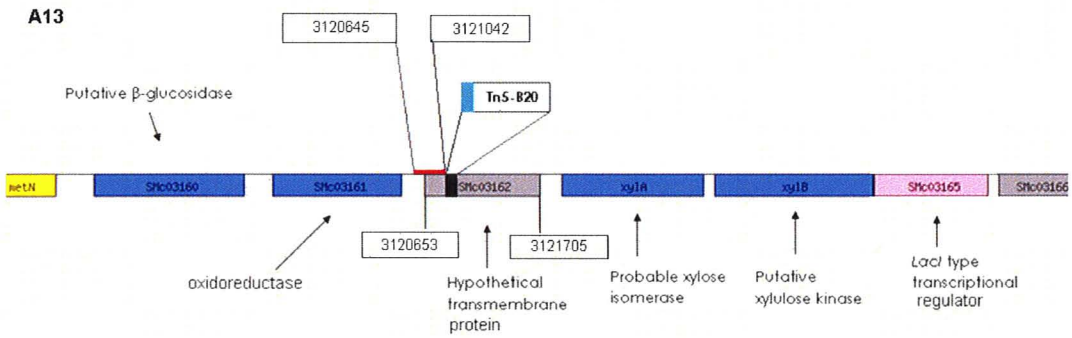


Figure 28: Tn5-B20 insertion sites in mutants C1, C2, A5, A8, and A13. Map of each mutant showing the precise location of Tn5-B20 transposon. The numbers in the box represent the location on the chromosome. The blue area on Tn5-B20 represents *lacZ* gene. The red line shows the location of genomic DNA that was sequenced using *lacZ* reverse primer. Mutants C1 and C2 show a slow growth phenotype when cellobiose is used as a source of carbon. Both of these have an insertion of Tn5-B20 within *smc04259* but at different locations, the putative periplasmic binding protein. Mutant A5 has Tn5-B20 integrated within *manB* gene and shows no growth when arbutin is used as a sole source of carbon. Mutants A8 and A13 both have Tn5-B20 inserted within *smc03162* and show no growth when arbutin is used as the sole source of carbon but do grow fairly well with cellobiose.

Both mutants A8 and A13 carried a Tn5-B20 insertion within the *smc03162* gene encoding a hypothetical transmembrane protein. These have shown that the transposon was inserted within a completely different gene cluster. Failure of A8 and A13 to grow on arbutin suggests that perhaps more than one cluster of genes is involved in metabolism of glycosidic sugars. The annotations of these genes suggest, that downstream of *smc03162* is a probable oxidoreductase protein (*smc03161*) and a gene *smc03160* encoding a β-glucosidase protein. According to annotation, there are no genes in that cluster that may be involved in transport of sugars.

The transposon insertions result in polar mutations, and thus if the transposon lies in a putative ABC transporter system, the rest of the downstream genes in that operon will also be rendered inactive, as they will not be transcribed. Therefore, the insertion within the *smc04259* (strains C1 and C2, Figure 28) will also result in complete inactivation of this cluster, and should have a phenotype similar to strain RmP2202 (deletion of *smc04259-smc04252* inclusive). It is important to note, that even with this mutation, the strains were still able to grow on cellobiose as a sole source of carbon. This suggests that the gene cluster in Figure 40 (Appendix) is not essential for metabolism of cellobiose.

In the case of arbutin, the strains showed complete lack of growth when they were selected for, when that sugar was used. Arbutin is an aryl-glycoside, and therefore it is probable that it will require a different mode of metabolism as hydrolysis of its β-(1-4) glycosidic bond releases a glucose and a hydroquinone. The inability to use arbutin when some of the genes have been lost from the gene cluster in Figure 40 suggests that it is essential for the metabolism of aryl-glycosides. Mutant A5, that showed a transposon insertion within *manB* gene, also shows no growth. It may suggest that either ManB is involved in metabolism of that sugar, or any other gene that is downstream from *manB*: *smc04254, smc04253, smc04252*. From the growth curve experiment with the library strains (Chapter 2), it was shown that the strains that had an insertion in either *smc04253* or *smc04252* (SmFL4430 or SmFL4262) were able to grow just as well as the wild type RmP110 on arbutin. This would suggest that it is possible that either *manB* or *smc04254* are responsible for the ability of *S. meliloti* to use aryl-glycosides such as arbutin or

salicin. Because the ManB activity assays showed that ManB is strictly a β-mannosidase, it suggests *smc04254* is required for complete metabolism of aryl-glycosides.

Growth Curves of Strains Obtained through Transposon Mutagenesis

The strains shown in Figure 28 were selected on the bases of their poor growth on cellobiose or arbutin but normal growth on glucose and succinate. To investigate the growth phenotypes more fully, growth curves were carried out in M9-minimal media containing the sole carbon sources cellobiose, gentiobiose, mannobiose, arbutin and salicin at 5 mM, and glucose at 10 mM.

Once again, to address the issue of different growth rates with different carbon sources in M9 minimal media, the doubling times were compared (Table 8).

Table 8: Doubling time of wild type *S.meliloti* strain RmP110 in M9 minimal media with each sugar as a sole source of carbon. The data was obtained from Figure 29.The mean growth rate constant k was calculated using the equation $k=(\log N_t-\log N_o)/0.301$, where N is the number of cells measured at OD600 at the start and at the end of measurement. Mean generation time g was calculated per hour using equation $g= 1/k$. g is the doubling time, which is the time it takes for one generation to double.

Carbon Source	Doubling time (g)
10 mM Glucose	4.0
5 mM Cellobiose	4.2
5 mM Gentiobiose	4.6
5 mM Salicin	11.3
5 mM Arbutin	15.5

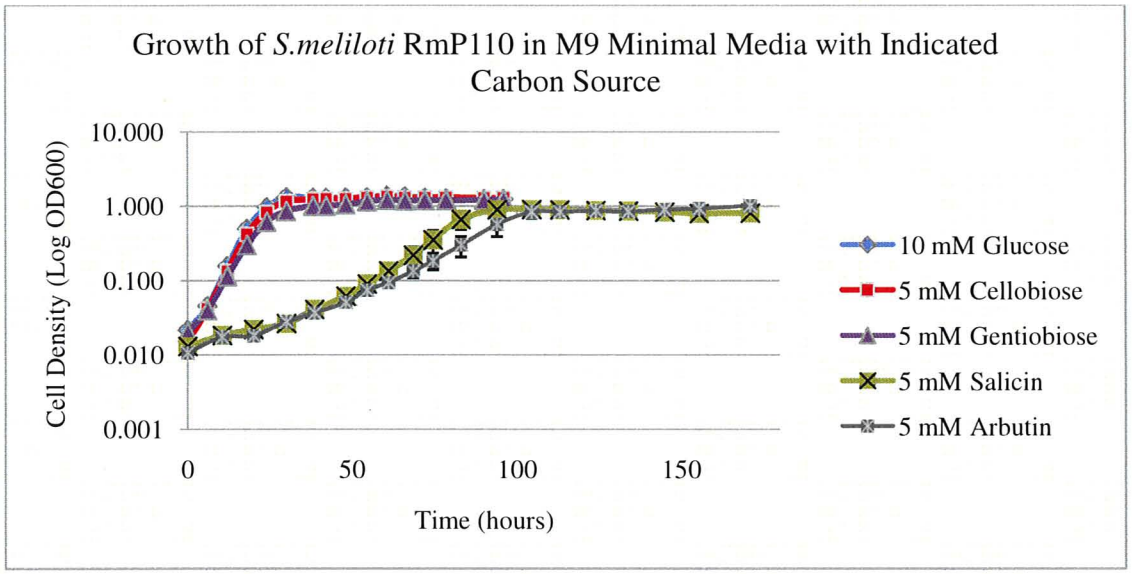


Figure 29: Growth rate of wild type *S. meliloti* strain RmP110 in M9 minimal media with an indicated sugar as a sole source of carbon. Growth with arbutin and salicin is measured over a longer time period as it proceeds very slowly compared to other sugars.

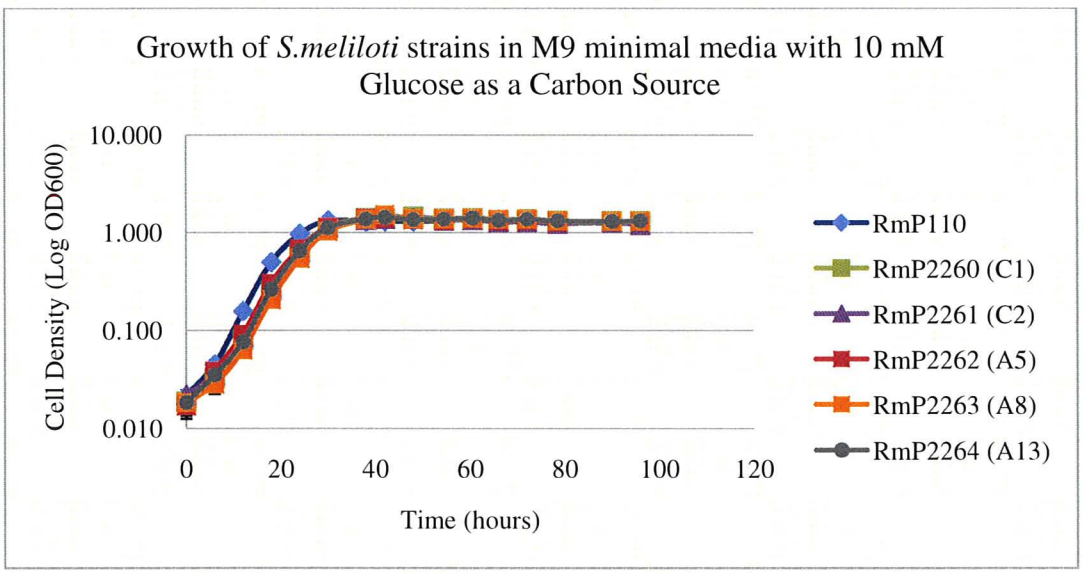


Figure 30: Growth curve of *S.meliloti* Tn5-B20 mutants with 10 mM glucose as a sole source of carbon.

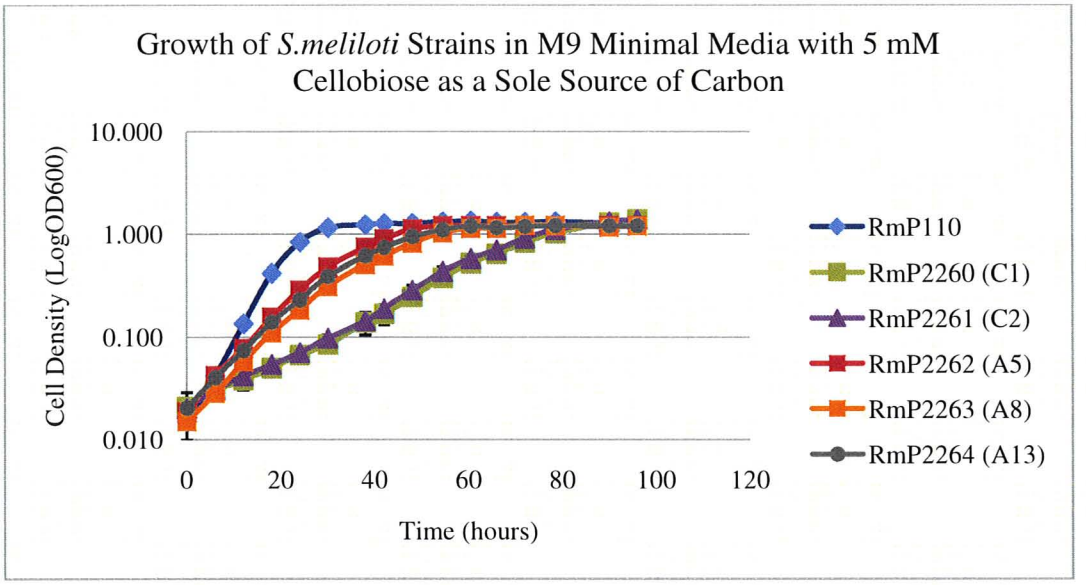


Figure 31: Growth curve of *S.meliloti* Tn5-B20 mutants with 5 mM of cellobiose as a sole source of carbon.

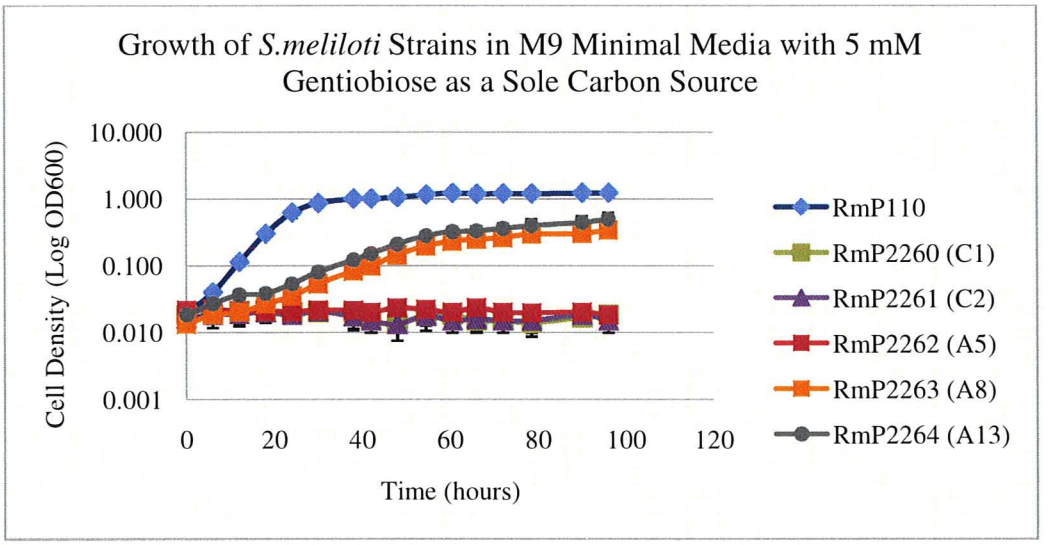


Figure 32: Growth of *S.meliloti* Tn5-B20 mutants with 5 mM of gentiobiose as a sole source of carbon.

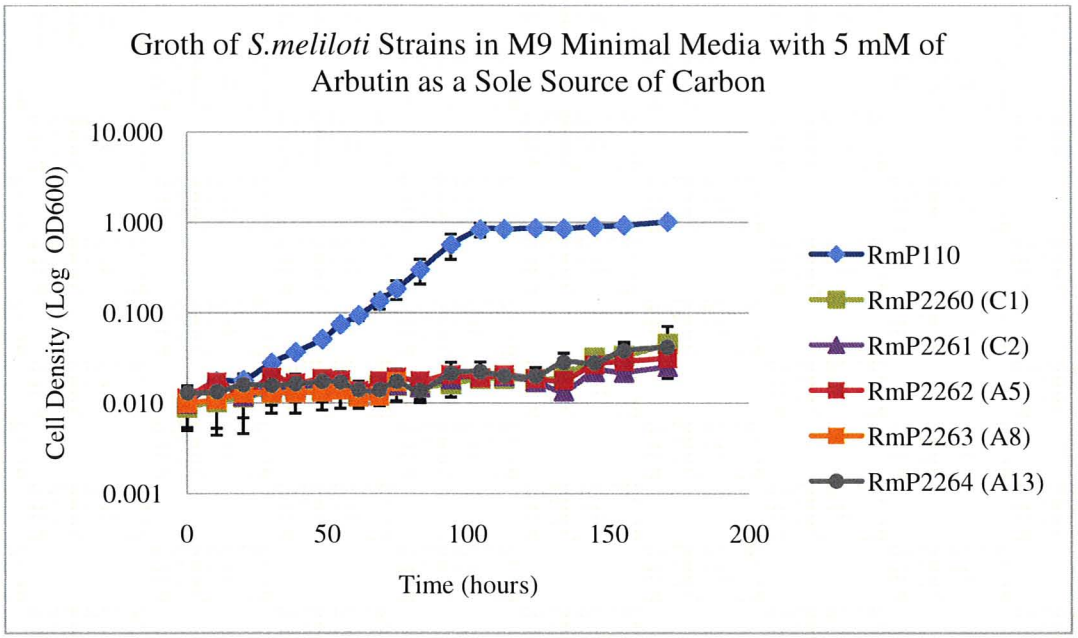


Figure 33: Growth of *S.meliloti* Tn5-B20 mutants with 5 mM of arbutin as a sole source of carbon.

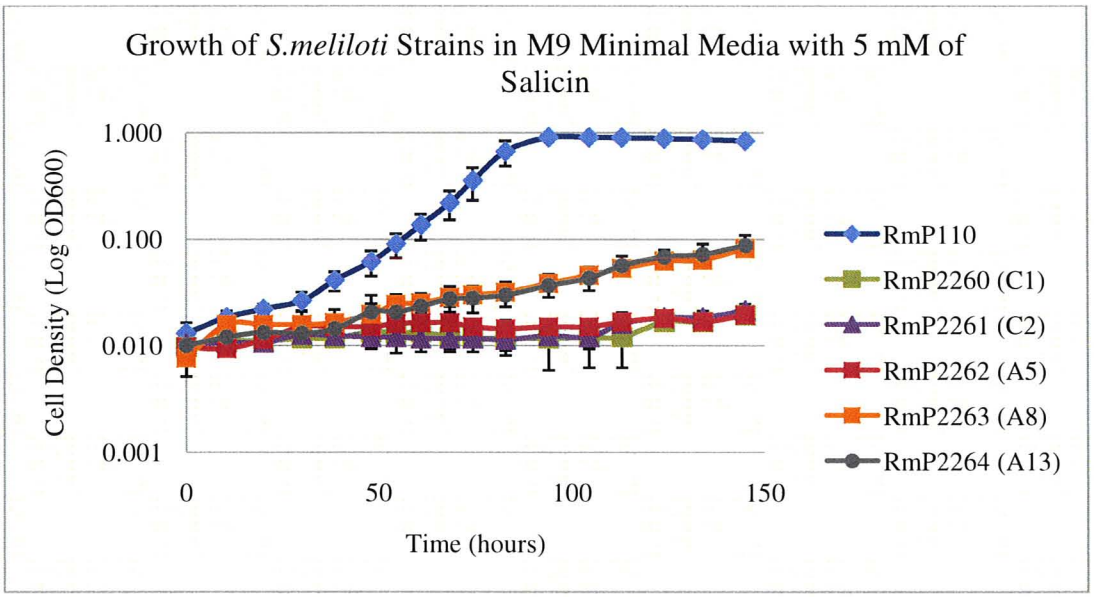
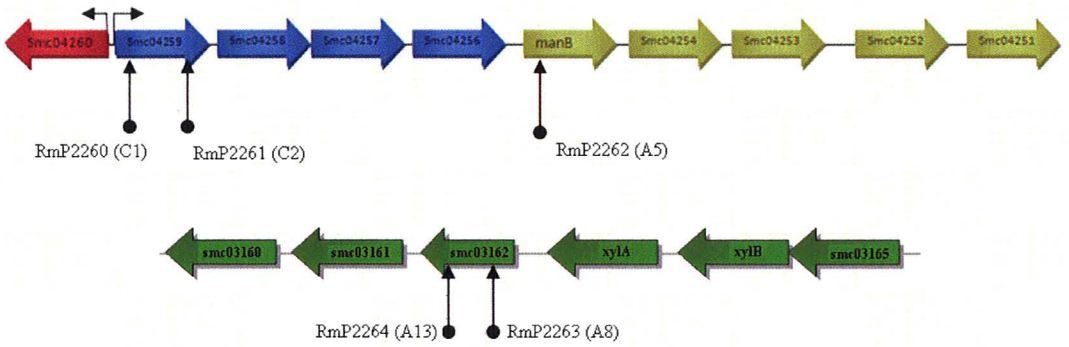


Figure 34: Growth of *S.meliloti* Tn5-B20 mutants with 5 mM of salicin as a sole source of carbon.

Table 9: Summary of growth of the Tn5-B20 mutant strains with the selected sources of carbon. Data was obtained from Figure 30, Figure 31, Figure 32, Figure 33 and Figure 34. +++ = growth comparable to wild type strain RmP110, ++ = growth slower than wild type strain RmP110, += growth is minimal compared to wild type strain RmP110, - = no growth. The figure shows the *smc04260-smc04251* gene cluster as well as another gene cluster, *smc03160 – smc03165,* where the transposon has inserted. The arrows point out the precise location of the transposon, while the strain that has that insertion is indicated next it each arrow.



Carbon Source	RmP110	RmP2260 (C1)	RmP2261 (C2)	RmP2262 (A5)	RmP2263 (A8)	RmP2264 (A13)
Glucose	+++	+++	+++	+++	+++	+++
Cellobiose	+++	+	+	++	++	++
Gentiobiose	+++	-	-	-	++	++
Salicin	+++	-	-	-	++	++
Arbutin	+++	-	-	-	-	-

As shown in Figure 31, when cellobiose is used as a source of carbon all strains were eventually able to grow and reach similar cell densities, although some mutants grew at a much slower rate. Mutants A5 (*manB*::Tn5-B20), A8 (*smc03162*::Tn5-B20) and A13

(*smc03162*::Tn5-B20) grew similarly and reached the stationary stage after 54 hours, compared to 30 hours for the wild type RmP110. On the other hand, C1 (*smc04259*::Tn5-B20) and C2 (*smc04259* ::Tn5-B20) carrying similar transposon insertions reached the stationary phase after 80 hours of growth. It is interesting that even despite very slow growth, as compared to RmP110, these strains still do grow when cellobiose is used. This suggests that cellobiose is metabolized independent of the two gene clusters, *smc04260-smc04251* and *smc03160-smc03165*, even though both clusters are clearly involved. *S.meliloti* then may have more than one pathway for breakdown of this sugar. That is not surprising, considering how widespread cellobiose is, and that it is most likely available in abundance in the soil – plant environment.

When the strains were grown with gentiobiose, as shown in Figure 32, it was evident that the gene cluster *smc04259- smc04252* is essential for metabolism of this sugar as mutant A5 (*manB*::Tn5-B20) failed to grow on this sugar while mutants C1 (*smc04259* ::Tn5-B20) and C2 (*smc04259*::Tn5-B20) grew very slowly. It is also clear that mutation within genes *smc03162* to *smc03160* is also important, since the insertion of transposon within *smc03162* results in a slower growth when gentiobiose is used as a source of carbon. The stationary stage was not reached by any of the mutant strains, although the A8 (*smc03162*::Tn5-B20) and A13 (*smc03162*::Tn5-B20) mutants grew to an OD600 of 0.4-0.5. This suggests that the *smc04260-smc04251* gene cluster is essential for metabolism of gentiobiose. The genes *smc03162* to *smc03160* may also be involved in hydrolysis or gentiobiose, although they may not be essential.

Tn5-B20 insertion mutants at both gene clusters failed to grow on arbutin. Insertion in the *smc04260-smc04251* gene cluster also grew poorly on salicin whereas insertion mutants in the *smc03162* gene did not grow on salicin (Figure 34).

Complementation of transposon mutagenesis strains using pLAFR1 cosmid library

A random library contained on a pLAFR1 cosmid vector in *E.coli* strains HB101 contains *S.meliloti* gene fragments, 23 kilobases in size on average, inserted at the *Eco*RI site of the pLAFR1 cosmid vector (Friendman, Long, Brown, Buikema, & Ausubel, 1982). pLAFR1 cosmid is tetracycline resistant, and about 20.3 kilobases in size (Vanbleu, Marchal, & Vanderleyden, 2004), (Freidman, Long, Brown, Buikema, & Ausubel, 1982). It replicates in both *E.coli* and *S.meliloti* and two primers, P129 and P128 (Table 2) have been designed to extend into the site of insertion of the *S.meliloti* DNA fragment cloned into pLAFR1. Hence the insertion boundaries can be determined by DNA sequencing.

None of the Tn5-B20 mutants were able to grow well with salicin alone as a sole source of carbon (Figure 34, Table 9). The pLAFR1 library was transferred into the three mutants C1 (*smc04259*::Tn5-B20), A5 (*manB*:: Tn5-B20) and A8 (*smc03162*::Tn5-B20), which are entered into the library as strains RmP2260, RmP2262 and RmP2263 respectively. Transconjugants that grew on M9-salicin were selected.

Seven transconjugants from each mating were purified and the cosmid was transferred from these *S.meliloti* strains into *E.coli* DH5α cells. Tc^r, Cm^s transconjugants

were purified to select for the presence of pLAFR1 cosmid and the absence of pRK600. Cosmid DNA was isolated from each complemented strain. The DNA sequencing reaction was run with each using primers P129 and P128.

Of eight cosmid clones obtained from RmP2260 (C1, *smc04259*::Tn5-B20) transconjugants, all contained matching sequences to the borders around the same region from the *S.meliloti* genome, i.e. *smc04260-smc04251*, Figure 35. This region includes the genes that were inactivated by the Tn5-B20 insertion in *smc04259*. No other cosmids that complemented this mutation were isolated.

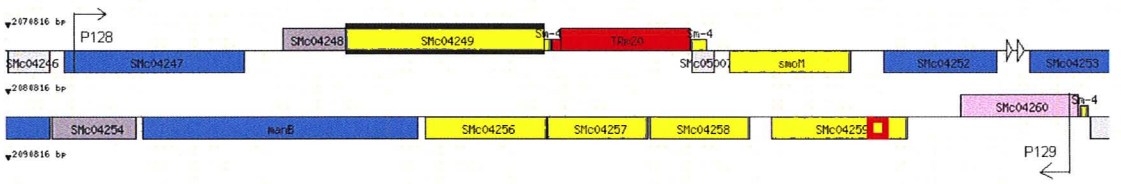


Figure 35: Fragment of *S.meliloti* genome showing the region that complemented mutant C1, resulting in its ability to grow with salicin as a sole source of carbon. P128 and P129 are showing the direction in which the primers amplified the borders of this fragment. The red box indicates the initial location of the mutation in *S. meliloti* strain RmP2260 (C1). This mutation was not present in the cosmid library regions, but is only shown to indicate that the same genes have complemented this mutation.

A representative *S. meliloti* transconjugant was designated as RmP2265 and the cosmid was assigned pTH2615.

The RmP2262 (mutant A5, *manB*::Tn5-B20) pLAFR1 transconjugant plasmid had exactly the same borders as pTH2615. This is not a surprising result, given that the

transposon insertion in RmP2262 was within the *manB* gene, which is in the same operon as *smc04259*. The complemented strain was entered into the library as RmP2266.

Two different but overlapping cosmids were found to complement the A8 mutants, *smc03162*::Tn5-B20 insertion in RmP2263. Several cosmids were found to complement this mutation, one insert region spanned from *smc03173* to *smc03156* (Figure 36), while the other spanned from *smc02977* to *smc03163* (Figure 37). The two regions included genes from *smc03132* to *smc03160* which overlaps the A8 (*smc03162*::Tn5-B20) insertion site.

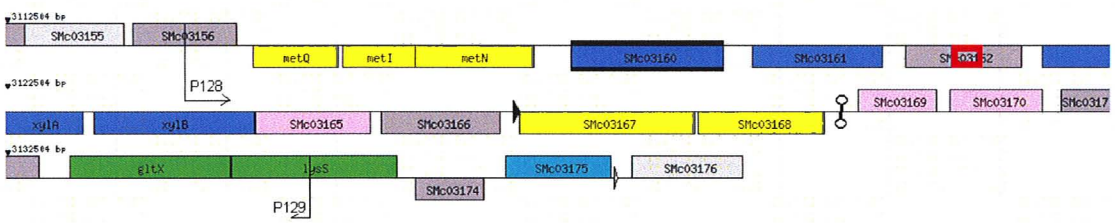


Figure 36: The region from a pLAFR1 cosmid pTH2616 that complemented the mutation A8 (*smc03162*::Tn5-B20). The location of this mutation is indicated in a red box within gene *smc03162*. The borders around this region that have been sequenced are indicated as P128 and P129.

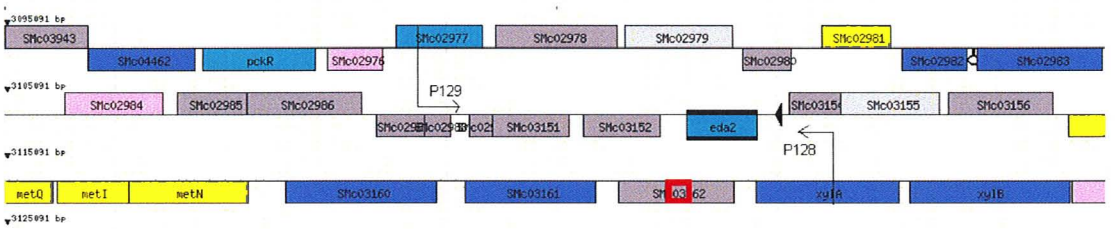


Figure 37: The region from a pLAFR1 cosmid pTH2617 that complemented the mutation A8 (*smc03162*::Tn5-B20). The location of this mutation is indicated in a red box within gene *smc03162*. The borders around this region that have been sequenced are indicated as P128 and P129.

These isolated cosmids were designated as pTH2616 and pTH2617 and the complemented *S.meliloti* strains as RmP2267 and RmP2268.

This data confirms that gene clusters *smc04259 – smc04252* and *smc03132 – smc03160* are necessary for growth with sugars such as cellobiose, gentiobiose, mannobiose, arbutin and salicin. No other gene clusters were found to complement the mutations in these gene clusters.

CHAPTER 6

Conclusions

The *smc04260-smc04251* gene cluster includes an ABC transport system, which seems to be specific for transport of sugars into the *S.meliloti* cell. Based on the results obtained here, it seems to be specific for aryl-glycosidic sugars such as arbutin and salicin and sugars such as cellobiose, gentiobiose and mannobiose. The *smc04260-smc04251* gene cluster also encodes for proteins that are involved in metabolism of these sugars. Growth curves with library fusion strains and with in-frame deletion strains showed slower growth of these strains on cellobiose, mannobiose and gentiobiose as well as lack of growth for some strains on gentiobiose, mannobiose, salicin and arbutin. This suggests that there is a complex metabolic pathway for the breakdown of these sugars present in *S.meliloti*. Growth with aryl-glycosides arbutin and salicin seems to be especially affected. The gene product of *smc04254* plays an important role in further metabolism of aryl-glycosides and gentiobiose. The absence of this gene results in slow growth of bacteria when cellobiose is used and no growth at all with gentiobiose and aryl-glycosides. Growth with mannobiose in the absence of *smc04254* was not affected, suggesting that *smc04254* is not involved in β-mannoside metabolism.

ManB has been shown to have activity only when mannobiose was used as a substrate in an assay with glucose oxidase. Other sugars, when used as a substrate, showed minimal activity when the reaction was allowed to proceed for 60 minutes. When p-nitrophenyl-mannopyranoside and p-nitrophenyl –glucopyranoside were used in an assay,

ManB was found to only be active on p-nitrophenyl-mannopyranoside. This confirms that ManB is a true β-mannosidase enzyme, probably belonging to family 2 of glycoside hydrolases.

It was demonstrated that *S.meliloti* has more than one system for transport and metabolism of these sugars. Transposon mutagenesis study has shown that there is at least another gene cluster that is involved in metabolism of these sugars – *smc03160-smc03161.*

The ability of *S.meliloti* to break down these sugars should be further explored. Specifically, gene *smc04254* seems to be very important in breakdown of the sugars studied here, yet it lacks any clear annotation. It would be useful to establish its precise role and activity in the future. Cluster *smc03160-smc03165* has been shown to be involved in metabolism of these sugars. It would also be interesting to look at the function of the genes in that cluster with respect to sugar metabolism.

Metabolism of sugars is important for understanding metabolic capacities of soil organisms such as *S.meliloti.* It seems that there are more genes devoted to the metabolism and transport than was originally thought. In addition, the enzymes that are able to break down larger sugars into simple monomers are in high demand, as this can have many useful applications in a biotechnology field.

Bibliography

Beki, Nagy, Vandeleyden, Jager, Kiss, Fulop, et al. (2003). Cloning and Heterologous Expression of a B-D-Mannosidase (E.C. 3.2.1.25) - Ecnoding Gene from *Thermobifida fusca* TM51. *Applied and Environmental Microbiology , 69* (4), 1944-1952.

Bergmann, Rimon, & Sigal. (1958). Effect of pH on the Activity of Eel Esterase Towards Different Substrates. *Biochemcal Journal , 68* (3), 493-499.

Bergmeyer, & Bernt. (1974). Determination with Glucose Oxidase and Peroxidase. In H. U. Bergmeyer, *Methods of Enzymatic Analysis* (pp. 1205-1214). New York: New York Academic Press, 2nd ed.

Biggs, A. (1954). A Spectrophotometric Determination of the Dissociation Constants of p-Nitrophenol and Papaverine. *Transactions of the Faraday Society , 50*, 800-802.

Bowers, McComb, Christensen, & Schaffer. (1980). High Purity 4-Nitrophenol: Purification, Characterization and Specifications for Use as a Spectrophotometric Reference Material. *Clinical Chemistry , 26* (6), 725-729.

Boyer, & Roulland-Dssouix. (1969). A Complementation Analysis of the Restriction and Modification of DNA in Escherichia coli. *Journal of Molecular Biology , 41*, 459-472.

Bradford, M. M. (1976). A Rapid and Sensitive Method for the Quantitation of the Microgram Quantities of Protein Utalizing the Principle of Protein-dye Binding. *Analytical Biochemistry* , 248-254.

Cantarel, Coutinho, Rancurel, Bernard, Lombard, & Henrissat. (1998-2010). *Glycoside Hydrolase Family 2*. Retrieved 2008-2010, from CAZy - Carbohydrate Active Enzymes: www.cazy.org

Cantarel, Coutinho, Rancurel, Bernard, Lombard, & Henrissat. (2009). The Carbohydrate Active Enzymes Database (CAZy) - an Expert Resource for Glycogenomics. *Nucleic Acids Research , 37*, D233-D238.

Cheng, Sibley, Zaheer, & Finan. (2007). A *Sinorhizobium meliloti minE* Mutant Has an Altered Morphology and Exibits Defects in Legume Symbiosis. *Microbiology , 153*, 375-387.

Cowie, A., Cheng, J., Sibley, Fong, Zaheer, R., Patten, C., et al. (2006). An Integrated Approach to Functional Genomics: Contstruction of a Novel Reporter Gene Fusion Library for *Sinorhizobium meliloti. Applied and Environmental Microbiology , 72* (11), 7156-7167.

Cowie, Cheng, Sibley, Fong, Zaheer, Patten, et al. (2006). An Integrated Approach to Functional Genomics: Contstruction of a Novel Reporter Gene Fusion Library for *Sinorhizobium meliloti. Applied and Environmental Microbiology , 72* (11), 7156-7167.

Stalbrand, & Warren. (1999). Mannan-Degrading Enzyme from *Cellulomonas fimi. Applied and Environmental Microbiology , 65* (6), 2598-2605.

Datsenko, & Wanner. (2000). One Step Inactivation of Chromosomal Genes in *Escherichia coli* K-12 Using PCR Products. *PNAS , 97* (12), 6640-6645.

Davidson, Dassa, Orelle, & Chen. (2008). Structure, Function, and Evolution of Bacterial ATP-binding Casette Systems. *Microbiology and Molecular Biology Reviews , 72* (2), 317-364.

Driessen, Rosen, & Konings. (2000). Diversity of Transport Mechanisms: Commong Structural Principles. *Trends in Biochemical Sciences , 25*, 297-401.

Duffaud, McCutchen, Leduc, Parker, & Kelly. (1997). Purification and Characterization of Extremly Thermostable B-Mannanase, B-Mannosidase, and A-Galactosidase From the Hyperthermophilic Eubacterium *Thermatoga neapolitana* 5068. *Applied and Environmental Microbiology , 63* (1), 169-177.

Ebringerova, Hromadkova, & Heinze. (2005). Hemicellulose. *Advance Polymer Science* , 1-67.

Falip, Delalande, Carapito, Goubet, Hatsch, Leize-Wagner, et al. (2005). Diversity of the Exoproteome of *Fusarium graminearum* Grown in Plant Cell Wall. *Current Genetics* , 48:366-379.

Finan, Hartweig, LeMieux, Bergman, Walker, & Signer. (1984). General Transduction in *Rhizobium meliloti*. *Journal of Bacteriology , 159* (1), 120-124.

Finan, Kunkel, DeVos, & Signer. (1986). Second Symbiotic Megaplasmid in *Rhizobium meliloti* Carrying Exopolysaccharide and Thiamine Synthesis Gene. *Journal of Bacteriology , 167* (1), 66-72.

Fraenkel, & Vinopal. (1973). Carbohydrate Metabolism in Bacteria. *Annual Review of Microbiology , 27*, 69-100.

Freidman, Long, Brown, Buikema, & Ausubel. (1982). Construction of a Broad Host Range Cosmid Cloning Vector and its Use in the Genetic Analysis of *Rhizobium* Mutants. *Gene , 18*, 289-296.

Galibert, Finan, Long, Puhler, Abola, Ampe, et al. (2001). The Composite Genome of a Legume Symbiont *Sinorhyzobium meliloti*. *Science , 293*, 668-672.

Gasteiger, Gattiker, Hoogland, Ivanyl, Appel, & Bairoch. (2003). *ExPASy Proteomics Server*. Retrieved Nomeber 2008, 2010, from EXPASy: The Proteomics Server for In-depth Protein Knowledge and Analysis: http://ca.expasy.org/tools/pi_tool.html

Henrissat, & Bairoch. (1993). New Families in the Classification of Glycosil Hydrolases Basen on Amino Acid Sequence Similarities. *Biochemistry Journal , 293*, 781-788.

Henrissat, B. (1991). A Classification of Glycosil Hydrolases Basen on Amino Acid Sequence. *Biochemistry Journal* , 309-316.

Hirsch, & Beringer. (1984). A Physical Map of pPH1JI and pJB4JI. *Plasmid , 12*, 139-141.

Iacazio, Perissol, & Faure. (2000). A New Tannanase Substrate for Spectrophotometric Assay. *Journal of Microbiological Methods , 42*, 209-214.

Jez, Bennett, Schlegel, Lewiss, & Penninng. (1997). Comparative Anatomy of the Aldo/Keto Reductase Superfamily. *The Biochemical Journal* , 625-636.

Kingston, Scopes, & Baker. (1996). The Structure of Glucose-Fructose Oxidoreductase from *Zymomonas mobilis*: an Osmoprotective Periplasmic Enzyme Containing Non-dissociable NADP. *Structure , 4*, 1413-1428.

Lehninger, A., Nelson, D., & Cox, M. (2005). *Lehninger Principles of Biochemistry* (4th ed., Vol. 1). New York: W.H. Freeman.

Liepman, Nairn, Willats, Sorensen, Roberts, & Keegstra. (2007). Functional Genomic Analysis Supports Conservation of Function Among Cellulose Synthase-Like A Gene Family Members and Suggests Diverse Roles of Mannans. *Plant Physiology , 143*, 1881-1893.

Mauchline, Fowler, East, Sartor, Zaheer, Hosie, et al. (2006). Mapping the *Sinorhizobium meliloti* 1021 Solute-binding Protein Dependent Transportome. *Proceeding of the National Academy of Sciences , 103* (47), 17933-17938.

Minic, & Jouanin. (2006). Plant Glycoside Hydrolases Involved in Cell Wall Polysaccharide Degradation. *Plant Phisiology and Biochemistry , 44*, 435-449.

Nucleotide Basic Local Alignment Sequence Tool. (2009-2010). Retrieved 2009, from National Center for Biotechnology Information: http://blast.ncbi.nlm.nih.gov/Blast.cgi

Palmer, & Anderson. (1972). Metabolism of Gentiobiose in *Aerobacter aerogenes. Journal of Bacteriology , 112* (3), 1316-1320.

Parker, Chhabra, Lam, Callen, Duffaud, Snead, et al. (2001). Galactomannanases Man2 and Man5 From *Thermatoga* Species: Growth Physiologhy on Galactomannans, Gene Sequence Analysis and Biochemical Properties of Recombinant Enzymes. *Biotechnology and Bioengineering , 75*, 322-333.

Rabus, Jack, Kelly, & Saier. (1999). TRAP Transporters: an Ancient Family of Extracytoplasmic solute-receptor-dependant secondary active transporters. *Microbiology* , 3431-3445.

Sambrook Joseph, Russel David W. (2001). Protocol 25. In R. D. Sambrook Joseph, *Molecular Cloning: A Laboratory Manual* (pp. 1.116-1.118). Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Sartor, A. (2006). *Expression Analysis of the Transporters of Sinorhizobium meliloti.* Hamilton: McMaster University.

Schmidt, & Skerra, A. (2007). The Strep Tag System for One Step Purification and High Affinity Detection or Capturing of Proteins. *Nature Protocols , 2* (6), 1528-1535.

Schweiser, Klassen, & Hoang. (1996). Improved Methods for Gene Analysis and Expression in *Pseudomonas*. In H. Schweiser, T. Klassen, & T. Hoang, *Molecular Biology of Pseudomonas* (pp. 229-237). Washington DC: ASM Press.

Simon, Quandt, & Klipp. (1989). New Derivatives of Transposon Tn5 Suitable for Mobilization of Replicons, Gneration of Operon Fusions and Introduction of Genes in Gram Negative Bacteria. *Gene , 80* (1), 161-169.

Somers, Vanderleyden, & Srinivasan. (2004). Rhizosphere Bacterial Signalling: A Love Parade Beneath Our Feet. *Critical Reviews in Microbiology* , 30:205-230.

Stowers, M. D. (1985). Carbon Metabolism in Rhizobium Species. *Annual Review in Microbiology , 39*, 89-108.

Tailford, Money, Smith, Dumon, Davies, & Gilbert. (2007). Mannose Foraging by *Bacteriodes thetaiotaomicron*, Structure and Specificity of the B-Mannosidase, *BtMan2A*. *The Journal of Biological Chemistry , 282* (15), 11291-11299.

Tam, & Saier. (1993). Structural Functional and Evolutionary Relationships Among Extracellular Solute - Binding Receptors in Bacteria. *Microbiological Reviews , 57* (2), 320-346.

Vanbleu, Marchal, & Vanderleyden. (2004). Genetic and Physical Map of the pLAFR1 Vector. *DNA Sequence , 15* (3), 225-227.

Watanabe, Kodak, & Makinos. (2006). Cloning, Expression and Characterization of Bacterial L-Arabinose 1 Dehydrogenase Involved in an Alternative Pathway of L-Arabinose Metabolism. *The Journal of Biological Chemistry , 281*, 2612-2623.

Yuan, Zaheer, Morton, & Finan. (2006). Genome Prediction of PhoB Regulated Promoters in *Sinorhizobium meliloti* and Twelve Proteobacteria. *Nucleic Acids Research , 34* (9), 2686-2697.

APPENDIX

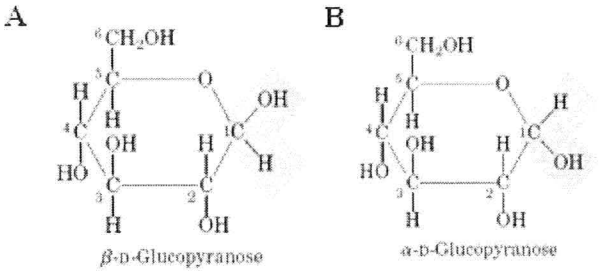


Figure 38: Two possible configurations of a glycosidic bond: the configuration around a hemiacetal carbon (C1) can take one of two forms, β or α, which differ in the orientation of hydroxyl group around C_1 carbon. In this case, glucopyranoside is shown in both configurations. The glycosidic bond is highlighted in pink. Figure adopted from (Lehninger, Nelson, & Cox, 2005).

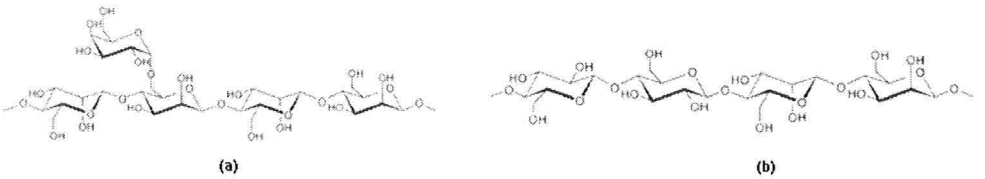


Figure 39: Two types of hemicelluloses with a mannan backbone. (a) Galactomannan. The galactose is branching off the main chain. The main chain consists of mannose units linked via β (1-4) glycosidic bond (b) Glucomannan – has mannan and glucose units connected via β (1-4) glycosidic bond. This arrangement resembles cellulose. This polysaccharide can be broken down by synergistic action of both mannosidase and glucosidase.

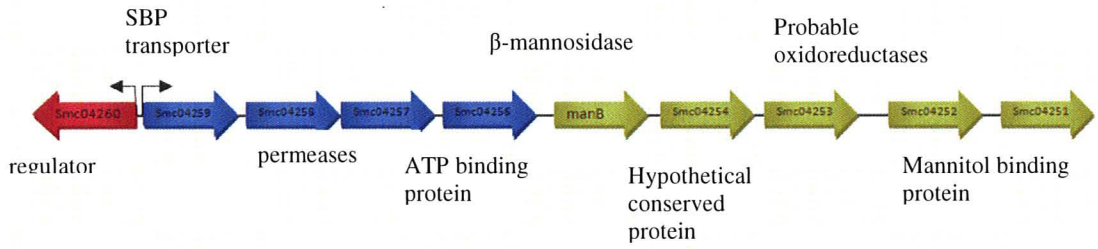


Figure 40: The *smc04260-smc04251* gene cluster. The arrow denotes the direction in which the gene is thought to be transcribed. The black arrows indicate the promoters. The genes are annotated as follows: *smc04260* – putative transcriptional regulator, *smc04259* – putative gentiobiose, cellobiose and salicin ABC transporter, periplasmic solute – binding component, *smc04258* and *smc04257* – putative gentiobiose, cellobiose and salicin ABC transporter, permease component, *smc04256* - putative gentiobiose, cellobiose and salicin ABC transporter, ATP component; *smc04255* – putative beta – mannosidase; *smc04254* – conserved hypothetical protein; *smc04253* – oxidoreductase; smc04352 – probable oxidoreductase; *smc04251* – probable periplasmic mannitol – binding protein. Red colour of the gene designates a putative transcription regulator protein, blue colour designates all putative transporter genes, and green colour designates all the genes believed to be involved in metabolism.

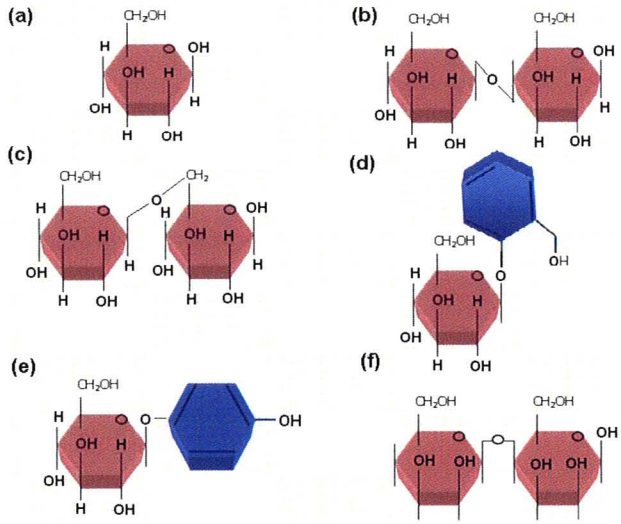


Figure 41: Sugars in this study. (A) β-glucopyranoside. This monosaccharide consists of one glucose unit in a β configuration at its anomeric (C1) carbon. (B) Cellobiose: this molecule consists of identical glucose molecules (such as shown in (A)) linked together via β-1,4-glycosidic bond. These disaccharides are byproducts of incomplete cellulose degradation (C) Gentiobiose: like cellobiose, this molecule consists of two glucose units, this time linked via β-1,6-glycosidic bond. (D) Salicin: this molecule is a glucose – based aryl – β- glycoside. Its glucose unit is linked via β-1,4-glycosidic bond to 2-hydroxy-methylphenyl molecule (E) Arbutin: glucose unit linked via β-1,4-glycosidic bond to hydroquinone (F) Mannobiose: two mannose units connected via β-1,4-glycosidic bond, which means it is a β-mannobiose. These disaccharides are a result of incomplete hemicellulose degradation.

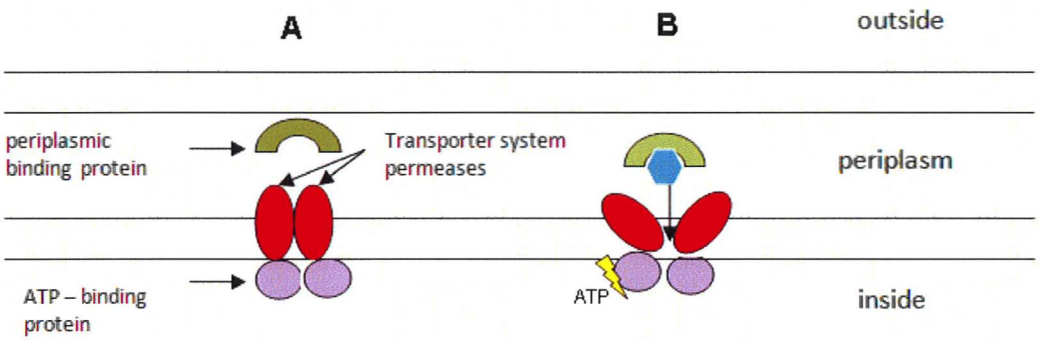
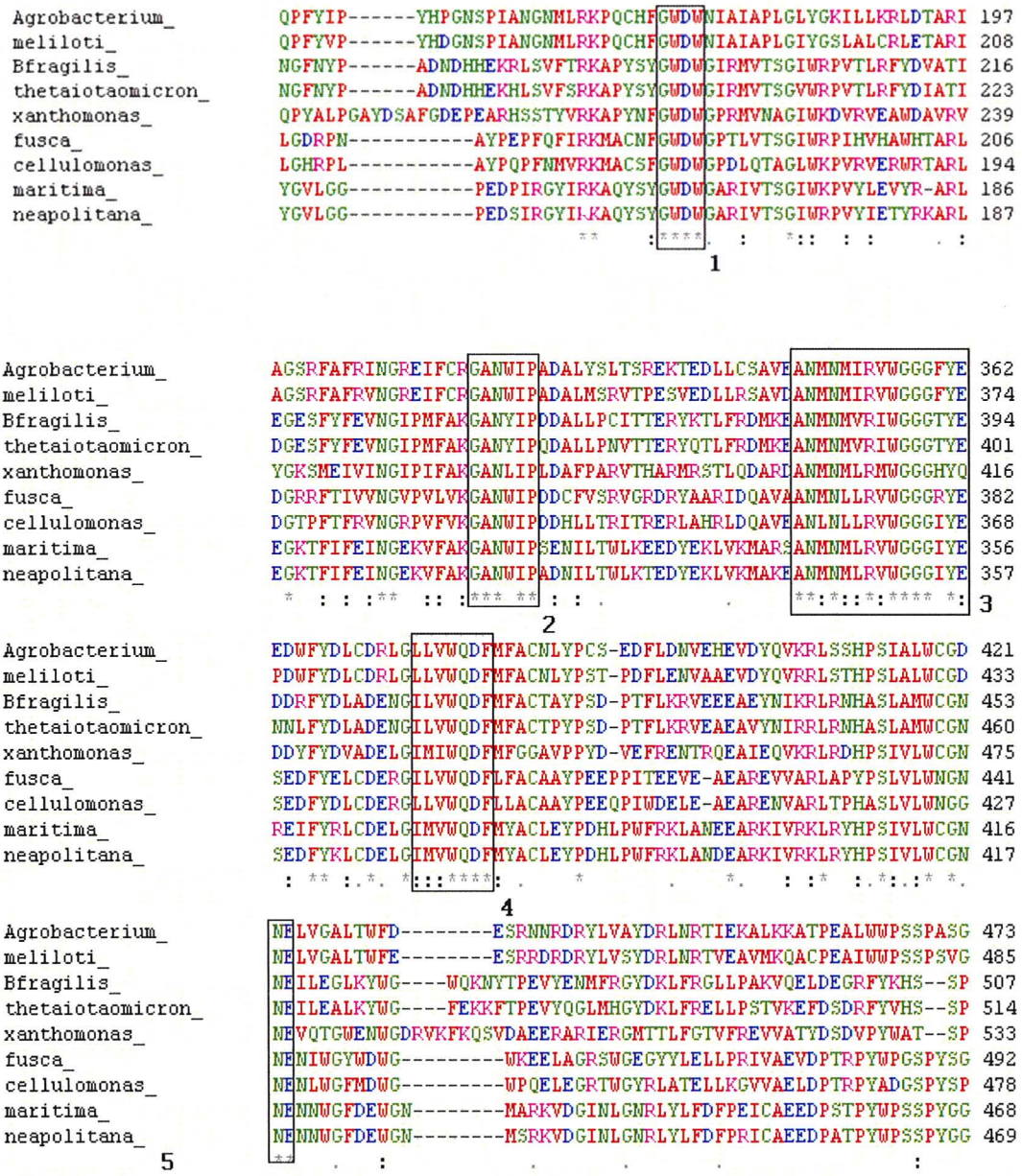


Figure 42: Solute binding protein – dependant ABC transporter system. The ABC-type transporter consists of at least the following components: a solute binding protein that binds a ligand in the periplasm, two transmembrane domains and an ATPase domain that provides energy for the transport system. (A) the system is shown in uninduced/nonfunctional form (B) when the inducing ligand is present (sugar, amino acid, vitamin or others) the solute binding protein binds to the molecule and delivers it to the transmembrane domains that transport this molecule inside the cell using energy derived from ATP hydrolysis. Transport via this system is unidirectional.


Agrobacterium_ QPFYIP------YHPGNSPIANGNMLRKPQCHFGWDWNIAIAPLGLYGKILLKRLDTARI 197
meliloti_ QPFYVP------YHDGNSPIANGNMLRKPQCHFGWDWNIAIAPLGIYGSLALCRLETARI 208
Bfragilis_ NGFNYP------ADNDHHEKRLSVFTRKAPYSYGWDWGIRMVTSGIWRPVTLRFYDVATI 216
thetaiotaomicron_ NGFNYP------ADNDHHEKHLSVFSRKAPYSYGWDWGIRMVTSGVWRPVTLRFYDIATI 223
xanthomonas_ QPYALPGAYDSAFGDEPEARHSSTYVRKAPYNFGWDWGPRMVNAGIWKDVRVEAWDAVRV 239
fusca_ LGDRPN----------AYPEPFQFIRKMACNFGWDWGPTLVTSGIWRPIHVHAWHTARL 206
cellulomonas_ LGHRPL----------AYPQPFNMVRKMACSFGWDWGPDLQTAGLWKPVRVERWRTARL 194
maritima_ YGVLGG----------PEDPIRGYIRKAQYSYGWDWGARIVTSGIWKPVYLEVYR-ARL 186
neapolitana_ YGVLGG----------PEDSIRGYIRKAQYSYGWDWGARIVTSGIWRPVYIETYRKARL 187
1
Agrobacterium_ AGSRFAFRINGREIFCRGANWIPADALYSLTSREKTEDLLCSAVEANMNMIRVWGGGFYE 362
meliloti_ AGSRFAFRVNGREIFCRGANWIPADALMSRVTPESVEDLLRSAVDANMNMIRVWGGGFYE 374
Bfragilis_ EGESFYFEVNGIPMFARGANYIPQDALLPCITTERYKTLFRDMKEANMNMVRIWGGGTYE 394
thetaiotaomicron_ DGESFYFEVNGIPMFARGANYIPQDALLPNVTTERYQTLFRDMKEANMNMVRIWGGGTYE 401
xanthomonas_ YGKSMEIVINGIPIFARGANLIPLDAFPARVTHARMRSTLQDARDANMNMLRMWGGGHYQ 416
fusca_ DGRRFTIVVNGVPVLVRGANWIPQDCFVSRVGRDRYAARIDQAVAANMNLLRVWGGGRYE 382
cellulomonas_ DGTPFTFRVNGRPVFVRGANWIPQDHLLTRITRERLAHRLDQAVEANLNLLRVWGGGIYE 368
maritima_ EGKTFIFEINGEKVFARGANWIPSENILTWLKEEDYEKLVKMARSANMNMLRVWGGGIYE 356
neapolitana_ EGKTFIFEINGEKVFARGANWIPADNILTWLKTEDYEKLVKMAKEANMNMLRVWGGGIYE 357
2 3
Agrobacterium_ EDWFYDLCDRLGLLVWQDFMFACNLYPCS-EDFLDNVEHEVDYQVKRLSSHPSIALWCGD 421
meliloti_ PDWFYDLCDRLGLLVWQDFMFACNLYPST-PDFLENVAAEVDYQVRRLSTHPSLALWCGD 433
Bfragilis_ DDRFYDLADENGILVWQDFMFACTAYPSD-PTFLKRVEEEAEYNIKRLRNHASLAMWCGN 453
thetaiotaomicron_ NNLFYDLADENGILVWQDFMFACTPYPSD-PTFLKRVEAEAVYNIRRLRNHASLAMWCGN 460
xanthomonas_ DDYFYDVADELGIMIWQDFMFGGAVPPYD-VEFRENTRQEAIEQVKRLRDHPSIVLWCGN 475
fusca_ SEDFYELCDERGILVWQDFLFACAAYPEEPPITEEVE-AEAREVVARLAPYPSLVLWNGN 441
cellulomonas_ SEDFYDLCDERGLLVWQDFLLACAAYPEEQPIWDELE-AEARENVARLTPHASLVLWNGG 427
maritima_ REIFYRLCDELGIMVWQDFMYACLEYPDHLPWFRKLANEEARKIVRKLRYHPSIVLWCGN 416
neapolitana_ SEDFYKLCDELGIMVWQDFMYACLEYPDHLPWFRKLANDEARKIVRKLRYHPSIVLWCGN 417
4
Agrobacterium_ MELVGALTWFD--------ESRNNRDRYLVAYDRLNRTIEKALKKATPEALWWPSSPASG 473
meliloti_ MELVGALTWFE--------ESRRDRDRYLVSYDRLNRTVEAVMKQACPEAIWWPSSPSVG 485
Bfragilis_ MEILEGLKYWG----WQKNYTPEVYENMFRGYDKLFRGLLPAKVQELDEGRFYKHS--SP 507
thetaiotaomicron_ MEILEALKYWG----FEKKFTPEVYQGLMHGYDKLFRELLPSTVKEFDSDRFYVHS--SP 514
xanthomonas_ MEVQTGWENWGDRVKFKQSVDAEERARIERGMTTLFGTVFREVVATYDSDVPYWAT--SP 533
fusca_ MENIWGYWDWG---------WKEELAGRSWGEGYYLELLPRIVAEVDPTRPYWPGSPYSG 492
cellulomonas_ MENLWGFMDWG---------WPQELEGRTWGYRLATELLKGVVAELDPTRPYADGSPYSP 478
maritima_ MENNWGFDEWGN--------MARKVDGINLGNRLYLFDFPEICAEEDPSTPYWPSSPYGG 468
neapolitana_ MENNWGFDEWGN--------MSRKVDGINLGNRLYLFDFPRICAEEDPATPYWPSSPYGG 469
5

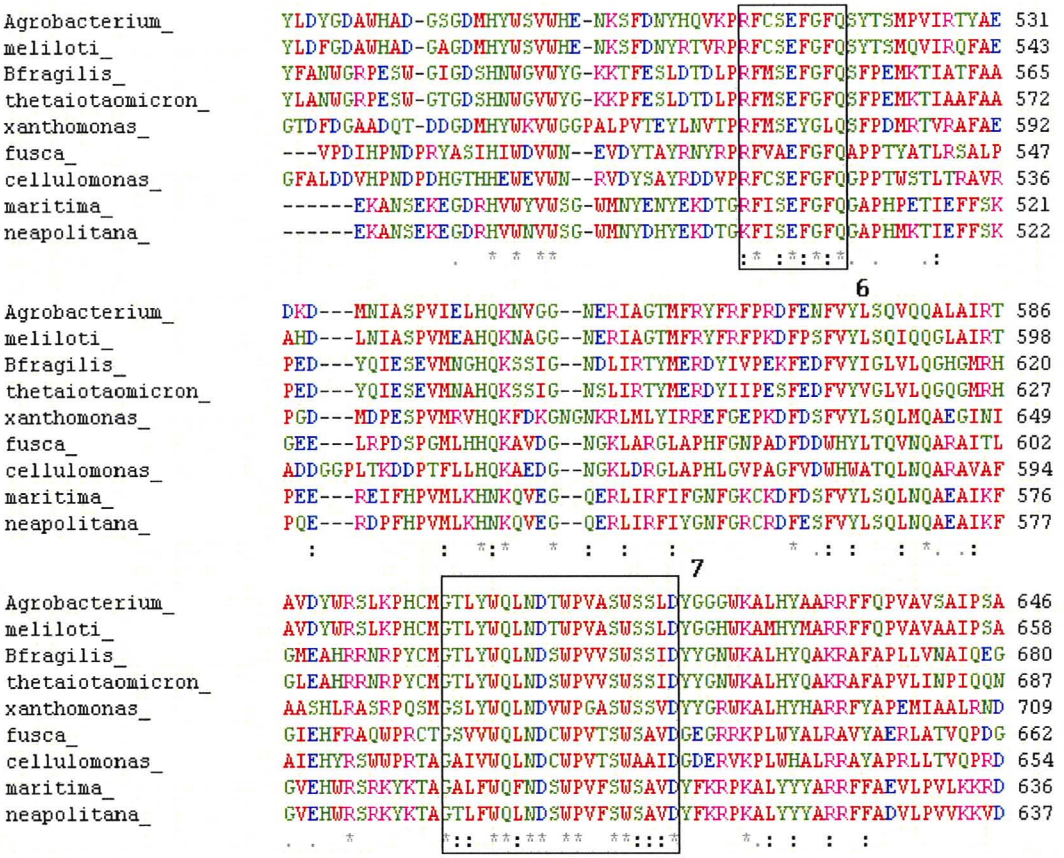


Figure 43: Multiple alignment of known and putative β-mannosidases. The organisms are listed in this order (from top to bottom): *Agrobacterium tumefaciens* putative, *S.meliloti* putative, *Bacteroides fragilis* putative, *Bacteroides thetaiotaomicron*, known (Tailford, Money, Smith, Dumon, Davies, & Gilbert, 2007), *Xanthamonas campestris* putative, *Thermobifida fusca* known (Beki, et al., 2003), *Cellulomonas fimi* known (Stoll, Stalbrand, & Warren, 1999), *Thermatoga maritima* known (Parker, et al., 2001) and *Thermatoga neapolitana* known (Duffaud, McCutchen, Leduc, Parker, & Kelly, 1997). Only the areas of most importance are shown. The boxed sequences are sequences of highest similarity. Site 1 is showing a WDW motif that was found to be conserved in *B.thetaiotaomicron* and is believed to be important in interactions with substrate (Tailford, Money, Smith, Dumon, Davies, & Gilbert, 2007). The multiple alignment has been made using ClustalW tool on EMBL web site (http://www.ebi.ac.uk/Tools/msa/clustalw2/).

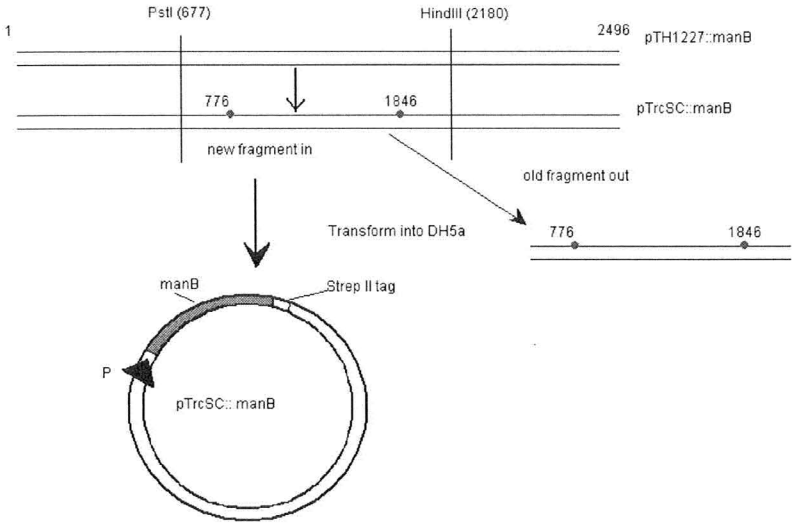


Figure 44: Correction of point mutations at sites 776 and 1846 in *manB* gene. A small fragment of each vector where *manB* has been cloned is shown. Both of these are cut with the same restriction sites, *PstI* and *HindIII.* The fragment with mutations is therefore substituted with the fragment that is error free. The result is vector pTH2493.

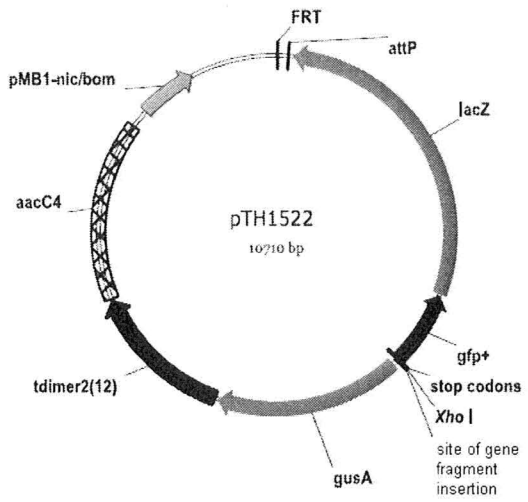
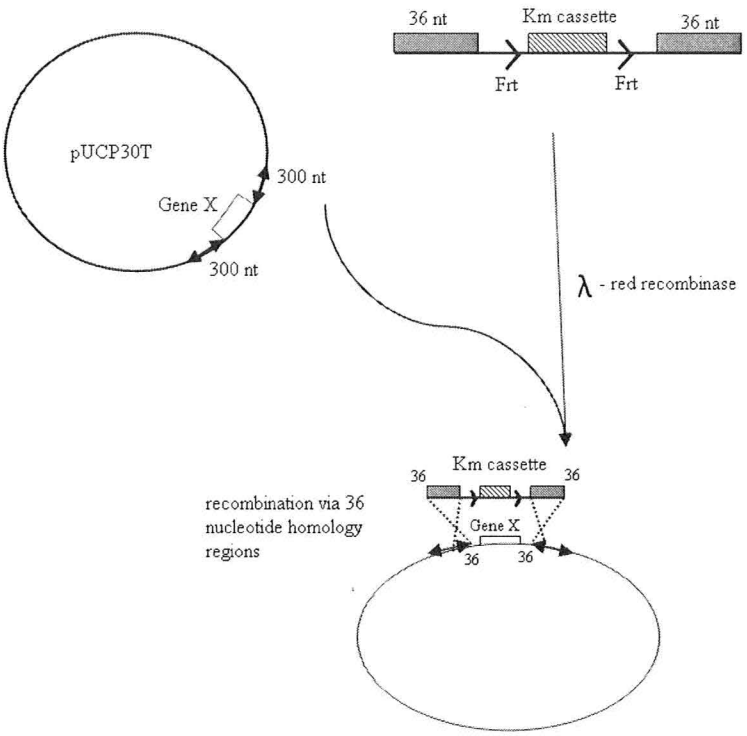
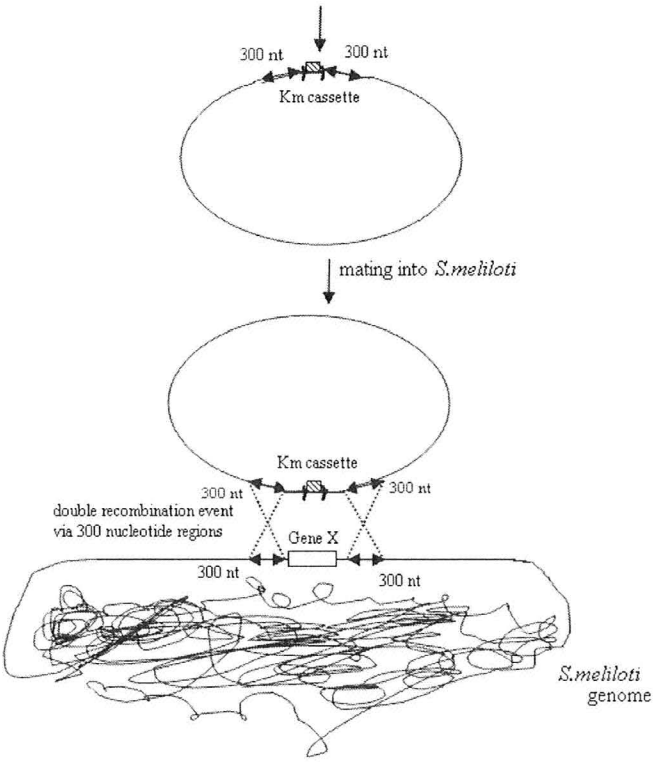


Figure 45: Plasmid pTH1522 used to construct library fusion strains. The site at which the *S.meliloti* fragment was inserted is indicated. The reporter genes are located around that site: *gusA* and *tdimer2* in one direction, and *gfp* and *lacZ* in another direction. Figure adopted from (Cowie A. , et al., 2006).

A.


36 nt
Km cassette
36 nt
Frt
Frt
pUCP30T
300 nt
Gene X
300 nt
λ - red recombinase
Km cassette
36
36
recombination via 36
nucleotide homology
regions
Gene X
36
36

B.


300 nt
300 nt
Km cassette
mating into S.meliloti
Km cassette
300 nt
300 nt
double recombination event
via 300 nucleotide regions
Gene X
300 nt
300 nt
S.meliloti
genome

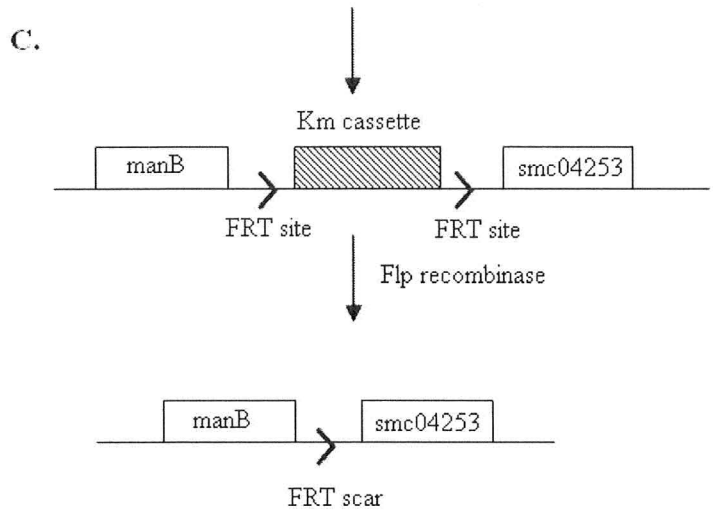


Figure 47: In frame deletion of gene *smc04254*. The diagrams are not to scale relative to each other. **A.** Plasmid pTH2615 has a large region of *S.meliloti* DNA inserted at the *Eco*RI site. That region spans *smc04260* to *smc04247* and includes gene *smc04254*. DNA fragment was amplified via primers ML-09-2919 and ML-09-2995 that contains 36 nuceotides directly upstream and downstream from *smc04254* and has FRT sites surrounding 790 nucleotide Kanamycin cassette from Tn5 via template plasmid pKD4 (Datsenko & Wanner, 2000). Recombination via 36 nucleotide regions occurred between the DNA fragment and pTH2615 in the presence of λ red recombinase. **B.** The resultant plasmid pTH2636 is pTH2615 (previously described in Materials and Methods section), except *smc04254* has been substituted with Km cassette flanked by FRT sites. pTH2636 is mated into *S. meliloti* along with plasmid pPH1JI, which is incompatible with pTH2636. In order to be able to remain in *S. meliloti*, pTH2636 must undergo homologous recombination. The double homologous recombination event was selected for because pTH2636 is not present within the genome although *smc04254* has been replaced with Km cassette. **C.** The region of *S.meliloti* DNA is shown, where the recombination has occurred. Using Flp recombinase (via pTH2505), Km cassette is removed via FRT sites. This results in the FRT scar where *smc04254* had been.

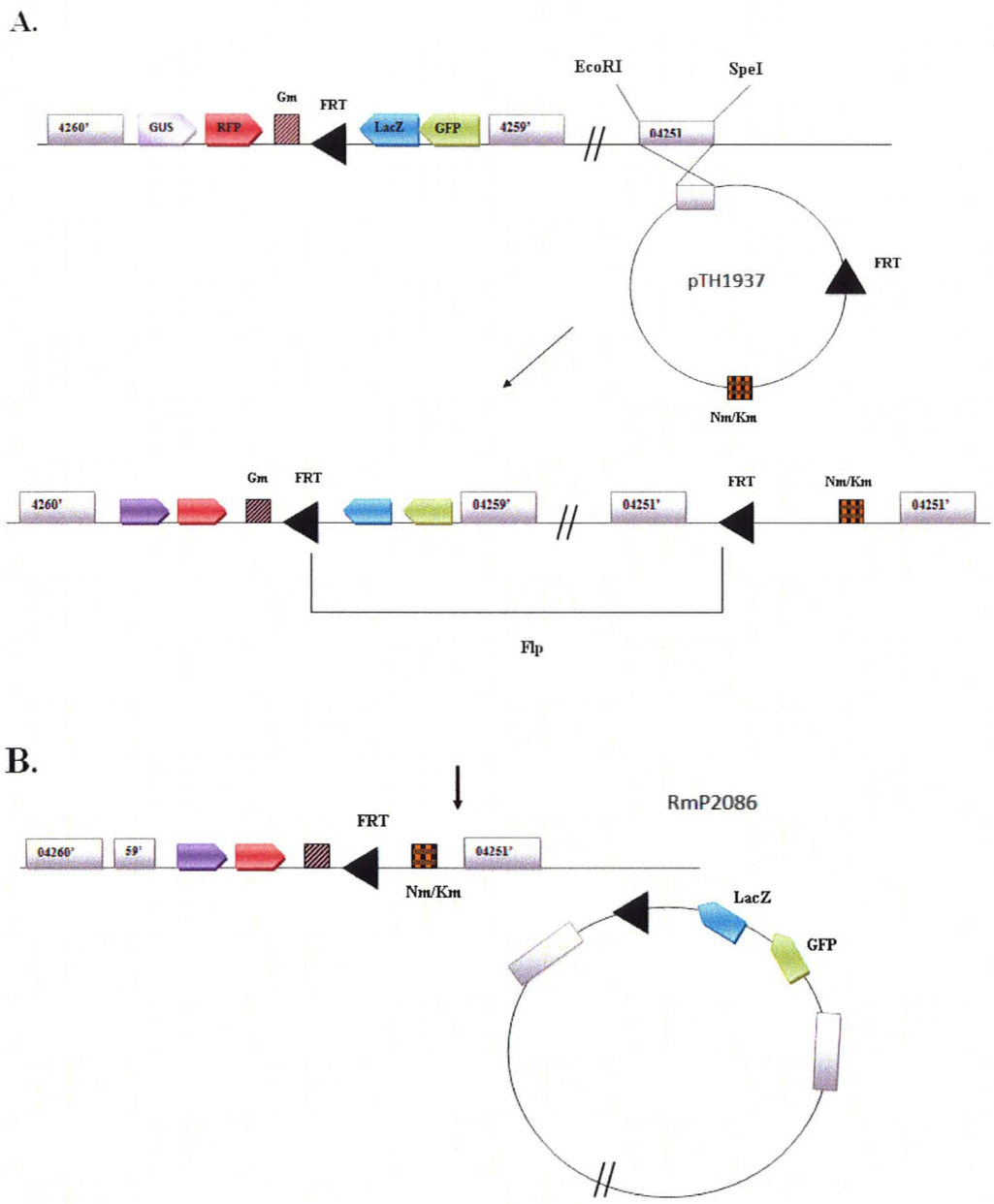


Figure 48: Construction of strains RmP2086 and RmP2202. A. The initial strain used for construction was fusion library strain SmFL1580, which contains an FRT site. Sites *Eco*RI and *Spe*I were used to clone a region around *smc04251* into pTH1937. This construct was later recombined into SmFL1580 to introduce a second FRT site. B. FlP recombinase was used in order to delete all the genes in-between the two FRT sites. These genes included *smc04259* to *smc04252*, inclusive.

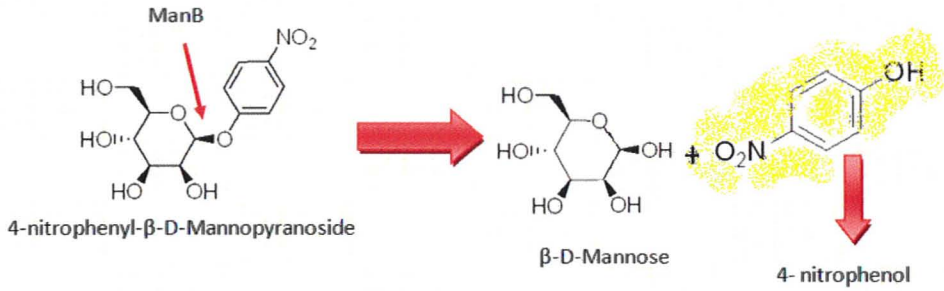


Figure 49: Principle of an assay with p-nitrophenyl linked substrates to test for activity of ManB. ManB cleaves the substrate 4-nitrophenyl-β-D-mannopyranoside (PM) and the products are mannose and 4-nitrophenol. The rate at which 4-nitrophenol is released was measured at 350 nm, the isosbestic point, due to the range of pH at which the assay was run.

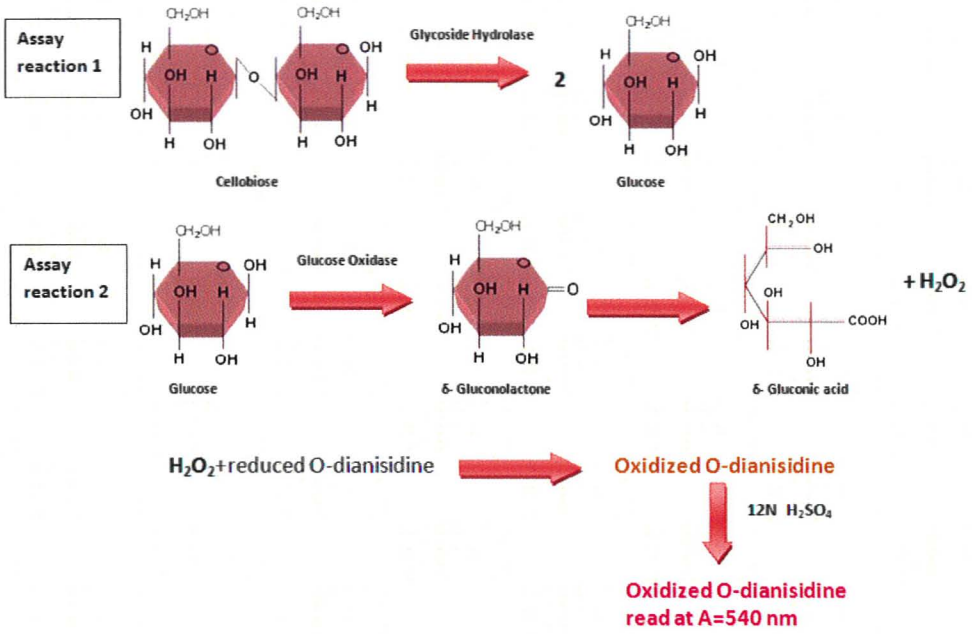


Figure 50: A diagram showing the procedure for detecting glucose using glucose oxidase/peroxidase method. Assay reaction 1 illustrates the breakdown of cellobiose substrate. The product of this reaction, glucose, can be quantified as a measure of the activity of the enzyme that breaks it down. In order to stop the reaction at a particular time point, the reaction tube was boiled at 100 °C for 5 minutes in order to inactivate all glycoside hydrolyzing enzymes. After sufficient cooling the glucose oxidase assay was carried out, as per manufacturer's instructions to quantify the amount of glucose liberated from breakdown of cellobiose (and other) substrates.

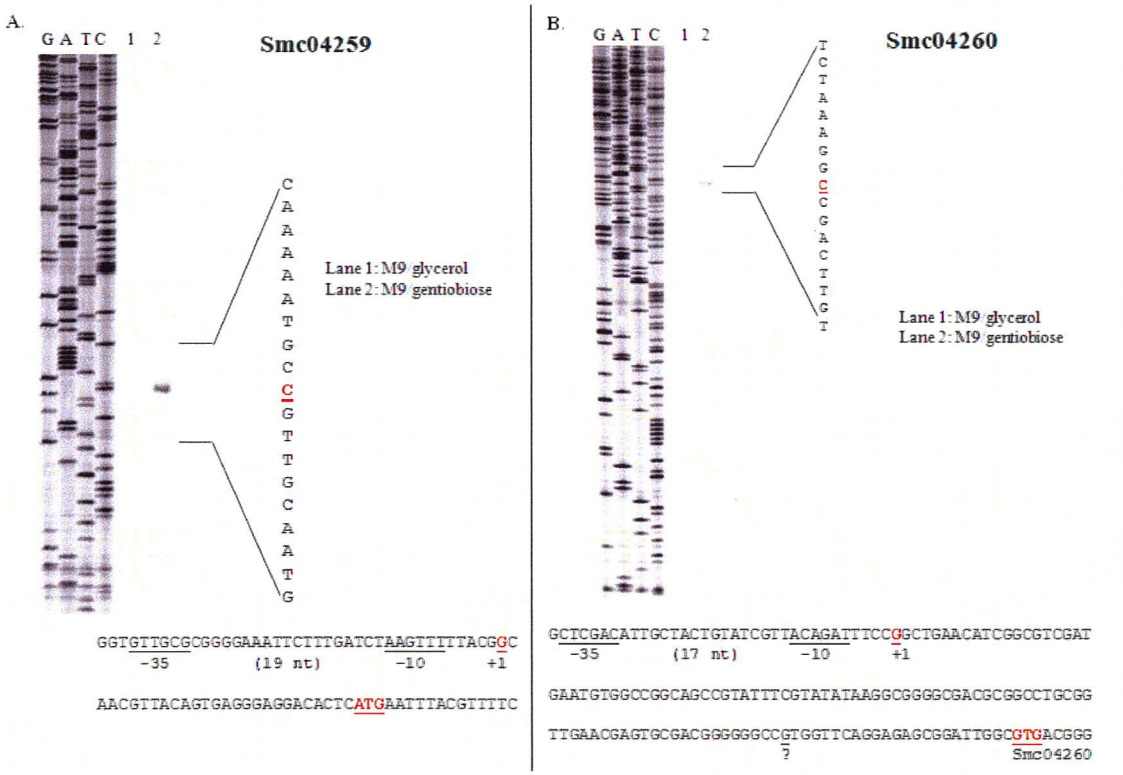


Figure 51: Primer extension to map the promoters upstream of *smc04259* and *smc04260*. Lane 1: M9 minimal media with 5% glycerol, Lane 2: M9 minimal media with 5 mM gentiobiose. **A.** Primer extension for upstream region of *smc04259* and **B.** Primer extension for upstream region of *smc04260*. Gentiobiose was used to induce the transcription. Primer extension was carried out by Dr. Cathy White, figure obtained from Dr. Cathy White.