EXPRESSION ANALYSIS OF ABC TRANSPORTERS IN SINORHIZOBIUM MELILOTI

EXPRESSION ANALYSIS OF ABC TRANSPORTERS IN SINORHIZOBIUM MELILOTI

By

JANE ELIZABETH FOWLER, B.Sc

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Science

McMaster University

© Copyright by Jane Fowler, December 2004

MASTER OF SCIENCE
(Biology)

MCMASTER UNIVERSITY
Hamilton, Ontario

TITLE: Expression Analysis of ABC Transporters in *Sinorhizobium meliloti*

AUTHOR: Jane Elizabeth Fowler, B.Sc.

SUPERVISOR: Professor T.M. Finan

NUMBER OF PAGES: 130 pages

ABSTRACT

Soils contain a complex mixture of compounds many of which can be transported and metabolized by microorganisms. *Sinorhizobium meliloti* is a soil bacterium whose gene sequence was recently determined. The diversity of the carbon and nitrogen sources that can be utilized by this organism is reflected in the large number of annotated ATP-binding cassette transporters in its genome. Although many of these genes are not necessary for survival, it is hypothesized that they aid in the competitive fitness of *S. meliloti* in the field. Many of these transporters remain uncharacterized.

In this study a high throughput screen was developed to measure β-glucuronidase activity in a 96 well microtitre plate format to quantify expression of many reporter gene fusions under a variety of conditions. This system was used to analyze the expression of putative small molecule ABC transporters in *S. meliloti*. 45 *gusA* reporter gene transcriptional fusions to these transport genes were generated and recombined into the genome. These strains were grown in 96 well plates in minimal media containing a large number of carbon sources and various legume root and seed exudates to be tested as inducers of transporter gene expression. Two transport systems were found to be induced by glucosamine and galactosamine and others were found to be induced by various sugars including mannose, arabinose, xylose and palatinose as well as protocatechuate and hydroxybenzoate.

The bacteria-plant symbiosis of *S. meliloti* and alfalfa plays an important role in agriculture. To further understand the role of ABC transporters in the competition of *S. meliloti* the *gusA* reporter fusions strains will also be inoculated onto alfalfa roots and

nodules will be assayed for GusA activity to give a more complete picture of the role of ABC transporters in the competition and symbiosis of *S. meliloti*.

Two substrates, galactitol and hydroxyproline found to induce transport systems were studied in depth in order to more fully understand the transport, metabolism and regulation of these compounds.

ACKNOWLEDGEMENTS

I would like to thank Dr. Turlough Finan for the opportunity to work in this lab which has been a great experience for the past two years.

I thank all the members of the Finan lab and especially those who have become true friends. I greatly appreciate all the fun, laughter, great discussions and ideas we have shared.

I would also like to thank my family, friends and boyfriend who have given me love and support and have provided a distraction from the lab. A special thanks to my mother Louise who has given me endless love, emotional and financial support throughout my schooling and who has taught me the importance of education.

TABLE OF CONTENTS

TITLE	PAGE
Descriptive note	ii
Abstract	iii
Acknowledgements	v
Table of contents	vi
List of Figures	vii
List of Tables	xii
Abbreviations	xiii
Chapter 1: Introduction	1
Sinorhizobium meliloti	1
ABC transporters	2
Transport in *S. meliloti*	6
β-glucuronidase	7
This work	9
Chapter 2: Materials and Methods	11
Bacterial strains and growth conditions	11
β-glucuronidase assay	30
High-throughput β-glucuronidase assay	31
β-galactosidase assay	32
Plant growth conditions	32
β-glucuronidase assay of nodules	33
Protein determination in nodule sample	34
Preparation of seed and root exudates	35
Bacterial matings	35
Plasmid preparation	36
DNA modifications	36
Ligation reactions	36
Competent cell preparation	37
Transformation	37
Agarose gel electrophoresis	37
Genomic DNA preparation	38
Southern transfer	38
Preparation of random primed DNA probe	39
Hybridization of radioactively labelled probe to DNA	40
Detection	40
PCR	41
Colony PCR	41
DNA sequencing and analysis	42

Chapter 3: Results
3.1. Deletion mutant screening 43
3.2. Properties of β-glucuronidase assay 45
3.3. Generation of *gusA* reporter gene fusions 49
3.4. The high-throughput β-glucuronidase screening 54
3.5. Results of high-throughput screens 55
3.6. Legume seed and root exudates 65
3.7. Plant screening 66

Chapter 4: Galactitol and Tagatose 68
4.1. Introduction 68
4.2. Results 69

Chapter 5: Hydroxyproline 75
5.1. Introduction 75
5.2. Results 80

Chapter 6: Discussion 96
6.1. Biolog and deletion mutant screening 96
6.2. Reporter fusions 98
6.3. The high-throughput β-glucuronidase assay 98
6.4. Screening results 100
6.5. Plant screening 105
6.6. Galactitol and Tagatose 109
6.7. Hydroxyproline 112
6.8. Conclusions 116

References 118

Appendix Biolog results 125

LIST OF FIGURES

Figure	Title	Page
3-1	ß-glucuronidase assay of RmP187 grown in LBmc and M9 minimal media with succinate or glucose as the sole carbon sources	46
3-2	β-glucuronidase assay comparing effects of washed vs. resuspended cell cultures in β-glucuronidase assays	46
3-3	β-glucuronidase assay comparing the effects of vortexing vs. no physical agitation after addition of permeabilizing agent	47
3-4	β-glucuronidase assay comparing the effects of reaction temperature	48
3-5	β-glucuronidase assay comparing the effects of culture storage time and temperature	48
3-6	Plasmid map of *gusA* reporter suicide vector, pTH1360	50
3-7	Single cross-over occurring in order to create *gusA* fusion strains under the control of the native ABC transport operon promoter	51
3-8	Single cross-over in order to dissociate a promoter from its operon to create a knock-out strain	52
3-9	Southern Blot of total DNA from RmP199, RmP187 and RmP189 reporter fusion strains using a *SmaI* and *BglII* digests	53
3-10	Predicted restriction maps of target regions after recombination of the plasmids used in the construction of strains RmP199, RmP187 and RmP189 for a *BglII* digest	54
3-11	ß-Glucuronidase assay of RmP193 (SMb21138::*gusA*) grown in M9 minimal media with various test carbon sources	56
3-12	ß-Glucuronidase assay of RmP198 (SMb20979::*gusA*) grown in M9 minimal media with various test carbon sources	57
3-13	ß-glucuronidase assay of RmP226 (SMb20428::*gusA*) grown in M9 minimal media with various test carbon sources	58
3-14	ß-glucuronidase assay of RmP203 (SMb21342::*gusA*) grown in M9 minimal media with various test carbon sources	58
3-15	. ß-Glucuronidase assay of RmP190 (SMb20235::*gusA*) grown in M9 minimal media with various test carbon sources	59
3-16	ß-glucuronidase assay of RmP206 (SMb20904::*gusA*) in order to confirm results obtained in the high-throughput method of screening	60
3-17	β-glucuronidase assay of RmP198 (SMb20979::*gusA*) grown in M9 minimal media with fucose and succinate as the sole carbon sources	61
3-18	β-glucuronidase assay of RmP266 (SMb20428::*gusA*) grown in M9 minimal media with fucose, mannose and succinate as the sole carbon sources	61
3-19	β-glucuronidase assay of RmP192 (SMb21707::*gusA*) grown in M9 minimal media with glucosamine, galactosamine and N-acetyl-D-glucosamine as the sole carbon sources	62
3-20	β-glucuronidase assay of RmP202 (SMb21342::*gusA*) grown in M9 minimal media sorbose, raffinose, glucose and succinate as the sole	62

	carbon sources	
3-21	β-glucuronidase assay of RmP222 (SMb21587::*gusA*) l grown in M9 minimal media arabinose, galactose and succinate as the sole carbon sources	63
3-22	ß-glucuronidase assay of RmP188 (SMb20784::*gusA*) grown in M9 minimal media with protocatechuate and p-hydroxybenzoate as the sole carbon sources	64
3-23	ß-glucuronidase assay of RmP230 (SMb20027::*gusA*) grown in M9 minimal media with protocatechuate and p-hydroxybenzoate as the sole carbon sources	64
3-24	ß-glucuronidase assay of RmP193 (SMb21138::*gusA*) and RmP194 (SMb21216::*gusA*) grown in alfalfa and pea seed exudates	65
3-25	ß-glucuronidase assay of RmP193 (SMb21138::*gusA*) grown in various carbon sources	66
3-26	ß-glucuronidase assay of nodules extracts from alfalfa plants inoculated with *S. meliloti* strains containing GusA fusions	67
3-27	ß-glucuronidase assay *gusA* fusion strains grown in LBmc media	67
4-1	ß-glucuronidase assay of RmP187 (SMb21373::*gusA*) and RmP195 (SMb21375::*gusA*) in M9 minimal media with galactose as the sole carbon source.	69
4-2	Gene map of the operon shown to be induced by galactose, galactitol, tagatose and sorbose and location of *gusA* fusions and ΩSp-Sm cassette insertion in SMb21372	69
4-3	ß-glucuronidase assay of RmP187 (SMb21373::*gusA*) in M9 minimal media with galactitol and tagatose.	70
4-4	ß-glucuronidase assay of RmP187 (SMb21373::gusA) in M9 minimal media with sorbose as the sole carbon source	71
4-5	ß-Glucuronidase assay RmP244 ($P_{SMb21377::gusA}$) grown in M9 minimal media with galactitol, sorbose and tagatose	72
4-6	Galactose pathway via galactitol and D-tagatose	72
4-7	ß-Glucuronidase assay of SMb21377::*gusA* with (RmP245) and without (RmP233) a Sp-Sm cassette interrupting SMb21372, a putative regulator grown in M9 minimal media containing galactitol and succinate	74
5-1	Reaction of proline to hydroxyproline catalyzed by prolyl 4-hydroxylase	77
5-2	ß-Glucuronidase high-throughput assay of RmP215 (SMb20263::*gusA*) showing induction by hydroxyproline	81
5-3	ß-Glucuronidase high-throughput assay of RmP217 (SMb20320::*gusA*) showing induction by hydroxyproline	81
5-4	Growth curve of Sm1021, RmP215 and RmP238 and RmF909 grown in M9 minimal media with hydroxyproline as the sole carbon source	82
5-5	Growth curve of Sm1021, RmP215 and RmP238 and RmF909 grown in M9 minimal media with allohydroxyproline as the sole carbon source	83
5-6	Growth curve of Sm1021, RmP215 and RmP238 and RmF909 grown in M9 minimal media with succinate as the sole carbon source	83

5-7	β-glucuronidase assay of RmP215and RmP217 grown in M9 minimal media with hydroxyproline or proline as the sole nitrogen/carbon source	84
5-8	Gene map of an operon on pSymB induced by hydroxyproline and location of gene fusions	85
5-9	Gene map of an operon on pSymB induced by hydroxyproline and location of gene fusions	85
5-10	ß-glucuronidase assay of intergenic sequences of SMb20261 (RmP243), SMb20262 (RmP242) and SMb20263 (RmP241) in a replicating GusA reporter vector grown in minimal media with succinate and hydroxyproline	86
5-11	ß-galactosidase assay of SmRL43	87
5-12	ß-glucuronidase assay of RmP239, RmP240, RmP215 and RmP238 grown in M9 minimal media with allohydroxyproline and succinate as the carbon sources	88
5-13	ß-glucuronidase assay of the promoter regions of SMb20261, SMb20262 and SMb20263 in a SMb20262 functional knock-out background) versus a wild-type background.	89
5-14	ß-glucuronidase assay of the promoter region of SMb20262 in a replicating gusA reporter plasmid (pTH1801) in a wild type background (RmP242) and a ntrC- background (RmP315) grown in M9 minimal media with hydroxyproline as the carbon source and succinate as the carbon source with and without nitrogen (NH_4Cl).	90
5-15	β-glucuronidase assay of RmP215and RmP238 grown in M9 minimal media containing hydroxyproline and glucose or succinate to investigate the catabolic repression of these two compound on hydroxyproline induction	91
5-16	β-glucuronidase assay of RmP215 (SMb20263::*gusA*) assay grown in alfalfa root and seed exudates.	91
5-17	Predicted restriction maps of target regions after recombination of the plasmids used in the construction of strain RmP239 (SMb20261::*gusA*) for an *EcoRV* digest.	93
5-18	Predicted restriction maps of target regions after recombination of the plasmids used in the construction of strain RmP240 for a *SmaI* and *BamHI* digest.	94
5-19	Predicted restriction maps of target regions after recombination of the plasmids used in the construction of strain RmP215for an *EcoRV* digest.	95
6-1	High-throughput β-glucuronidase assay developed for screening in 96-well microtitre plates	99
6-2	Gene map of SMb20904, part of an ABC transport system found to be induced by glucose, mannose and xylose	101
6-3	Gene map of SMb20784, part of a transport operon induced by pca and pob located upstream of the *pcaBGHCD*	105
6-4	Hydroxyproline metabolic pathway and annotation of operon with hypothesized gene function.	115

LIST OF TABLES

Table	Title	Page
1-1	Number of transport proteins and ABC transporters in bacteria	6
2-1	Plasmids used in this study	13
2-2	*E. coli* strains used in this study	16
2-3	*S. meliloti* strains used in this study	19
2-4	Primers used in this study	22
3-1	Carbon sources not utilized by pSymB deletion mutants	44
3-2	Substrates and exudates used in the high-throughput screen to investigate expression of fusion strains	55
4-1	Results of growth on M9 minimal media agar plates containing various carbon sources	74
4-2	Frequency of reversion mutants of galactitol mutants	74

ABBREVIATIONS

Km	kanamycin
Cm	chloramphenicol
Sp	spectinomycin dihydrochloride
Tc	tetracyclin
Gm	gentamicin sulphate
Rm	Rifampicin
PNPG	p-nitrophenyl β-D-glucuronide
ONPG	o-nitrophenyl β-D-galactopyranoside
X-Gluc	5-bromo-4-chloro-3-indoyl β-D-glucuronide
DTT	DL-Dithiothreitol
SDS	sodium dodecyl sulphate
EDTA	(ethylenedinitrilo)-tetraacetic acid
bp	base pair
kb	kilobase
OD	optical density
BSA	bovine serum albumin

CHAPTER 1. LITERATURE REVIEW

Sinorhizobium meliloti

Sinorhizobium meliloti is a gram negative, alpha-proteobacteria that forms nitrogen fixing root nodules on *Medicago sativa* (alfalfa). Through this mutually beneficial relationship the metabolic needs of the bacterium are met by the plant. *S. meliloti* belongs to the *Rhizobiacae* which also included the genera *Mesorhizobium* and *Rhizobium* and *Bradyrhizobium*. The members of this family are of great economic importance due to their important role in agriculture. Nitrogen fixation is the process where atmospheric N_2 gas is converted to ammonia (NH_3) which is the form usable by plants. The symbiotic relationship between nitrogen fixing rhizobia and their host-plant allows the plant to grow under nitrogen deficient conditions.

The symbiotic process begins with recognition and colonization of the host plant root. Root hair curling leads to infection thread formation and subsequently nodule formation. Within the central zone of the nodule, nitrogen fixation is carried out. Genes essential to symbiosis are comprised of the *nod*, *nif* and *fix* genes. The nodulation (*nod*) genes are needed for root hair curling, infection thread and nodule formation. The *nif* and *fix* genes are required for nitrogen fixation, as reviewed in (van Rhijn and Vanderleyden 1995). The process of symbiosis is relatively well characterized in the rhizobia but comparatively little is known as to how this soil bacterium competes and proliferates in the soil.

S. meliloti must adapt to a variety of extreme environments including the oligotrophic bulk soil, the nutrient rich rhizosphere and the highly specific environment

of the root nodule. This bacterium has evolved to become a highly adaptive and competitive microorganism of the soil. *S. meliloti* has a tri-partite genome made up of a chromosome and two megaplasmids, pSymA and pSymB. pSymA contains many of the genes needed for nitrogen fixation whereas the roles of pSymB are more elusive. Most of pSymB has been deleted and it contains very few essential genes (Charles and Finan 1991). This lends itself to the hypothesis that pSymB is needed for survival and competition in the soil environment.

The entire *S. meliloti* chromosome has been sequenced and annotated and pSymB was found to be mainly composed of transport (~ 20%) and metabolism genes (~ 25%) (Galibert et al. 2001; Finan et al. 2001). With the completion of the sequence and annotation of the *S. meliloti* genome we can now focus on the analysis of gene function.

ABC Transporters

The ABC (ATP binding cassette) transport system is a major class of cellular transport machinery. ABC transporters form the largest group of paralogous genes in bacterial and archaeal genomes (Tomii and Kanehisa 1998). ATP binding cassette domains are involved in physiological processes in all eukaryotic, archaebacterial and bacterial species. They are involved primarily in import and export across the cytoplasmic membrane. All ABC transporters are composed of 4 domains: 2 transmembrane permeases and 2 ATPases. These 4 domains can be expressed as 4 different proteins or in various combinations of fusions (Linton and Higgins 1998). The genes for the components of these transport systems are often found in an operon

(Higgins 1992). The two transmembrane domain (TMD) permeases contain membrane-spanning α-helices (usually 6 each) and together form a pore in the membrane. These proteins are what confer substrate specificity on the transporter. The only conserved motif present on permeases is on a cytoplasmic loop which is thought to interact with the ATP-binding domain. The two ATP or nucleotide–binding domain (NBD) proteins are hydrophilic and are found on the cytoplasmic side of the membrane. This domain contains the signature ABC domain by which this class of transporters is defined. ATP-binding domains from different transporters share between 30 and 50% similarity and the signature domain is more extensive than the two short Walker domains (Higgins 1992). This domain of approximately 215 amino acid residues contains Walker A and B motifs and also a C motif just upstream of the Walker B sequence and is responsible for the hydrolysis of ATP producing the energy needed to move the solute across the membrane (Jones and George 1999). Not every protein with a Walker domain is an ABC transporter although every ABC transporter contains a Walker domain. Although most ABC domains indicate a function in transport there are some cases in which the domain is not transport associated. The UvrA protein which functions in DNA repair contains two ABC domains and hydrolyses ATP as part of its repair function. Also the elongation factor EF-3 of yeast contains an ABC domain (Higgins 1992).

In gram negative bacteria ABC uptake systems contain an additional protein, the periplasmic solute binding protein. In gram positive bacteria a membrane bound lipoprotein acts similarly to the periplasmic binding protein whereas eukaryotic cells have no homologues to these two proteins. The presence of a solute binding protein is the

only indication of an import versus export system. The periplasmic binding proteins (PBPs) all share a similar structure with two globular domains with a cleft in between that contains the substrate binding site. All substrates have been found to bind to the PBP through hydrogen bonds no matter what the nature of the substrate (Higgins 1992). The PBP is the initial receptor for the transported ligand and brings the substrate to the membrane bound components. The *in vitro* binding specificities and affinities measured from purified PBP's correspond to those for the *in vivo* transport processes which suggests that the PBP is the rate limiting step for transport (Tam and Saier, Jr. 1993). ABC transporters can maintain very high solute concentrations because of the PBPs which can scavenge very low concentrations of solute in order to accumulate these to relatively high internal concentrations (Higgins 1992). PBPs have also been found to function in chemoreception and also sensory transduction. A substrate bound PBP which has undergone a conformational change may interact with transmembrane chemoreception proteins which span the membrane twice or with a transmembrane protein to transmit a signal to cytoplasmic constituents (Tam and Saier, Jr. 1993).

The ChvE PBP of *Agrobacterium tumefaciens* is an example of a protein that carries out multiple functions in which it may interact with at least 12 different sugars, including arabinose, galactose, glucose, ribose and fucose and 3 different membrane associated proteins. Monosaccharides released by a host plant wound are bound by the PBP and this then interacts with the periplasmic region of the membrane bound, sensor kinase VirA. VirA is part of the VirA-VirG sensor-regulator to activate transcription of the *vir* regulon. ChvE mutants are avirlulent. ChvE also mediates chemotaxis when

bound to a substrate by interacting with a membrane bound receptor. This multi-functional protein is also part of a high affinity uptake system (*ggu*) (Kemner et al. 1997).

A multitude of solutes are found to be imported by ABC transporters including sugars, amino acids, peptides, opines, phosphate, sulfate, metals (Higgins 2001). Bacterial efflux systems include various drugs such as daunomycin and erythromycin and toxins like hemolysin and leukotoxin. Proteases and complex carbohydrates like capsular polysaccharides are also transported by ABC transporters (Tam and Saier, Jr. 1993). In humans ABC transporters are often associated with genetic diseases like Cystic fibrosis and Tangiers disease (Higgins 1992). When comparing ABC transporters in various microbial genomes it is interesting to note the difference of transporter number in parasitic bacteria as opposed to those that exist in more diverse environments. *E. coli* and *Haemophilus influenzae* contain similar sets of transporters whereas *Mycoplasma genitalium*, *M. pneumoniae* and *Helicobacter pylori* lack many transporters (branched-chain amino acid, many simple sugar and peptide transporters) which is consistent with the observation that these microorganisms are partially or entirely lacking the biosynthetic pathways for amino acids (Tomii and Kanehisa 1998).

Analysis of the *E. coli* genome sequence revealed that approximately 5% of the genome was made up of ABC transport genes and this was thought to be a very high percentage compared with other genomes that had been sequenced at that time (Linton and Higgins 1998). When the *B. subtilis* genome annotation was analysed for the presence of ABC transport systems it was also found that 5% of the genome annotation was composed of ABC transporters with half of those hypothesized to be involved with

uptake (Quentin et al. 1999). Table 1-1 compares the number of ABC transport proteins present in seven bacteria.

Table 1-1. Number of transport proteins and ABC transport proteins in a variety a bacteria (www.membranetransport.org).

	Sm	*Ml*	*At*	*Pp*	*Ec*	*Mt*	*Hi*
Genome Size(Mb)	6.7	7.59	5.3	6.1	4.6	4.4	1.83
Total Transporter Proteins	382	433	357	380	351	148	137
No. of ABC family transporters	200	216	219	117	67	44	67

Sm = S. meliloti 1021
Ml = Mesorhizobium loti MAFF303099
At = A. tumefaciens C58
Pp = Pseudomonas putida KT2440
Ec = E. coli K12
Mt = Mycobacterium tuberculosis H37Rv
Hi = H. influenzae KW20

In *S. meliloti* 12% of the genome has been annotated as ABC transporter genes and on pSymB alone they account for a striking 17% of this megaplasmids' annotation. The diversity of carbon and nitrogen sources that can be utilized by this organism may reflect the large number of annotated ABC transporters. This potential for the transport and metabolism of such a wide diversity of substrates and molecules is definitely an interesting aspect of *S. meliloti's* ability to compete and thrive in a multitude of environments (Finan et al. 2001).

Transport in *S. meliloti*

A small number of ABC transporters in *S. meliloti* have been characterized. By screening transposon mutants on various carbon sources Lambert *et al.*, isolated a mutant that could not used fructose as the sole carbon source. Sequence analysis demonstrated 2 operons divergently transcribed with genes displaying similarity to the ribose transport

genes in *B. subtilis* and *R.. capsulatus.* Transport assays verified that the cluster was involved in fructose uptake as well as that of mannose and ribose. *frcBCA* was determined to be the sole fructose transporter in *S. meliloti* due to the abolishment of growth on this substrate which was interesting because there is often more that one transporter for a particular substrate. Although *frcBCA* is involved with mannose and ribose, transport mutants of the *frc* cluster did not display a phenotype when grown on these substrates which would indicate multiple transporters for these compounds (Lambert et al. 2001).

The trehalose transport and utilization genes (*thuEFGK*) are another set of genes that make up an ABC transport system. Trehalose has been shown to be an osmoprotectant in many organisms including *S. meliloti*. *thuE* mutants were unable to grow on trehalose and had impaired growth on maltose, although *thuEFGK* are induced only by trehalose. *aglEFGAK* is the ABC transport system for the alpha-glucosides sucrose, maltose and trehalose, but a mutant in this transporter can still grow using all these carbon sources as the sole carbon source. When a double mutant in the *thu* and *agl* loci was tested in plants it was still able to nodulate and fix nitrogen in alfalfa but it was impaired in its ability to colonize alfalfa. These findings again demonstrated that there can be a certain amount of redundancy in the transporters of *S. meliloti* and that the uptake of disaccharides (especially those found in alfalfa root exudates) can affect the competitiveness in the rhizosphere (Jensen et al. 2002; Willis and Walker 1999)

ß-glucuronidase

Reporter genes have been used for the past forty years to monitor gene expression. One of the earliest reports in which gene fusions were employed is that of Miller, where β-galactosidase was used as an assayable gene expression reporter (Miller 1972). When used as a transcriptional fusion reporter, enzymes may be used to quantify gene expression or as a tool to mark strains in the environment. Frequently used reporters for the genetic study of microorganisms are *lacZ*, *luxAB*, *gfp* and *gusA*, encoding the proteins ß-galactosidase, bacterial luciferase, green fluorescent protein and ß-glucuronidase (Meighen 1993). The green fluorescent protein (Gfp) from *Aequorea Victoria* is a more recent development in reporting systems and has an advantage over other reporter enzymes which need more complicated assays for their detection. *gfp* needs only the correct excitation wavelength to be detected and needs no substrate to act upon as do many of the other reporters. This is a newer reporting system and is not yet accepted as a quantitative reporter of gene expression although is very useful for *in-situ* localization of proteins (Tsien 1998). *lacZ* has been commonly used in bacteria for its reproducibility and sensitivity as a reporter of gene expression, although the timed assay requires a substrate and specific conditions to optimize ß-galactosidase activity. The other limitation of a *lacZ* reporting system is that most bacterial genomes encode a gene for ß-galactosidase. Hence unless a *lacZ* mutant strain is used as a backbone, there may be a background of enzyme activity.

ß-glucuronidase catalyses the hydrolysis of a wide variety of glucuronides, which are compounds formed by the condensation of a glucuronic acid with a variety of aromatic hydroxyl compounds including borneol, sterols and phenol (Fruton JS and

Simmonds S 1959). This enzyme is mainly found in vertebrates and is absent from plants, most bacteria, fungi and insects. This has made *gusA* an excellent reporter because very small quantities of ß-glucuronidase can be measured accurately due to no background activity for this enzyme in most bacteria and plants (Jefferson et al. 1987a). One of the bacteria containing ß-glucuronidase (*gusA*) is *E. coli* which is the donor of the gene used in bacterial reporting systems. There are many commercially available substrates for *gusA* including 5-bromo-4-chloro-3-indoyl β-D-glucuronide (X-Gluc) which is used for histochemical assays and ρ-nitrophenyl β-glucuronide (PNPG) for spectrophotometric analyses. The *gusA* gene product of *E. coli* is a homotetramer with each subunit having a molecular mass of 68000 daltons. ß-glucuronidase is very stable, has a wide optimal pH range (pH 5.0 – 7.5) and is not thermally inactivated at 50°C, making it a great candidate for a reporter enzyme (Jefferson et al. 1986). ß-glucuronidase was first used as a reporter in plants to localize certain proteins within the cell or cells within tissues (Jefferson et al. 1987b). Subsequently *gusA* was used to monitor gene expression of *S. meliloti* symbiotic genes within the nodule using transposon Tn5-*gusA* to gain knowledge on the expression patterns within alfalfa (Sharma and Signer 1990). This demonstrated the usefulness of *gusA* not only in plant studies but also for plant-microbe interactions due to the absence of ß-glucuronidase in both *S. meliloti* and alfalfa. In an additional report, an *in vitro* ß-glucuronidase assay was developed in order to study the effects of various stresses on the expression of a transcriptional regulator, *phrR*. This study showed that *phrR* was induced by exposure to low pH and other stresses and also demonstrated the ability of *gusA* to report gene expression in vitro just as one would with *lacZ* (Reeve et al. 1998).

This Work

The purpose of this work was to generate expression data on the ABC transporters of pSymB in order to infer the substrate specificity of the various transporters. ß-glucuronidase (GusA) was used as a reporter of gene expression under the control of the native promoter. A suicide, GusA reporter vector, pTH1360 was used to clone PCR amplified 3' ends of ACB transport operons. Through a single crossover event, a single copy of *gusA* was inserted under control of the native promoter with no disruption to the gene function of the operon.

A high-throughput ß-glucuronidase assay was developed in order to assay the transport genes after growth in a variety of substrates. Many inducers of ABC transporters were found including arabinose, galactose, galactitol, protocatechuate, p-hydroxybenzoate, hydroxyproline and various root and legume seed exudates. The reporter fusions were also inoculated onto alfalfa plants and nodule extracts assayed for GusA activity to determine if the transporters were induced during symbiosis.

Two substrates, galactitol and hydroxyproline found to induce ABC transport systems were studied in depth in order to more fully understand the transport, metabolism and regulation of these compounds.

CHAPTER 2. MATERIALS AND METHODS

Bacterial strains and growth conditions

S. meliloti was grown at 30°C and *E. coli* was grown at 37°C. Cultures were inoculated with single colonies that had been streak purified three times on selective media. Small scale cultures (2 ml) used for genetic experiments and for plasmid DNA isolation were grown in test tubes on a rotary mixer. Larger cultures were grown in Erlenmeyer flasks in a rotary shaker. *S. meliloti* and *E. coli* were routinely grown in Luria-Bertani broth (LB) which contains 10 g tryptone (Difco), 5 g Yeast extract (Difco), and 5 g NaCl per litre of double distilled water. For growth of *S. meliloti* LB broth was supplemented with $MgSO_4$ (2.5 mM) and $CaCl_2$ (2.5 mM). Solid media was prepared by the addition of 15 g agar (Difco) to 1L of LB before sterilization. Defined growth medium was M9 minimal media. This media contains 5 x M9 salts (Difco) which consist of Na_2HPO_4 (33.9 g/L), KH_2PO_4 (15 g/L), NaCl (2.5 g/L) and NH_4Cl (5 g/L), a carbon source, 0.3 mg/ml biotin, 10 ng/ml cobalt chloride and 1.0 mM $MgSO_4$ and 0.25 $CaCl_2$. Media was sterilized at 15 pounds/square inch at 121°C for 30 minutes. Temperature labile compounds (sugars, amino acids and plant exudates) were filter sterilized through a 0.45 or 0.20 μM filter. Carbon sources were filter sterilized and used at concentrations of 15 mM (succinate and glucose), 10 mM (galactitol), 5 mM (all amino acids including hydroxyproline and allohydroxyproline), 1% casein amino acids, 1% glycerol and 0.2% (all others mentioned). When necessary, thiamine was added to a final concentration of 5 μM.

Antibiotics were obtained from Sigma or Boehringer Mannhiem and were stored

at –20°C as stock solutions in ethanol (tetracycline, chloramphenicol), 50% methanol (rifampicin) and the remaining in water. They were filter sterilized and used at the following concentrations for the growth of *E. coli* on solid agar media (µg/ml): kanamycin sulphate (Km), 20; chloramphenicol (Cm), 10; spectimomycin dihydrochloride (Sp), 100; tetracycline (Tc), 10; gentamicin sulphate (Gm), 10. For *E. coli* grown in liquid media the indicated concentrations were halved. For growth of *S. meliloti* on solid agar media the following antibiotic concentration were used (µg/ml): streptomycin sulphate (Sm), 200; neomycin sulphate (Nm), 200; spectimomycin dihydrochloride (Sp), 200; tetracycline (Tc), 5; gentamicin sulphate (Gm), 60; and rifampicin (Rm), 20. For *S. meliloti* growth in broth antibiotic concentration used were half those for growth on agar. All plasmids, strains and primers used in this study are listed in Tables 2-1 through 2-3.

Table 2-1. Plasmids used in this study

Plasmid	Relevant Characteristics	Reference	Primer sets (5' – 3')	Stock Number
pRK600	pRK2013 *npt*::Tn9, Cm^R	(Finan et al. 1986)		
pHP45Ω	pBR322 derivative carrying the ΩSm-Sp^R intersposon	(Prentki and Krisch 1984b)		
pUCP3OT	suicide cloning vector, Gm^R	(Schweizer HP et al. 1996)		
pTH1582	GusA from pFUS1 (*pstI* sites) into pTH1581 (modified pJP2), Tc^R	Z. Yuan, Finan lab (Prell et al. 2002)		M462
pTH1360*	pV0155 with *gusA* cassette from pFUS1	R. Zaheer, Finan lab		M216
pTH1639	309 bp 3' end *hutU* (PmlI-XbaI into pTH1360)	This study	ML1823 – ML1829	M527
pTH1640	421 bp 3' end SMb21373 (PmlI-XbaI into pTH1360)	This study	ML1822 – ML1828	M528
pTH1641	314 bp 3' end SMb20784 (PmlI-XbaI into pTH1360)	This study	ML1824 – ML1830	M529
pTH1642	425 bp 3' end SMb21097 (PmlI-XbaI into pTH1360)	This study	ML1825 – ML1831	M530
pTH1643	392 bp 3' end SMb20030 (PmlI-XbaI into pTH1360)	This study	ML1821 – ML1827	M531
pTH1644	385 bp 3' end SMb20124 (PmlI-XbaI into pTH1360)	This study	ML1826 – ML1832	M532
pTH1645	354 bp 3' end SMb21707 (PmlI-XbaI into pTH1360)	This study	ML2463 – ML2464	M533
pTH1646	362 bp 3' end SMb21138 (HindIII-XbaI into pTH1360)	This study	ML2461 – ML2462	M534
pTH1647	389 bp 3' end SMb21216 (HindIII-XbaI into pTH1360)	This study	ML2465 – ML2466	M535
pTH1648	450 bp 3' end SMb21375 (HindIII-XbaI into pTH1360)	This study	ML2467 – ML2468	M536
pTH1649	479 bp 3' end SMb20235 (HindIII-XbaI into pTH1360)	This study	ML2469 – ML2470	M537
pTH1650	435 bp 3' end SMb20854 (HindIII-XbaI into pTH1360)	This study	ML2646 – ML2647	M538
pTH1651	453 bp 3' end SMb20979 (HindIII-XbaI into pTH1360)	This study	ML2471 – ML2472	M539
pTH1652	470 bp 3' end SMb20318 (HindIII-XbaI into pTH1360)	This study	ML2473 – ML2474	M540

pTH1653	335 bp 3' end SMb20488 (SalI-XbaI into pTH1360)	This study	ML2644 – ML2645	M541
pTH1654	336 bp 3' end SMb20506 (HindIII -XbaI into pTH1360)	This study	ML2642 – ML2643	M542
pTH1655	425 bp 3' end SMb21016 (HindIII -XbaI into pTH1360)	This study	ML2475 – ML2476	M543
pTH1656	526 bp 3' end SMb21342 (HindIII -XbaI into pTH1360)	This study	ML2572 –ML2573	M544
pTH1657	483 bp 3' end SMb21592 (HindIII -XbaI into pTH1360)	This study	ML2574 – ML2575	M545
pTH1658	336 bp 3' end SMb21602 (HindIII -XbaI into pTH1360)	This study	ML2576 – ML2577	M546
pTH1659	497 bp 3' end SMb20904 (HindIII -XbaI into pTH1360)	This study	ML2648 – ML2649	M547
pTH1660	486 bp 3' end SMb21424 (HindIII -XbaI into pTH1360)	This study	ML2650 – ML2651	M548
pTH1661	483 bp 3' end SMb21458 (HindIII -XbaI into pTH1360)	This study	ML2652 – ML2653	M549
pTH1662	427 bp 3' end SMb20720 (HindIII -XbaI into pTH1360)	This study	ML2656 – ML2657	M550
pTH1663	357 bp 3' end SMb20697 (PmlI-XbaI into pTH1360)	This study	20697F – 20697R	M551
pTH1664	392 bp 3' end SMb20018 (PmlI-XbaI into pTH1360)	This study	20018F – 20018R	M552
pTH1665	395 bp 3' end SMb20112 (PmlI-XbaI into pTH1360)	This study	20112F – 20112R	M553
pTH1666	356 bp 3' end SMb20568 (PmlI-XbaI into pTH1360)	This study	20568F – 20568R	M554
pTH1667	309 bp 3' end SMb20571 (PmlI-XbaI into pTH1360)	This study	20571F – 20571R	M555
pTH1668	394 bp 3' end SMb20263 (PmlI-XbaI into pTH1360)	This study	20263F – 20263R	M556
pTH1669	462 bp 3' end SMb02793 (PmlI-XbaI into pTH1360)	This study	02793F – 02793R	M557
pTH1670	353 bp 3' end SMb20320 (PmlI-XbaI into pTH1360)	This study	20320F – 20320R	M558
pTH1671	424 bp 3' end SMb20036 (PmlI-XbaI into pTH1360)	This study	20036F – 20036R	M559
pTH1672	533 bp 3' end SMb20354 (PmlI-XbaI into pTH1360)	This study	20354F – 20354R	M560
pTH1673	411 bp 3' end SMb20349 (PmlI-XbaI into pTH1360)	This study	20349F – 20349R	M561
pTH1674	525 bp 3' end SMb21151 (PmlI-XbaI into pTH1360)	This study	21151F – 21151R	M562
pTH1675	437 bp 3' end SMb21587 (PmlI-XbaI into pTH1360)	This study	21587F – 21587R	M563
pTH1676	435 bp 3' end SMb20931 (PmlI-XbaI into pTH1360)	This study	20931F – 20631R	M564
pTH1677	336 bp 3' end SMb20158 (PmlI-XbaI into pTH1360)	This study	20158F – 20158R	M565
pTH1678	399 bp 3' end SMb20433 (PmlI-XbaI into pTH1360)	This study	20433F – 20433R	M566
pTH1679	317 bp 3' end SMb20428 (PmlI-XbaI into pTH1360)	This study	20428F – 20428R	M567
pTH1680	413 bp 3' end SMb20442 (PmlI-XbaI into pTH1360)	This study	20442F – 20442R	M568
pTH1681	366 bp 3' end SMc00954 (PmlI-XbaI into pTH1360)	This study	ML1823 – ML1829	M569
pTH1682	324 bp 3' end SMb20502 (PmlI-XbaI into pTH1360)	This study	20502F – 20502R	M570
pTH1683	420 bp 3' end SMb20027 (PmlI-XbaI into pTH1360)	This study	20027F – 20027R	M571

pTH1684	414 bp 3' end SMb20476 (PmlI-XbaI into pTH1360)	This study	20476F – 20476R	M572
pTH1685	322 bp 3' end SMb21130 (PmlI-XbaI into pTH1360)	This study	21130F – 21130R	M573
pTH1686	387 bp 5' end SMb21377 (PmlI-XbaI into pTH1360)	This study	ML3145 – ML3146	M653
pTH1687	430 bp 5' end SMb21094 (PmlI-XbaI into pTH1360)	This study	ML3143 – ML3144	M654
pTH1688	338 bp 5' end SMb21345 (PmlI-XbaI into pTH1360)	This study	ML3925 – ML3926	M655
pTH1689	564 bp 5' end SMb20902 (PmlI-XbaI into pTH1360)	This study	ML3995 – ML3996	M656
pTH1690	465 bp 5' end SMb20263 (PmlI-XbaI into pTH1360)	This study	20263KOF – 20263KOR	M657
pTH1691	348 bp 5' end SMb20322 (PmlI-XbaI into pTH1360)	This study	20322KOF – 20322KOR	M658
pTH1798	431 bp 5' end SMb20261 (PmlI-XbaI into pTH1360)	This study	20261F – 20261R	M659
pTH1799	510 bp 5' end SMb20262 (PmlI-XbaI into pTH1360)	This study	20262F – 20262R	M660
pTH1800	292 bp of intergenic region b/w SMb20262 and SMb20263 (BglII-HindIII into pTH1582)	This study	HLPINF – HLPINR	M661
pTH1801	330 bp of intergenic region b/w SMb20261 and SMb20262 (BglII - HindIII into pTH1582)	This study	PUINFF – PUINFR	M662
pTH1802	340 bp of intergenic region b/w SMb20262 and SMb20261 (BglII - HindIII into pTH1582)	This study	PUINRF - PUINRR	M663
pTH1803	288 bp of intergenic region b/w SMb21378 and SMb21377 (BglII - HindIII into pTH1582)	This study	ML6989 – ML6990	M664
pTH1804	922 bp SMb21372 (nucleotides 10 - 931) (HindIII – BamHI into pUCP3OT)	This study	ML7348 – ML7349	M665
pTH1805	PCR amplified ΩSp-Sm from pHP45 cloned into pTH1804 into SalI sites at 466 and 535 bp of SMb21372	This study	ML7350 – ML7350	M666
pTH1811	510 bp 5' end SMb20262 (SmaI – XbaI into pUCP30T)	This study	20262F – 20262R	M674

* See Figure 3-5 for a plasmid map of pTH1360

Table 2-2. *E. coli* Strains used in this study

Strain	Relevant Characteristics	Reference
DH5α	F^-, *end*A1, *hsd*R17 (r_K^-, m_K^-), *sup*E44, *thi*-1, *rec*A1, *gyr*A96, *rel*A1, Δ(*arg*F-*lac*ZYA) U169, Φ80d*lacZ*, ΔM15	B.R.L. Inc
MT616	MT607/pRK2013 *npt*::Tn9	(Finan et al. 1986)
J299	pHP45Ω	(Prentki and Krisch 1984c)
J252	pUCP3OT	(Schweizer HP et al. 1996)
M462	pTH1582	Z. Yuan, Finan lab
M216	pTH1360	R Zaheer, Finan lab
M527	DH5α (pTH1639)	This study
M528	DH5α (pTH1640)	This study
M529	DH5α (pTH1641)	This study
M530	DH5α (pTH1642)	This study
M531	DH5α (pTH1643)	This study
M532	DH5α (pTH1644)	This study
M533	DH5α (pTH1645)	This study
M534	DH5α (pTH1646)	This study
M535	DH5α (pTH16470	This study
M536	DH5α (pTH1648)	This study
M537	DH5α (pTH1649)	This study
M538	DH5α (pTH1650)	This study
M539	DH5α (pTH1651)	This study
M540	DH5α (pTH1652)	This study
M541	DH5α (pTH1653)	This study
M542	DH5α (pTH1654)	This study
M543	DH5α (pTH1655)	This study
M544	DH5α (pTH1656)	This study
M545	DH5α (pTH1657)	This study
M546	DH5α (pTH1658)	This study

M547	DH5α (pTH1659)	This study
M548	DH5α (pTH1660)	This study
M549	DH5α (pTH1661)	This study
M550	DH5α (pTH1662)	This study
M551	DH5α (pTH1663)	This study
M552	DH5α (pTH1664)	This study
M553	DH5α (pTH1665)	This study
M554	DH5α (pTH1666)	This study
M555	DH5α (pTH1667)	This study
M556	DH5α (pTH1668)	This study
M557	DH5α (pTH1669)	This study
M558	DH5α (pTH1670)	This study
M559	DH5α (pTH1671)	This study
M560	DH5α (pTH1672)	This study
M561	DH5α (pTH1673)	This study
M562	DH5α (pTH1674)	This study
M563	DH5α (pTH1675)	This study
M564	DH5α (pTH1676)	This study
M565	DH5α (pTH1677)	This study
M566	DH5α (pTH1678)	This study
M567	DH5α (pTH1679)	This study
M568	DH5α (pTH1680)	This study
M569	DH5α (pTH1681)	This study
M570	DH5α (pTH1682)	This study
M571	DH5α (pTH1683)	This study
M572	DH5α (pTH1684)	This study
M573	DH5α (pTH1685)	This study
M653	DH5α (pTH1686)	This study
M654	DH5α (pTH1687)	This study
M655	DH5α (pTH1688)	This study
M656	DH5α (pTH1689)	This study

M657	DH5α (pTH1690)	This study
M658	DH5α (pTH1691)	This study
M659	DH5α (pTH1798)	This study
M660	DH5α (pTH1799)	This study
M661	DH5α (pTH1800)	This study
M662	DH5α (pTH1801)	This study
M663	DH5α (pTH1802)	This study
M664	DH5α (pTH1803)	This study
M665	DH5α (pTH1804)	This study
M666	DH5α (pTH1805)	This study

Table 2-3. *S. meliloti* strains used in this study

Strain	Relevant Characteristics	Reference
Rm1021	SU47, *str*-21	(Meade et al. 1982)
RmK763	Rm5000 Rif^R isolate of SU47	(Finan et al. 1984)
RmF909	ΔΩ5085-5047::TnV	(Charles and Finan 1991)
RmF117	ΔΩ5060-5033::Tn5-233	(Charles and Finan 1991)
RmG462	ΔΩ5011-5033::Tn5-*oriT*	(Charles and Finan 1991)
SRmG373	ΔΩ5177-5079::Tn5-*oriT*	(Charles and Finan 1991)
RmG506	ΔΩ5040::Tn5-233 ΔG506	(Charles and Finan 1991)
RmG470	ΔΩ5025-5007::Tn5	(Charles and Finan 1991)
RmP32	Sm1021 ΦpTH1639, *dme*::*gusA* in pTH1360	R. Zaheer, Finan lab
RmK1001	Sm1021 (pTH1584), *katA*::gusA, *phoC*-	Z. Yuan, Finan lab
RmP62	Sm1021 (pTH1582), empty vector	Z. Yuan, Finan lab
SmP319	nifH::gusA-rfp in pTH1522	J. Cheng, Finan lab
SmRL43	Library clone #43 containing 1920 bp from SMb20268 and SMb20269, P110 background	Finan Lab unpublished
RmP186	Sm1021 ΦpTH1639	This study
RmP187	Sm1021 ΦpTH1640	This study
RmP188	Sm1021 ΦpTH1641	This study
RmP189	Sm1021 ΦpTH1642	This study
RmP190	Sm1021 ΦpTH1643	This study
RmP191	Sm1021 ΦpTH1644	This study
RmP192	Sm1021 ΦpTH1645	This study
RmP193	Sm1021 ΦpTH1646	This study
RmP194	Sm1021 ΦpTH1647	This study
RmP195	Sm1021 ΦpTH1648	This study
RmP196	Sm1021 ΦpTH1649	This study
RmP197	Sm1021 ΦpTH1650	This study
RmP198	Sm1021 ΦpTH1651	This study

RmP199	Sm1021 ΦpTH1652	This study
RmP200	Sm1021 ΦpTH1653	This study
RmP201	Sm1021 ΦpTH1654	This study
RmP202	Sm1021 ΦpTH1655	This study
RmP203	Sm1021 ΦpTH1656	This study
RmP204	Sm1021 ΦpTH1657	This study
RmP205	Sm1021 ΦpTH1658	This study
RmP206	Sm1021 ΦpTH1659	This study
RmP207	Sm1021 ΦpTH1660	This study
RmP208	Sm1021 ΦpTH1661	This study
RmP209	Sm1021 ΦpTH1662	This study
RmP210	Sm1021 ΦpTH1663	This study
RmP211	Sm1021 ΦpTH1664	This study
RmP212	Sm1021 ΦpTH1665	This study
RmP213	Sm1021 ΦpTH1666	This study
RmP214	Sm1021 ΦpTH1667	This study
RmP215	Sm1021 ΦpTH1668	This study
RmP216	Sm1021 ΦpTH1669	This study
RmP217	Sm1021 ΦpTH1670	This study
RmP218	Sm1021 ΦpTH1671	This study
RmP219	Sm1021 ΦpTH1672	This study
RmP220	Sm1021 ΦpTH1673	This study
RmP221	Sm1021 ΦpTH1674	This study
RmP222	Sm1021 ΦpTH1675	This study
RmP223	Sm1021 ΦpTH1676	This study
RmP224	Sm1021 ΦpTH1677	This study
RmP225	Sm1021 ΦpTH1678	This study
RmP226	Sm1021 ΦpTH1679	This study
RmP227	Sm1021 ΦpTH1680	This study
RmP228	Sm1021 ΦpTH1681	This study
RmP229	Sm1021 ΦpTH1682	This study

RmP230	Sm1021 ΦpTH1683	This study
RmP231	Sm1021 ΦpTH1684	This study
RmP232	Sm1021 ΦpTH1685	This study
RmP233	Sm1021 ΦpTH1686	This study
RmP234	Sm1021 ΦpTH1687	This study
RmP235	Sm1021 ΦpTH1688	This study
RmP236	Sm1021 ΦpTH1689	This study
RmP237	Sm1021 ΦpTH1690	This study
RmP238	Sm1021 ΦpTH1691	This study
RmP239	Sm1021 ΦpTH1798	This study
RmP240	Sm1021 ΦpTH1799	This study
RmP241	Sm1021 ΦpTH1800	This study
RmP242	Sm1021 ΦpTH1801	This study
RmP243	Sm1021 ΦpTH1802	This study
RmP244	Sm1021 ΦpTH1803	This study
RmP245	RmP233 ΦpTH1805 (pUCP30T recombined out to leave ΩSp-Sm in SMb21372)	This study
RmP249	Sm1021 pTH1811	This study
RmP311	SmP249 pTH1800	This study
RmP312	SmP249 pTH1801	This study
RmP313	SmP249 pTH1802	This study
RmP314	RmF932 pTH1800	This study
RmP315	RmF932 pTH1801	This study
RmP316	RmF932 pTH1802	This study

Table 2-4. Primers used in this study

5' Primer / Sense (5' – 3')	3' Primer / Antisense (5' – 3')	Restriction sites	Notes
ML1821 GCCTCGCTTTCACGTCGAGAAGG	ML1827 AA**TCTAGA**TCAGTGCCTCAGCTTG	N/A – XbaI	Amplify the 3' end (384 bp) of *sig* (SMb20030)
ML1822 GTGCTCAGCAAGACGAGCTTCGG	ML1828 AA**TCTAGA**TCAGCGCTTTGCCTGCTT GGC	N/A - XbaI	Amplify the 3' end (413 bp) of SMb21373
ML1823 CGATGAAGGACGGCTCGGATGC	ML1829 AA**TCTAGA**TCACTCCCCCAATATGCC CGG	N/A - XbaI	Amplify the 3' end (309 bp) of *hutU* (SMb21163)
ML1824 CGAGCAGCAGATGTTGTCTACCG C	ML1830 AA**TCTAGA**TCAATGCAGCCCGCCAA CCC	N/A - XbaI	Amplify the 3' end (306 bp) of SMb20784
ML1825 CGGCAAGATAGTCGGCATGAGCC	ML1831 AA**TCTAGA**TCACTTCACCGCTTCGCC ATTGACG	N/A – XbaI	Amplify the 3' end (417 bp) of SMb21097
ML1826 TGATCTCATGCGCCGCTTCAAGG	ML1832 AA**TCTAGA**TCAAGCTGCCGATAGCG GAG	N/A – XbaI	Amplify the 3' end (375 bp) of SMb20124
ML2461 CCC**AAGCTT**CAAGGTGAAGAAG ATGAGC	ML2462 GC**TCTAGA**TCAGTTGAACGTCGAGG TC	HindIII – XbaI	Amplify the 3' end (345 bp) of SMb21138
ML2463 CCC**AAGCTT**CGGTACTTCAGACG ATG	ML2464 GC**TCTAGA**TCAAACGGTTAGATGCC GG	HindIII – XbaI	Amplify the 3' end (337 bp) of SMb21707
ML2465 CCC**AAGCTT**CCGATGAACATTCT GGACG	ML2466 GC**TCTAGA**CTATTCCATTCTAAGCGA GCG	HindIII – XbaI	Amplify the 3' end (372 bp) of SMb21216

ML2467 CCC**AAGCTT**CTTCGTCATCAATC AGATCG	ML2468 AA**TCTAGA**TCAGCGCTTTGCCTGCTT GG	HindIII – XbaI	Amplify the 3' end (433 bp) of SMb21375
ML2469 CCC**AAGCTT**CAATGTCGAGCAG ATCG	ML2470 AA**TCTAGA**CTACGCACTTGCGGGCG	HindIII – XbaI	Amplify the 3' end (462 bp) of SMb20235
ML2471 CCC**AAGCTT**CCTTCGGTTTCTAT CTCG	ML2472 AA**TCTAGA**CTATTGGACCGCCGTCG	HindIII – XbaI	Amplify the 3' end (436 bp) of SMb20979
ML2473 CCC**AAGCTT**CCTATTTCCTCGTT CTGG	ML2474 AA**TCTAGA**TCAGCTTTTCGGCCTGAT GG	HindIII – XbaI	Amplify the 3' end (453 bp) of SMb20318
ML2475 CCC**AAGCTT**TATCCTGCAGACCT ATCC	ML2476 GC**TCTAGA**TCAGAAGTCGAAGTTGC C	HindIII – XbaI	Amplify the 3' end (408 bp) of SMb21016
ML2572 CCC**AAGCTT**GACTTCCTTGGCTT TCC	ML2573 AA**TCTAGA**TCAGACGTGTTCCTCGCC TG	HindIII – XbaI	Amplify the 3' end (509 bp) of SMb21342
ML2574 CC**AAGCTT**GCTTCGTTGCAGGTT TCATC	ML2575 AA**TCTAGA**TCAAGACGGGGAGAAGG TCG	HindIII – XbaI	Amplify the 3' end (466 bp) of SMb21592
ML2576 CCC**AAGCTT**CGATTTCCAGATCC TGTTC	ML2577 AA**TCTAGA**TCAGCCCTTGATGCCGG AC	HindIII – XbaI	Amplify the 3' end (318 bp) of SMb21602
ML2646 CC**AAGCTT**GGTGTCGTCCTCAAC TTC	ML2646 AA**TCTAGA**TCATCGCTTCAGAGAGC CGG	HindIII – XbaI	Amplify the 3' end (419 bp) of SMb20854
ML2648 CC**AAGCTT**AATCTCTATCTCGGC CGC	ML2649 AA**TCTAGA**TCACGCCATCTGCATGG CG	HindIII – XbaI	Amplify the 3' end (481 bp) of SMb20904

ML2650 CCC**AAGCTT**CGATGTCGTGATTTTCACC	ML2651 AA**TCTAGA**TTACTTCCCAGTGCCGACAG	HindIII – XbaI	Amplify the 3' end (468 bp) of SMb21424
ML2652 CC**AAGCTT**AATGTTCGAGGTGATCGGATG	ML2653 AA**TCTAGA**TCACCCCTTGACGGCGC	HindIII – XbaI	Amplify the 3' end (465 bp) of SMb21458
ML2654 CC**AAGCTT**GCTTTCACGAGTTCACCG	ML2655 AC**TCTAGA**TTAAAGGCCGTTGTCCTTCAG	HindIII – XbaI	Amplify the 3' end (424 bp) of SMb20671
ML2656 CC**AAGCTT**GCTCCAAGGTCGAAATCC	ML2657 AA**TCTAGA**TCAGAACGGCGAGTCGGG	HindIII – XbaI	Amplify the 3' end (409 bp) of SMb20720
ML2642 CC**AAGCTT**ATTTTCGCCCTCCAGGG	ML2643 AAA**TCTAGA**TCAGGCTCGCGGCCTGC	HindIII – XbaI	Amplify the 3' end (320 bp) of SMb20506
ML2644 ACG**GTCGAC**ATACCTATCACATGAACATCG	ML2645 AA**TCTAGA**TCAGATCAGGCCGTACTGCG	SalI – XbaI	Amplify the v3' end (319 bp) of SMb20488
ML3143 ATGACCGAGCCCACTCAGCTTTGG	ML3144 AA**TCTAGA**GCTTCTTCGATGGCGATCACG	N/A – XbaI	Amplify the 5' end (430bp) of SMb21094
ML3145 ATGAACATGCCCCGCATTCTCAAAACTC	ML3146 AA**TCTAGA**TGCCGATATGCGACGTCACCG	N/A – XbaI	Amplify the 5' end (387bp) of SMb21377
ML3925 GAACAGGATTTCACGTCGCGCG	ML3926 AA**TCTAGA**ATGACGGCGGGGATCAACG	N/A – XbaI	Amplify the 5' end (338 bp) of SMb21342 (nucleotides 1 – 338)

ML3995 GCCATCATAGCGTCCATGCATTCCG	ML3996 **AATCTAGA**TGGGCGTTCTCCGGCTTCC	N/A – XbaI	Amplify the 5' end (564bp) of SMb20902
20697F **AAGATATG**GGAAGGAAACGACCGCTCC	20697R **AATCTAGA**TCAGTCCCTGTGTCTGTCG	EcoRV – XbaI	Amplify the 3' end (357 bp) of SMb20697
20018F GACATCAACGGCGTGCTGGTGC	20018R **GCTCTAGA**TCACTTGAACAGCGCTTCCG	N/A – XbaI	Amplify the 3' end (392 bp) of SMb20697
20058F **AAGATATG**AGCGGCAAAGGCTGCACATTG	20058R **AATCTAGA**TCAGGCTCCGACGGCG	EcoRV – XbaI	Amplify the 3' end (373 bp) of SMb20058
20112F **AAGATATG**ATCGTCGCGTTGTTGATGAAGG	20012R **AATCTAGA**TATCCTCGCCTCTTCCCCTC	EcoRV – XbaI	Amplify the 3' end (395 bp) of SMb20112
20568F GCATGACTCACCGGAAAACAAGGC	20568R **AATCTAGA**TCATTTCACCGCGGCTAGACC	N/A – XbaI	Amplify the 3' end (356 bp) of SMb20568
20571F **AAGATATG**TATGGCGCGTCGCGTCTC	20571R **AATCTAGA**CTAGTTCGAGGCGTCGCTC	EcoRV – XbaI	Amplify the 3' end (309 bp) of SMb20571
20263F **AAGATATG**CTATGAGGCGATCGCACTCG	20263R **AATCTAGA**TCAACGGTAAACGCCGTTCACC	EcoRV – XbaI	Amplify the 3' end (394 bp) of SMb20263
02793F **GAGATATG**ATCGAGGCGTTGAACTTCACG	02793R **AATCTAGA**CTATCTAAGCCGCGGCCG	EcoRV – XbaI	Amplify the 3' end (462 bp) of SMc02793
20036F **AAGATATG**TCGAGACGGGAACGGTCG	20036R **AATCTAGA**TCATTGGCCGCGGGCG	EcoRV – XbaI	Amplify the 3' end (406 bp) of SMb20036

20320F **AAGATATG**CCATGTCGAACGGC CAGG	20320R **AATCTAGA**TCAGAGCGTTGCGCCGA C	EcoRV – XbaI	Amplify the 3' end (353 bp) of SMb20320
20354F **AAGATATG**CAGTGCGACGAAGG ACAATCC	20354R **AATCTAGA**TCACAGGATGTCGATGC GGG	EcoRV – XbaI	Amplify the 3' end (517 bp) of SMb20354
20349F **AAGATATG**AAGATGGTGGCGCA GCAGTC	20349R **AATCTAGA**TCAATCGGCCAGCGCGA AG	EcoRV – XbaI	Amplify the 3' end (395 bp) of SMb20349
20059F **AAGATATG**AGACCACCGCATCA TAGGTCTCG	20059R **GCTCTAGA**ATGGACATGCGCAACCA CG	EcoRV – XbaI	Amplify the 3' end (360 bp) of SMb20059
21153F GC**GATATG**CTTCCAGGACGATGT TCCAG	21153R GC**TCTAGA**TCAATCTGCGATGCGTTT CACC	EcoRV – XbaI	Amplify the 3' end (327 bp) of SMb21153
21151F GC**GATATG**GCTTACGAGCAGAA CGAAATCC	21151R GC**TCTAGA**TTGTACTGTTCGAGGTCG GC	EcoRV – XbaI	Amplify the 3' end (501 bp) of SMb21151
21587F GC**GATATG**ATCTGAAGATCGTCC GCAGC	21587R GC**TCTAGA**TCAGTAGCCGAGGCCTTT CT	EcoRV – XbaI	Amplify the 3' end (421 bp) of SMb21587
20931F **AAGATATG**GAAGGCAACTGGAC GGAAGTGG	20931R GC**TCTAGA**TTATTTCTGGATGCAGGT ATCCGC	EcoRV – XbaI	Amplify the 3' end (419 bp) of SMb20931
20144F **AAGATATG**GACAAGAATGCCTG CGAGTCG	20144R GC**TCTAGA**ATGCAAATGAAAAGAAC ACTGATTATGGG	EcoRV – XbaI	Amplify the 3' end (374 bp) of SMb20144
20140F **AAGATATG**TGCCGGGCTTCGAA GTCTATTCC	20140R **AATCTAGA**CACAGCAGCCGGTGCC	EcoRV – XbaI	Amplify the 3' end (346 bp) of SMb20140

20158F AA**GATATG**TACAAGCTCGGCATCTGCAAGG	20158R GC**TCTAGA**TCACTTGCTGTAATTGGCGC	EcoRV – XbaI	Amplify the 3' end (321 bp) of SMb20185
20267F AA**GATATG**CGGACGCAGATCACTTTCG	20267R GC**TCTAGA**TCAGAATCTTTGCGGCG	EcoRV – XbaI	Amplify the 3' end (383 bp) of SMb20267
20433F AA**GATATG**AGAGCGGCCATCACTGG	20433R GC**TCTAGA**TCAACTTTCGAATGTCGTACC	EcoRV – XbaI	Amplify the 3' end (416 bp) of SMb20433
20428F AA**GATATG**CTGAAGATGCTGCAGGACG	20428R AA**TCTAGA**TTATTTCGCGGCGCAGAGC	EcoRV – XbaI	Amplify the 3' end (299 bp) of SMb20428
20442F AA**GATATG**GAGGTCTATCTCGCGCTC	20422R AA**TCTAGA**TCACTTGATGGCGCGGATAG	EcoRV – XbaI	Amplify the 3' end (395 bp) of SMb20422
20295F GC**GATATG**GACTCGATGAAGTTCACC	20295R AA**TCTAGA**TTAACGGAGATCGAGCAGCG	EcoRV – XbaI	Amplify the 3' end (330 bp) of SMb20295
20476F AA**GATATG**GCTGTCGGCTTCAAGGTC	20476R CG**TCTAGA**CTACTCTTTCGTCACGCC	EcoRV – XbaI	Amplify the 3' end (414 bp) of SMb20476
20502F GGTGACGGTGCTGTGTGCG	20502R GC**TCTAGA**TCAGAGAGCCTTCCGC	N/A– XbaI	Amplify the 3' end (316 bp) of SMb20502
21130F AA**GATATG**CGTGCCGGGAATGAGCC	21130R AA**TCTAGA**CACGTGCCCGTGGCTG	EcoRV – XbaI	Amplify the 3' end (304 bp) of SMb21130
20027F AA**GATATG**CTGGTGGTCCTGAACCTGC	20027R AA**TCTAGA**TCAGTCCTCCTCGACGAAGA	EcoRV – XbaI	Amplify the 3' end (402 bp) of SMb20027

20015F AA**GATATG**GCTCATGCTCAACCGCACC	20015R AA**TCTAGA**TCATGCGCGCATCGCCC	EcoRV – XbaI	Amplify the 3' end (418 bp) of SMb20015
20263KOF GC**GATATG**CGACGAGAACAACGAACC	20263KOR GC**TCTAGA**TTGACGTTGGACTGCGC	EcoRV – XbaI	Amplify the 5' end (449 bp) of SMb20263
20322F GC**GATATG**CCTGAACGTCGTCACC	20322R AA**TCTAGA**GTCTGGCTGCGCTCGG	EcoRV – XbaI	Amplify the 5' end (332 bp) of SMb20322
20261F AA**GATATG**GCTTCAGGAGCGCGTCG	20261R AA**TCTAGA**CGACCGTGGCATAGCTCG	EcoRV – XbaI	Amplify the 5' end (415 bp) of SMb20261
20262F GC**GATATG**TCTGGACCTATGGCTATTCC	20262R AA**TCTAGA**CCCTGGATCAGCGAGAATACG	EcoRV – XbaI	Amplify the 5' end (492 bp) of SMb20262
HLPINF CCC**AAGCTT**GTCGCTTACCAGAACATGC	HLPINR GA**AGATCT**AACGAGGCCGAAGGCGAGG	HindIII – BglII	Amplify the intergenic region between SMb20262 and SMb20263 (273 bp)
PUINFF CC**AAGCTT**AGCGCGGTGGTCGTCAGG	PUINFR GG**AGATCT**CCGCGACGAGGTGTTTTCC	HindIII – BglII	Amplify the intergenic region between SMb20261 and SMb20262 (313 bp)
PUINRF CCC**AAGCTT**CCGCGACGAGGTGTTTTCC	PUINRR GGA**AGATCT**CTCGACGCGCTCCTGAAGC	HindIII – BglII	Amplify the intergenic region between SMb20262 and SMb20261 (329 bp)

ML6989 CCC**AAGCTT**TGGTTGCCTCCCCTTCC	ML6990 GA**AGATCT**CAGCAATGTGAGCGCGGTGC	HindIII – BglII	Amplify the intergenic region between SMb21378 and SMb21377 (267bp)
ML7348 CC**AAGCTT**GACGCTTCGACCGTGACG	ML7349 CGC**GGATCC**CGTTGTAGTTCTGCGC	HindIII – BamHI	Amplify SMb21372 (nucleotide 10 – 931)
ML7350 GC**GTCGAC**GGATCCGGTGATTGATTGAGC	ML7350 GC**GTCGAC**GGATCCGGTGATTGATTGAGC	SalI – SalI	AmplifyΩSpSm antibiotic resistance cassette from pHP45 (nucleotide 2332 – 4393)
ML758 TCAAGCTTGCATGCCTGC			Sequencing primer for pTH1360

β-glucuronidase assay

The β-glucuronidase assay was carried out as first described by Jefferson (Jefferson et al. 1986) and as adapted to *S. meliloti* by Reeve et al. (Reeve et al. 1998). Cultures for the β-Glucuronidase assay were grown overnight in 2 ml LBmc in test tubes (17 mm x 150 mm). Bacterial pellets were washed once with 0.85% NaCl and resuspended in 0.85% NaCl and subcultured into 2 ml of test media (M9 minimal media with various carbon sources) using a 1:50 or 1:100 dilution. Cultures were grown for 24 – 48 hours until an OD_{600} of approximately 0.3 to 0.6. Cultures (1.5 ml) were pelleted in microfuge tubes, washed with 0.85% saline and resuspended in GUS buffer (50 mM sodium phosphate, 50 mM DTT and 1 mM EDTA; pH 7). The OD_{600} of these cultures was measured for cell density either in 96 well microtitre plates or 1.5 ml cuvettes (1 cm cuvettes). The reaction was carried out in triplicate in microfuge tubes consisting of 790 μl of GUS buffer and 200 μl of the cultures resuspended in GUS buffer. One drop of toluene was added to each sample, vortexed and incubated with the lid off at 37°C for half an hour. The samples were equilibrated to room temperature (20°C) for five minutes before the addition of 10 μl of 35 mg/ml p-nitrophenyl β-glucuronide (PNPG). The start time of each reaction was recorded and the reactions were allowed to proceed for 20 minutes to an hour or until sufficient yellow product was observed. Reactions were stopped by removing 200 μl of the reaction into 700 μl of 0.46 M Na_2CO_3 with the stop time being noted. The stopped reaction mixture was centrifuged for 1 minute at 13.2 x 1000 rpm to sediment debris, then aliquoted into 96 well microtitre plates or 1.5 ml cuvettes for the OD_{405} reading. The absorbances were read in the Cary Varian

spectrophotometer or the Tecan Safire microtitre plate reader. Miller units were calculated to determine amount of β-Glucuronidase activity by using the following formula where the various dilutions are taken into account in the volume of culture used (Miller 1972):

$(1000 \times A_{405}) / (A_{600} \times$ reaction time in minutes $\times$ volume of culture used in ml).

High-throughput β-glucuronidase assay

The fusion strains and controls were grown in 1 ml of LBmc in deep well microtitre plates for 24 hours at 30°C (to an OD_{600} of approximately 1.0) in a rotary shaker. The strains were then subcultured into microtitre plates containing various test media using a replicating tool which inoculates approximately 5 µl of culture. Cultures were grown in test media for approximately 24 hours at 30°C. The MultiProbeII (PerkinElmer) was used to dispense reagents into the 96 well plates in order to achieve a more efficient screen. The actual reaction was as follows: 20 µl cell culture was added to a mixture containing: 78 µl GUS buffer, 1 µl p-nitrophenyl-ß-D-Glucuronide (PNPG) (35mg/ml) and 1 µl of 1% SDS (final concentration of 0.01%). The reaction was allowed to take place for approximately one hour or until sufficient yellow color was observed with the start and stop times for each row noted. The reaction was stopped by adding an equal volume (100 µl) of 1M Na_2CO_3. The original microtitre dish containing the cell cultures was used to generate an OD_{600} reading for cell density then the reaction plate was read at 405 nm using the Tecan Safire microtitre plate reader so that Miller Units could be calculated using the following formula:

(1000 x A_{405}) / (A_{600} x reaction time in minutes x volume of culture used in ml).

β-Galactosidase assay

Assays were adapted from the protocol described by Miller (Miller 1972) and carried out in microfuge tubes. The reaction was as follows: 950 μl of Z buffer (60 mM anhydrous Na2HPO4, 40 mM NaH2PO4, 10 mM KCl, 1.0 mM MgSO4, pH7), 50 μl of chloroform, 25 μl of 0.1% SDS and 50 μl of cells. This mixture was vortexed and the substrate o-nitrophenyl β-D-galactopyranoside (ONPG) was added. The reaction was allowed to proceed until sufficient yellow colour developed then was stopped by the addition of 0.5 ml Na_2CO_3 (1 M). The OD_{420} of the stopped reaction and the OD_{600} of the original culture were read on the Cary Varian spectrophotometer in 1.5 ml cuvettes (1.5 cm path length) in order to calculate Miller units by the following equation:

(1000 x A_{420}) / (A_{600} x reaction time in minutes x volume of culture used in ml).

Plant Growth Conditions

Alfalfa (*Medicago sativa*) seeds were surface sterilized in 95% ethanol for five minutes, in 2.5% sodium hypochlorite for 20 minutes and then rinsed repeatedly for 1 hour with sterile ddH_20. Seeds were spread on water agar plates to allow for germination for 2 days in the dark. Seeds were planted in Leonard assemblies which are a plastic pot with a hole in the bottom placed in a 250 ml glass beaker. A cotton wick extended from the hole in the plastic pot to the bottom of the beaker. Square plastic containers that stack together were also used in the same way with a hole in the top container and the cotton

wick extending from the top container to the bottom of the lower container. The top container was filled with a sand/vermiculite mixture (1:1 w/w) and 250 ml of Jensen's medium. The Leonard jars were covered on top and at the top-bottom junction with tinfoil and autoclaved for 1 hour on the wet cycle.

The Jensen's medium was prepared as a 2X solution, NaOH added to achieve a pH of 7 and diluted to 1 X. Jensen's medium is made up of: 1g $CaHPO_4$, 0.2 g K_2HPO_4, 0.2 g $MgSO_4{\bullet}7H_2O$, 0.2 g NaCl, 0.1g $FeCl_3$ and trace elements. 1000 x trace minerals contained per litre: 0.1 g H_3BO_3, 0.1g $ZnSO_4{\bullet}7H_2O$, 0.05 g $CuSO_4{\bullet}5H_2O$, 0.05 g $MnCl_2{\bullet}4H_2O$, 0.1 g $Na_2MoO_4{\bullet}2H_2O$, 1 g Na_2EDTA, 0.2 g NaFeEDTA.

Seven seedlings per pot were planted and the tinfoil covers were replaced and the pots then placed in a growth chamber on a 16 hour day cycle (22°C) and an 8 hour night (18°C). After two days growth, 0.2 ml of various *S. meliloti* overnight cultures inoculated from a single colony were diluted in 10 ml sterile ddH_20. The negative control was uninoculated water, positive controls included Sm1021 and RmP32 *(dme::gusA)* and RmP319 *nifH::gusA* reporter fusions. For every strain, each plant was inoculated in duplicate and allowed to grow uncovered for 5 weeks while plants were watered with sterile ddH_20 as needed.

β-glucuronidase assay of nodules

Nodules were removed from 5 week old plants. 5 nodules were placed in 0.75 ml of ice cold, sterile MMS buffer (40 mM MOPS, 20 mM KOH, 2 mM MgSO4, 0.3 M sucrose, pH 7) and ground with mini pestles and tubes (DiaMed). Plant tissue and debris

was sedimented by spinning at 2000 rpm for 2 minutes. 0.5 ml of the supernatant was removed and 1 % SDS added to a final concentration of 0.01% to allow for lyses of the bacteroids. β-Glucuronidase activity was assayed in triplicate as in Reeve et al. (Reeve et al. 1998) with a few changes. 100 μl of nodule squash products are added to 890 μl of GUS phosphate buffer (50 mM sodium phosphate, 50 mM DTT and 1 mM EDTA; pH 7) to an epindorf tube and vortexed briefly, no toluene is added for permeabilization because 0.01% SDS has already been added to the samples. 10 μl of PNPG (35 mg/ml) is added to the tubes and the reaction is allowed to proceed for approximately half an hour or until sufficient yellow color is observed. The OD_{405} reading was taken for the amount of GusA activity using the Tecan Safire microtitre plate reader. 10 μl of each nodule extract sample was assayed in triplicate to determine protein concentrations. Specific activity was calculated using the following equation:

(1000 x A_{405}) / reaction time in minutes x amount of protein in mg

Protein Determination in Nodule samples

The protein concentration of crude nodule extracts was determined by the method of Bradford (Bradford 1976) using the Bio-Rad Protein assay Microassay procedure for microtitre plates (Bio-Rad Laboratories). Bovine serum albumin (BSA) in concentrations ranging from 0.0 to 0.2 mg/ml were used to create a standard curve.

Preparation of Seed and Root Exudates

Alfalfa, sweet clover, pea, lentil and bean seeds were surface sterilized in 95% ethanol (1 minute) and 1.25% (w/v) hypochlorite (15 minutes) and rinsed five times with sterile water. For preparation of seed exudates the surface sterilized seeds were imbibed in sterile double distilled water in the dark for 6 hours. Two seeds per ml of water were used for large seeds (pea, lentil and bean) and for small seeds (alfalfa and sweet clover) the volume of water was four times the volume of seeds used. Alfalfa, sweet clover and pea root exudates were prepared by germinating seeds on water agar for 2 days then imbibing seeds in sterile water in the dark for 5 days. Approximately 40 pea, lentil and bean seeds were used for 20 ml of sterile water and a volume of 10 ml of alfalfa, sweet clover seeds were used for 40 ml of sterile water. All exudates were filtered though Whatman filter paper (No. 7) to remove plant debris, then though a 0.45 μm syringe filter to ensure sterility. If not used immediately exudates were stored at –20°C (Rosenblueth et al. 1998).

Bacterial matings

Plasmids were transferred to *S. meliloti* from *E. coli* by triparental mating using overnight cultures of the recipient strain, donor strain and helper strain MT616 (which carries the plasmid pRK600 and provides the transfer functions). These strains were centrifuged for 1 minute at 13 000 rpm, washed and resuspended in 0.5 ml of sterile 0.85% NaCl. 20 μl of each culture were then spotted onto an LB agar plate in the following order: donor, helper, then recipient and incubated overnight at 30°C. The

mating spot was resuspended in 1 ml of 0.85% NaCl and dilutions were plated on appropriate media for selection of the *S. meliloti* recipient and the plasmid.

Plasmid preparation

Plasmid DNA was isolated with the QIAquick miniprep kit (Qiagen) and the Gene Elute miniprep kit (Sigma) by following the manufacturer's directions.

DNA modifications

Restriction enzyme digests were carried out by following the manufacturer's directions (Roche, NEB, Invitrogen). 5' overhangs were removed with Mung Bean Nuclease (NEB) and 5' fill-in reactions were carried out with the Klenow fragment of DNA polymerase I (Invitrogen) according to manufacture's recommendations.

Ligation Reactions

Restriction enzymes and any other necessary modifying enzymes were used on plasmid and insert DNA. An excess of purified PCR product and the plasmid DNA were passed through a QIAquick PCR purification kit or purified by phenol extraction and ethanol precipitation. If the DNA remained in a large final volume it was evaporated from the sample in an Eppendorf Vacufuge at 45°C until no buffer remained. The ligation reaction was carried out in a 10 or 20 µl reaction containing ligase and ligation buffer as suggested by the manufacturer (NEB) and incubated overnight at 16°C.

Competent Cell Preparation

Competent cells were prepared by using 1 ml of an overnight culture of DH5α grown in LB to inoculate 100 ml of LB in an Erlenmeyer flask. The culture was grown to an OD_{600} of 0.4 and aliquoted into two 50 ml Falcon tubes and placed on ice for 0.5 hours. The cultures were pelleted by centrifugation for 10 minutes at 5000 rpm and gently resuspended in 10 ml of ice cold 100 mM sterile $CaCl_2$. The cultures were then stored on ice for a minimum of 30 minutes then again pelleted and resuspended in 2 ml of cold 100 mM $CaCl_2$. The competent cells were either used within 18 hours or mixed (2 ml) with 70 µl of DMSO and frozen down in microfuge tubes by liquid nitrogen and stored at –70°C (Sambrook J and Russell DW 2001).

Transformation

100 µl of fresh or frozen competent cells were mixed with a ligation reaction or plasmid DNA and incubated on ice for 30 minutes. Heat shock treatment was at 42°C for one minute then on ice for 2 minutes. 0.5 ml of LB broth was added then incubated in a rotary wheel at 37°C for approximately 2 hours. The transformation was then plated on the appropriate selective media and incubated overnight at 37°C.

Agarose gel electrophoresis

Agarose gels were used to visual DNA fragments. 0.8 to 1.2% agarose gels were used (BioShop) depending on the size of DNA for electrophoresis with 0.5X Tris-borate-EDTA (TBE) running buffer. DNA Samples were mixed with 6X loading buffer (0.25%

xylene cyanol FF and 15% Ficoll) before loading into the gel. Samples were electrophoresed at 60 to 110 Volts for 1 to 2 hours. To visualize DNA, gels were stained for approximately 20 minutes in water and ethidium bromide (0.05 μg/ml) and destained for 15 minutes in ddH_2O then photographed on a transilluminator.

Genomic DNA preparation

Overnight cultures of *S. meliloti* were grown in 5 ml of LBmc and centrifuged for 10 minutes at 5000 rpm. The pellet was then washed in 5 ml 0.85% NaCl and resuspended in 1 ml of T_{10} E_{25} (10 mM Tris-HCl, 25 mM EDTA, pH 8), then the following were added: 50 μl of 25%SDS, 25 μl of Proteinase K (10 mg/ml) and 125 μl NaCl (5 M). This mixture was gently mixed, incubated at 65°C for 30 minutes then cooled to room temperature. Three phenol:chloroform extractions were carried out followed by one extraction with a large volume of chloroform. The aqueous layer was removed and precipitated with ammonium acetate to a final concentration of 2 M and then added to an equal volume of 100% ethanol. Centrifugation was carried out to precipitate the DNA and the supernatant is removed. The pellet was washed with 70% ethanol and dried at 37°C. DNA was resuspended in 50 μl $T_{10}E_1$ with 20 μg/ml RNaseA and incubated at 65°C for 30 minutes .

Southern transfer.

Genomic DNA (5 μg) was digested with the appropriate restriction enzymes for 5 hours and electrophoresed on a 0.8% agarose gel overnight at 15 Volts. The gel was

incubated in alkaline transfer buffer (0.4 M NaOH, 1M NaCl) for 30 minutes (while changing the buffer once). The blotting apparatus was assembled by stacking paper towels, 4 Whatman filter papers, 2 filter papers which had been soaked in transfer buffer, the nylon membrane (Bio-Rad Zeta-probe Blotting membrane), the gel, another 2 filter papers soaked in transfer buffer and a long piece of filter paper to be used as the wick which had been wet with transfer buffer with its ends in receptacles filled with more alkaline transfer buffer. After transfer for approximately 6 hours the membrane was soaked in 0.5 M Tris-HCl (pH 7.5), 1 M NaCl for 15 minutes then air dried. The membrane was then wrapped in saran wrap and fixed under UV light on the transilluminator for two minutes (Sambrook J and Russell DW 2001).

Preparation of Random Primed DNA probe

The plasmid to be used for a probe was linearized and passed through a Qiagen PCR purification column then radiolabeled using the Roche Random Primed DNA Labeling Kit. The plasmid DNA was brought up to a volume of 9 μl, boiled in a water bath for 10 minutes then immediately placed on ice. Added to the denatured probe was: 3 μl dNTPs (0.5 mM each except for dATP), 2 μl of hexanucleotide mixture in 10x reaction buffer, 1 μl of Klenow enzyme and 5 μl of 50μCi [α-^{33}P]dATP (3000 Ci/mmol). This mixture was centrifuged briefly and incubated at 37°C for one hour then purified in a Qiagen PCR purification column.

Hybridization of Radioactively labeled probe to target DNA

In a 50 ml Falcon tube or hybridization tube the membrane and 10 ml of prehybridization solution (5X SSC, 5X Denhardt's Reagent, 0.5% SDS, 100 µg/ml sheared Herring sperm DNA, 50% formamide, 0.1% Na pyrophosphate) were incubated at 42°C for 2.0 hours. The labeled probe was then boiled for 10 minutes and denatured by being immediately placed on ice. The probe was added to the membrane and prehybridization mixture and incubated at 42°C for 18 hours.

Detection

The hybridization solution was removed then the membrane washed with 10 ml of the prehybridization solution at room temperature for 15 minutes. The membrane was then washed in 4 washes. The first (2x SSC, 0.5% SDS) and the second washes (2x SSC, 0.1% SDS) were for 15 minutes at room temperature, the third (0.1x SSC, 0.1% SDS) at 65°C for at least 0.5 hours and the fourth (0.1% SSC) briefly at room temperature. The membrane was air dried, wrapped in saran wrap and exposed to Kodak Scientific Imaging Film (Cat no. 165 1454) in a cassette with an intensifying screen. The film was developed in a Kodak X-OMAT 2000A Processor and the molecular weights of the bands were calculated by referring to the photograph of the agarose gel that was taken with the ruler along side the 1 kb ladder. Also the membrane was exposed to a Storage Phosphor Screen (Amersham Biosciences).

PCR

Primers were synthesized (Mobix, Sigma Genosys) and resuspended in sterile ddH_2O to 100 pmoles/µl. PCR reactions were carried out in a PerkinElmer GeneAmp PCR system 2400 or an Eppendorf Mastercycler epgradient S. A mix of 100 µl was used for each reaction containing: 10 µl 10x buffer, 16 µl dNTPs (1.25 mM stock), 1 µl of each primer, 1.0 to 2.5 mM $MgCl_2$, 0.2 µl Platinum Taq polymerase (Invitrogen) brought up to 95 µl with ddH_2O and lightly vortexed. This mixture was added to 5 µl of template DNA (2 ng/µl genomic DNA). Each reaction began with an initial melting for 2 minutes at 95°C followed by 30 cycles of amplification with 30 seconds of melting (95°C), 30 seconds of extension (72°C) and 40 seconds of annealing ranging from 55 to 62°C depending on the melting temperatures of the primers. The final step was an extension for 7 minutes at 72°C. Successful reactions were confirmed by electrophoresing 2 to 5 µl of the reaction on an agarose gel and visualizing. Before use in a cloning procedure the PCR product was passed through a Qiagen PCR purification kit.

Colony PCR

PCR was used to screen recombinant plasmids for insert DNA in some cases. Using a 200 µl pipetman tip a fresh colony was picked, patched onto a plate and then placed into a microfuge tube containing 50 µl ddH_2O. The tubes containing tips were vortexed then placed in a boiling water bath for 5 minutes. The tubes were then spun in a microfuge for 2 minutes at high speed. 10 µl of the supernatant was added to a PCR tube to be used as template DNA. A master mix was prepared containing buffer, primers,

$MgCl_2$ and dNTPs with 40 µl aliquots being added to the template DNA to undergo a standard PCR reaction. To screen for inserts 5 µl of the reaction was electrophoresed on an agarose gel. Positive clones were then inoculated into LB media from the patched plate and recombinants confirmed by restriction enzyme digest.

DNA sequencing and analysis

DNA sequencing was carried out using dye terminator chemistry and cycle sequencing on an ABI 373 Stretch automatic sequencer (Mobix). The pTH1360 sequencing primer ML758 (5'- TCAAGCTTGCATGCCTGC -3') was used for sequencing *gusA* reporter fusion clones.

CHAPTER 3 – RESULTS

3.1. Deletion mutant screening

Deletion mutant strains of wild type *S. meliloti*, RmG462, Rm5406, Rm5408, RmG373, RmF909, RmF117 were previously generated with large regions of pSymB ranging from 120 – 600 kb deleted. These strains were made by homologous recombination occurring between the IS50 elements of transposons that had previously been inserted. The deletion strains were screened for loss of growth on various carbon sources such as succinate, protocatechuate and galactitol, although many of these defined deletions yielded no phenotype (Charles and Finan 1991). The exact locations of the deletions were determined by sequencing outwards from the transposons so that the endpoints could be defined (Susan Lehman, Undergraduate Thesis). Several of these strains whose deletions together cover 1.4 Kb of pSymB were previously screened by the Biolog © MicroArray system (PMs), (Punita Anja & Finan, unpublished data) and analyzed in this study. The Biolog system consists of four 94 well microtitre plates with each well containing a different substrate. Two of the plates contain carbon sources, one contains nitrogen sources and the other is made up of sulphur and phosphorous sources. Cellular respiration is used as a reporter using redox chemistry. If a phenotype is positive the cells respire and a tetrazolium dye is reduced forming a dark colour. If the substrate cannot be used or is only weakly used there will be little or no respiration therefore little or no formation of color, indicating a negative phenotype (Bochner et al. 2001). By comparing the phenotype of wild type strain Sm1021 to the various deletion mutants one can decipher the region where genes involved in the uptake and/or metabolism of the

substrate in question is located on pSymB. The analysis of the Biolog data is listed in Appendix A.

To verify the Biolog PMs data the deletion mutants were then screened on M9 minimal media agar plates containing various carbon and nitrogen sources for which non-utilization phenotypes were suggested. These results are shown in Table 1. The compounds whose metabolism and/or transport genes were deduced to be present on pSymB due to the observation of a utilization phenotype were used as a starting point for compounds to screen the ABC transporter reporter fusions.

Table 3-1. Sole Carbon sources not utilized deletion by pSymB deletion mutants

Strain (deletion)	Deletion Size (bp)	Location (nucleotide)	Sole Carbon Source
RmG462 (Ω5033Δ5011)	288480	889262 - 1177742	L-histidine L-lysine Galactitol D-tagatose
RmG373 (Ω5177Δ5069)	293645	1452882 - 63194	D-melibiose D-raffinose Thymidine L-leucine L-valine L-isoleucine L-serine L-arginine L-ornithine Glycine Succinate Fumarate Malate
RmF909 (Ω5085Δ5047)	629383	106128 - 735511	Palatinose D-galactosamine Hydroxyproline
RmF117 (Ω5060Δ5033)	119173	770089 - 889262	D-galactosamine Thymidine

3.2. Properties of β-glucuronidase assay

In order to develop a high-throughput GusA screening assay various parameters of the assay were studied. Factors such as permeabilizing agent (toluene vs. chloroform) and assay temperature were varied in order to optimize the ß-glucuronidase assay for high-throughput screening. The ß-glucuronidase assay developed by Jefferson (Jefferson et al. 1986) and used in *S. meliloti* by Reeve et al. (Reeve et al. 1998) has several time consuming steps which are not conducive to high-throughput methods. Thus cells are washed with saline and resuspended in the GUS buffer (50 mM phosphate buffer, 50 mM DTT, 1 mM EDTA, pH 7) before subsequent addition of more GUS buffer, permeabilization and the enzyme assay. Figure 3-1 demonstrates that this wash step is unnecessary and using cell culture directly from growing cultures in the various test media will maintain the integrity of the assay as only a slight reduction of activity is observed when the wash step is omitted. Optimization of the high-throughput ß-Glucuronidase assays were carried out using the *gusA* fusion in strain RmP187 (SMb21373::*gusA*) as this was highly induced in LBmc and not expressed in M9 minimal media containing glucose or succinate. Casamino acids (1%) were also added to M9 minimal media containing glucose or succinate to determine if it was the presence of an amino acid in the LB media which was inducing this gene. As seen in Figure 3-1 SMb21373 was only induced when grown in LBmc and not in M9 minimal media containing casamino acids.

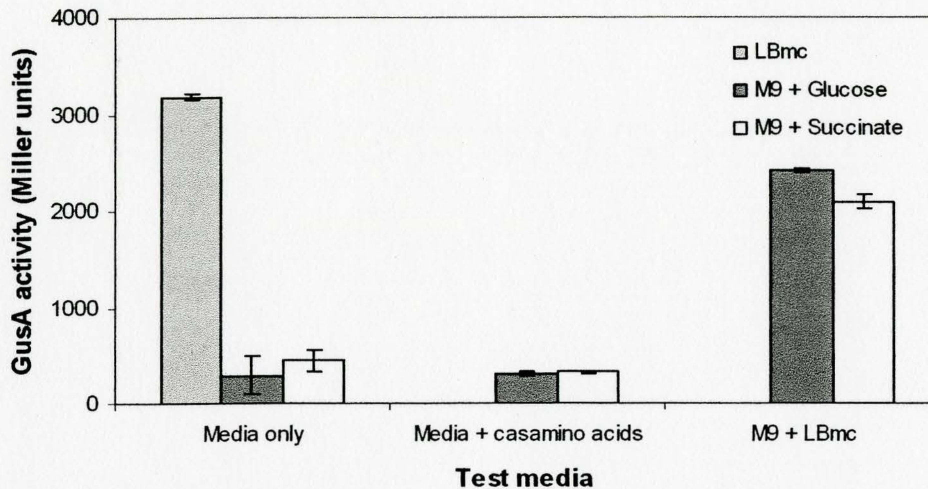

Figure 3-1. ß-glucuronidase assay of RmP187 (SMb21373::*gusA*) grown in LBmc and M9 minimal media with succinate or glucose as the sole carbon sources and with and without 1% casamino acids.

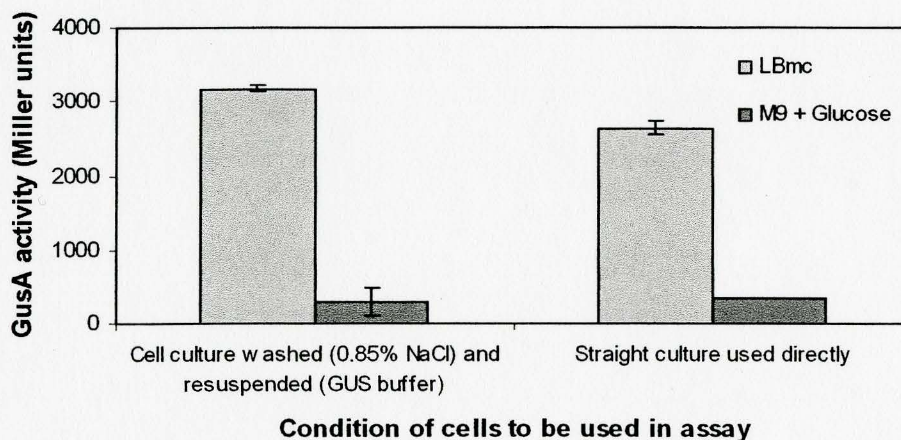

Figure 3-2. ß-glucuronidase assay of RmP187 (SMb21373::*gusA*) comparing the effect of using cell culture directly as opposed to cells washed with 0.85% NaCl and resuspended in GUS buffer cell culture in the assay.

In the traditional assay, permeabilization of the cells is carried out by the addition of the permeabilizing agent toluene followed by vortexing. This step was another

potential road block to a high-throughput method. In Figure 3-2 a comparison between vortexing and no agitation of samples after addition of a drop of permeabilizing agent shows no discernable difference between the two treatments.

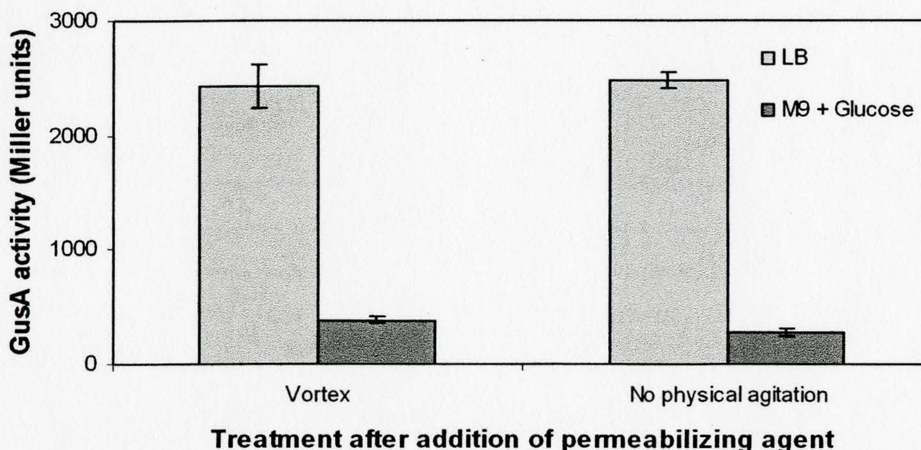

Figure 3-3. ß-glucuronidase assay of RmP187 (SMb21373::*gusA*) comparing the effects of vortexing versus using no physical agitation after the addition of a drop of permeabilizing agent (toluene) to the sample.

When using the PerkinElmer MultiprobeII liquid handling system to carry out the assay it was not convenient to allow the reaction to take place in a water bath or incubator to achieve the optimum temperature of 37°C. The variation of temperature on GusA activity was measured using different temperature conditions to determine the effects of temperature on the assay. As seen in Figure 3-3 carrying out the assay in the 37°C incubator or at room temperature (~20°C) slightly reduced the enzyme activity in both the LBmc grown cells and the M9 minimal media and glucose grown cells. Although carrying out the assay in a 37°C water bath yielded higher GusA activity it did not greatly affect the outcome of the results.

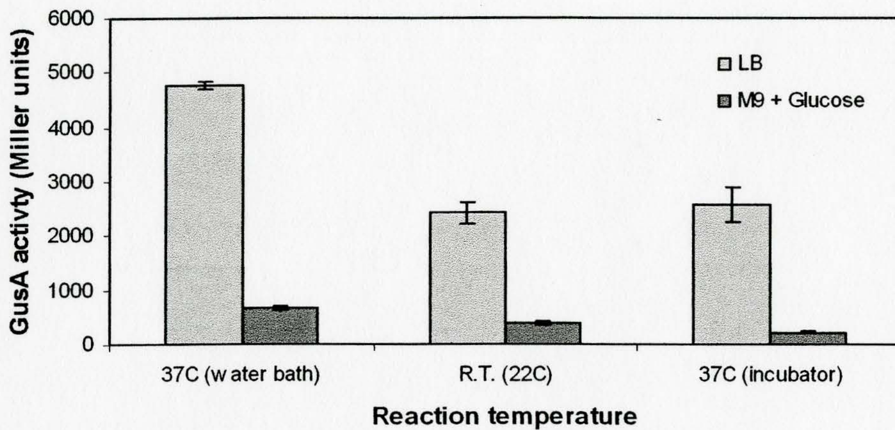

Figure 3-4. ß-glucuronidase assay of RmP187 (SMb21373::*gusA*) comparing the effects of reaction temperature on the levels of GusA activity.

When screening many different strains under several conditions it is expected that some cultures will have to be placed on ice or frozen before the assay is carried out. Figure 3-4 shows different storage conditions for varying amounts of time (hours) and the effects on GusA activity.

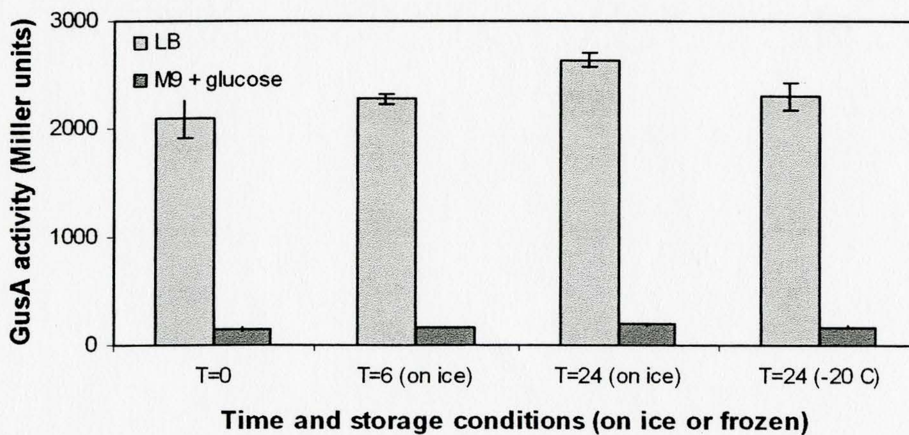

Figure 3-5. ß-glucuronidase assay of RmP187 (SMb21373::*gusA*) comparing the effects cold storage conditions and length of storage time on GusA activity.

The high-throughput assay was carried out on polystyrene plates that are solubilized by agents such as toluene or chloroform therefore for an efficient permeabilizing step a non-organic solution must be used. A final concentration of 0.01% SDS was found to be a suitable replacement as a cell permeabilizing agent instead of toluene. When a strain containing a *gusA* fusion to the promoter region of *pckA* (a gene induced by succinate) the levels of β-glucuronidase after growth in M9 minimal media with glucose and with succinate were found to be comparable when permeabilized using a drop of toluene or a final concentration of 0.01% SDS. When grown with glucose as the carbon source both methods of permeabilization had β-glucuronidase levels of approximately 600 Miller Units. After growth with succinate as the sole carbon source, the toluene permeabilized sample and the SDS permeabilized sample had β-glucuronidase levels of approximately 1150 and 1100 Miller Units, respectively (Rahat Zaheer, personal communication).

3.3. Generation of *gusA* reporter gene fusions

In order to study the expression of putative ABC transporter genes located on pSymB, reporter fusion strains were created. A GusA reporter plasmid pTH1360 (pVO155 derivative) with the GusA cassette from pFUS1 was used for cloning, as seen in Figure 3-5 (Oke and Long 1999). This plasmid cannot replicate in *S. meliloti*; features of this plasmid include neomycin/kanamycin and ampicillin antibiotic resistance cassettes, transcriptional terminators in front of the multiple cloning site (in order for knock-out

mutants to be made) and the *uidA* gene which encodes the ß-glucuronidase (GusA) protein from *E. coli* as a reporter (Jefferson et al. 1986).

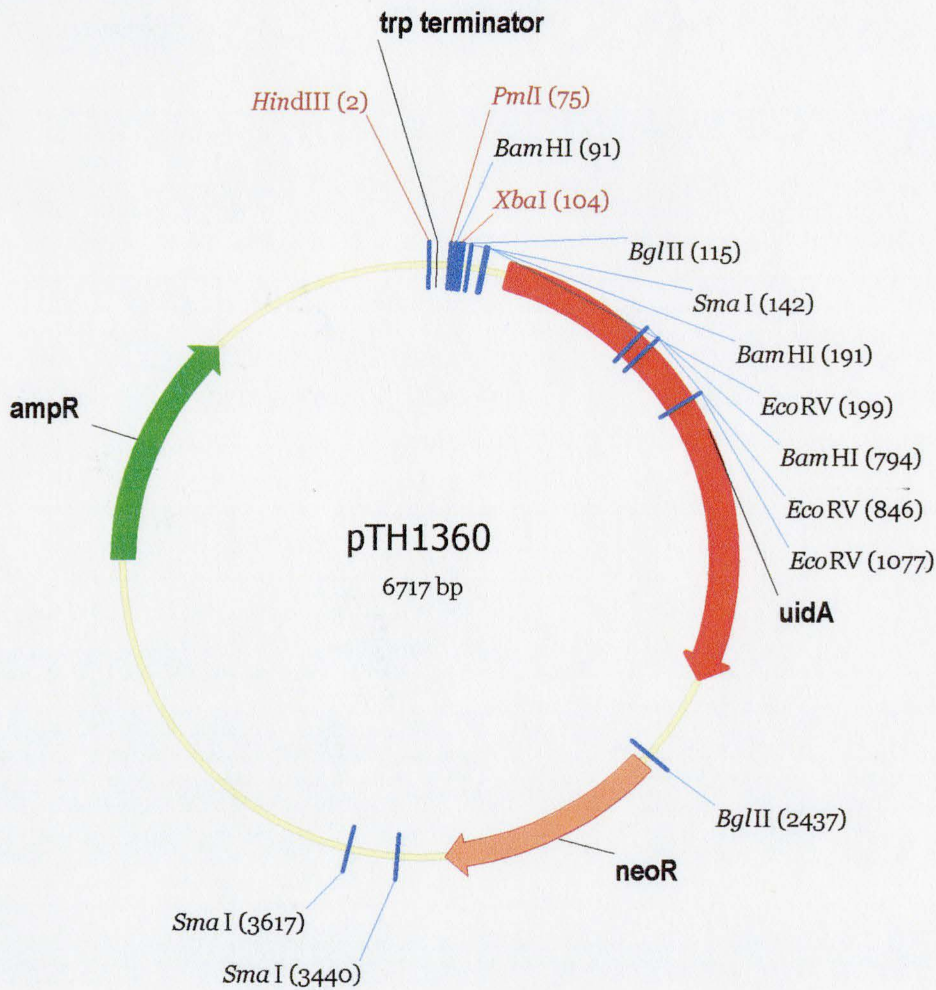

Figure 3-6. Plasmid map of *gusA* reporter suicide vector, pTH1360

Approximately 250 to 500 bases at the 3' regions of putative ABC transport clusters were amplified by PCR and cloned into the *PmlI* and *XbaI* restriction enzyme sites of the vector pTH1360 (Figure 3-5). Following transfer of this plasmid to *S. meliloti* and recombination into the genome the resulting strains had a single copy of the ß-glucuronidase gene downstream from the last gene in the operon, under control of the

native promoter, as shown in Figure 3-6. The transcriptional fusion strains were built to the 3' ends of the putative operons in order to retain a wild type functional operon.

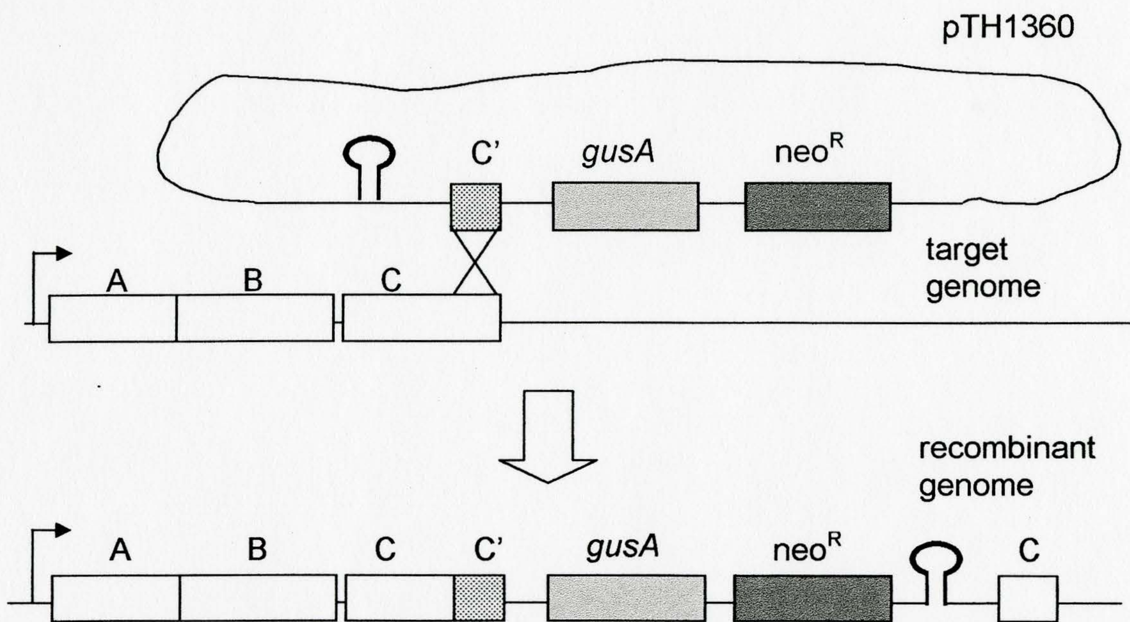

Figure 3-7. Single cross-over occurring in order to create *gusA* fusion strains under the control of the native ABC transport operon promoter

The vector pTH1360 was also used to make knock-out mutants by cloning a fragment internal to the first gene of the operon into the plasmid. This will result in a promoter dissociated from the rest of the operon under its control (Figure 3-7). The presence of the transcriptional terminator upstream of the pTH1360 cloning site eliminates any vector derived transcription of the gene cloned in pTH1360.

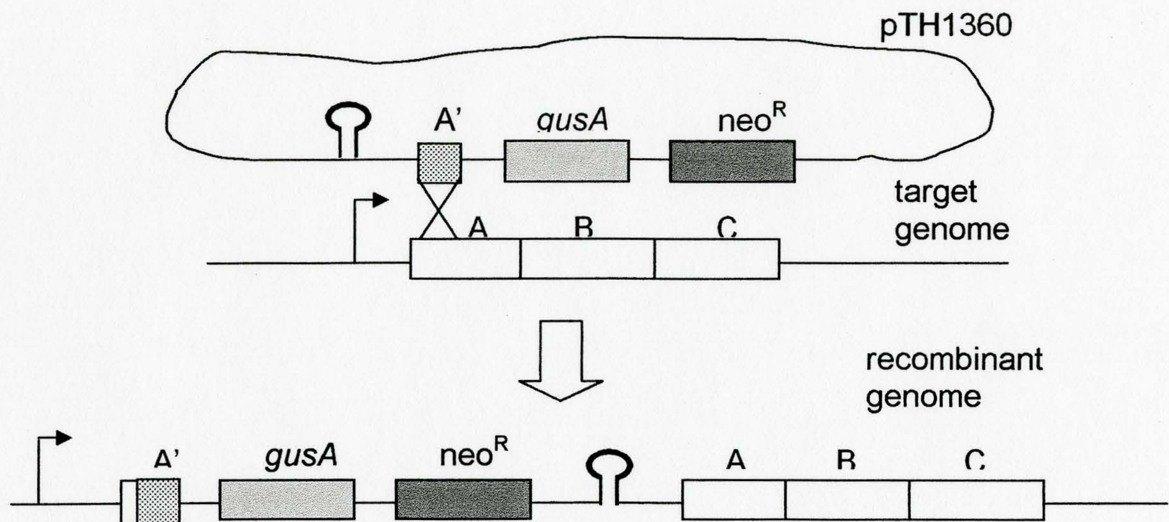

Figure 3-8. Single cross-over in order to dissociate a promoter from its operon to create a knock-out strain.

To determine the structure of the targeted gene region in the resulting recombinants, total DNA from three recombinants was analysed by southern hybridization with linearized pTH1360 as the labelled probe (Figure 3-8). Correct band sizes were obtained for the three strains tested, indicating that only one copy of pTH1360 had recombined. For each of the strains the *SmaI* digest would generate fragments of 2.9Kb and 4.3Kb in size. Refer to Figure 3-5 for location of *SmaI* sites within pTH1360. The *BglII* digest of RmP199 (SMb20318::*gusA*) would yield two fragments of 2.3 and 5.0 Kb, RmP187 (SMb21373::*gusA*) would yield 2.3 and 9.0 Kb and RmP189 (SMb21097::*gusA*) would yield 2.3 and 6.5 Kb. The position of the BglII sites in the genome was deduced from the genome sequence. Figure 3-9 shows the *BglII* restriction maps for RmP199, RmP187 and RmP189.

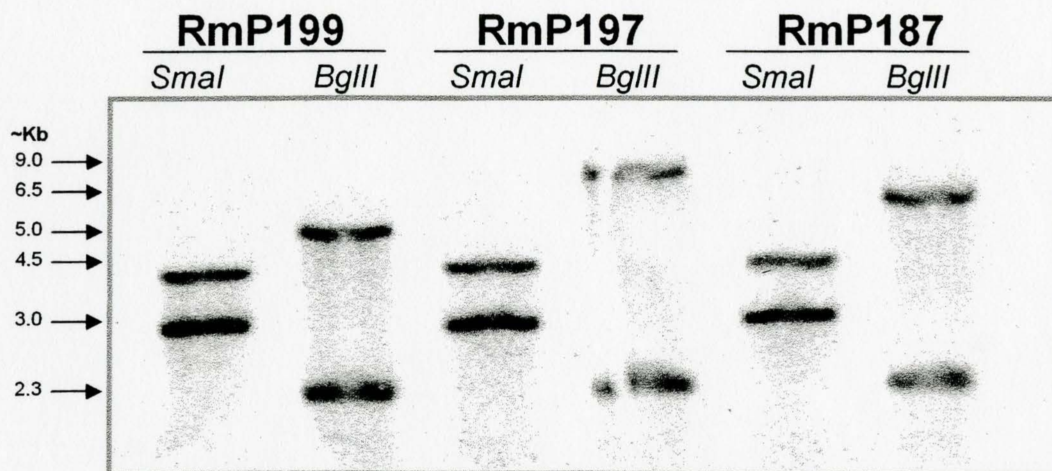

Figure 3-9. Southern Blot of total DNA from RmP199 (SMb20318::*gusA*), RmP187 (SMb21373::*gusA*) and RmP189 (SMb21097::*gusA*) reporter fusion strains using a *SmaI* and *BglII* digests.

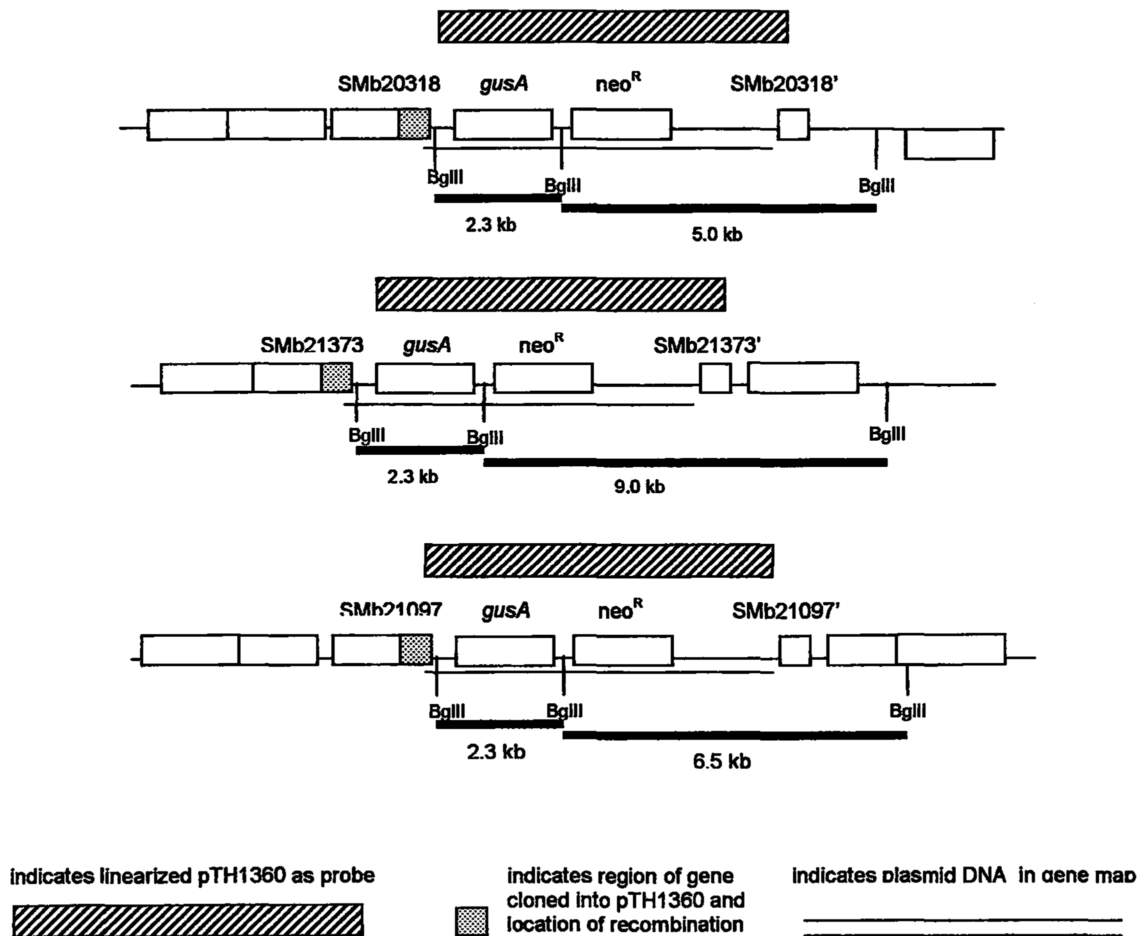

Figure 3-10. Predicted restriction maps of target regions after recombination of the plasmids used in the construction of strains RmP199 (SMb20318::*gusA*), RmP187 (SMb21373::*gusA*) and RmP189 (SMb21097::*gusA*) for a *BglII* digest

3.4. The high-throughput ß-glucuronidase screening strategy

In order to assay the expression of many genes under various conditions we optimized the assay to measure β-glucuronidase activity in a 96 well microtitre format (see section 3.1). This optimized assay has eliminated the separate steps of permeabilizing agent addition and substrate addition by using one pre-mixed reaction mixture. Any positive results obtained through the modified assay were repeated by the

standard lab bench assays (Reeve et al. 1998). Forty-five different fusion strains were examined in addition to a positive control strain RmP32 (*dme::gusA)* (Driscoll and Finan 1993), Rm1021 as a negative control and a blank (uninoculated control) to ensure no contamination. Table 3-2 lists the compounds tested for possible induction of the fusion strains, including seed and root exudates.

Table 3-2. Substrates and exudates used in the high-throughput screen to investigate expression of fusion strains.

LBmc media	Hydroxyproline
Casamino acids	Xylose
Glucose	Sucrose
Succinate	Methyl-pyruvate
Glycerol	Dextrin
Galactitol	Fucose
Raffinose	Trehalose
Galacturonic acid	Arabitol
Histidine	Protocatechuate
Mannose	p-hydroxybenzoic acid
Ribose	Alfalfa seed exudates
Arabinose	Alfalfa root exudates
Palatinose	Sweet clover seed exudates
Galactose	Sweet clover root exudates
Sorbose	Bean seed exudates
Sorbitol	Lentil seed exudates
Myo-inositol	Pea seed exudates
Glucosamine	Pea root exudates

3.5. Results of high-throughput screens

The following diagrams depict typical results of the high-throughput ß-glucuronidase assays. The assays were carried out with one replicate upon which any putative positive results were retested in the standard bench top assay in triplicate. Of the forty-five fusion strains, four were found to be induced by only one of the test

compounds used in a single screen. In the following results sections the various reporter fusions are referred to by their strain name (ie. RmP215) and also by the gene name as given by the *Sinorhizobium meliloti strain 1021* Genome Project website which is also the designation used by GenBank (ie. SMb21373). The reason for referring to strains as the actual gene name is to provide simplicity for the reader due to the large number of genes/strains included in this study. Figures 3-10 and 3-11 show the ß-glucuronidase activity measured when strains RmP193 (SMb21138::*gusA*) and RmP198 (SMb20979::*gusA*) were cultured under 27 different conditions. RmP193 (SMb21138::*gusA*) was induced by galactosamine and RmP198 (SMb20979::*gusA*) was induced by fucose.

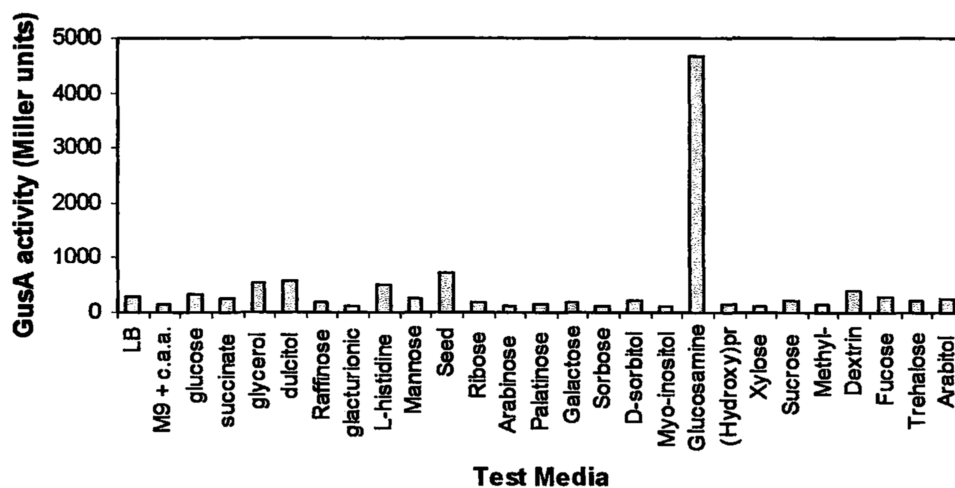

Figure 3-11. ß-Glucuronidase assay of RmP193 (SMb21138::*gusA*) grown in M9 minimal media with various test carbon sources.

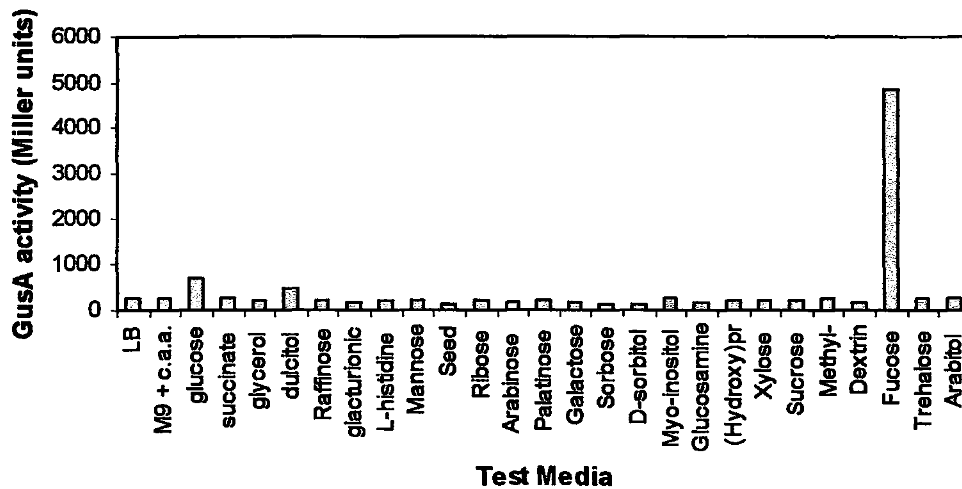

Figure 3-12. ß-Glucuronidase assay of RmP198 (SMb20979::*gusA*) grown in M9 minimal media with various test carbon sources.

Some of the fusion strains were induced by multiple compounds. The following two figures show the results of ß-glucuronidase assays for RmP226 (SMb20428::*gusA*) which was induced by mannose, xylose and fucose and RmP203 (SMb21342::*gusA*) which was induced by raffinose, alfalfa seed exudates and galactose. All together fourteen of the fusion strains demonstrated induction by two or more of the test compounds, including the various seed exudates.

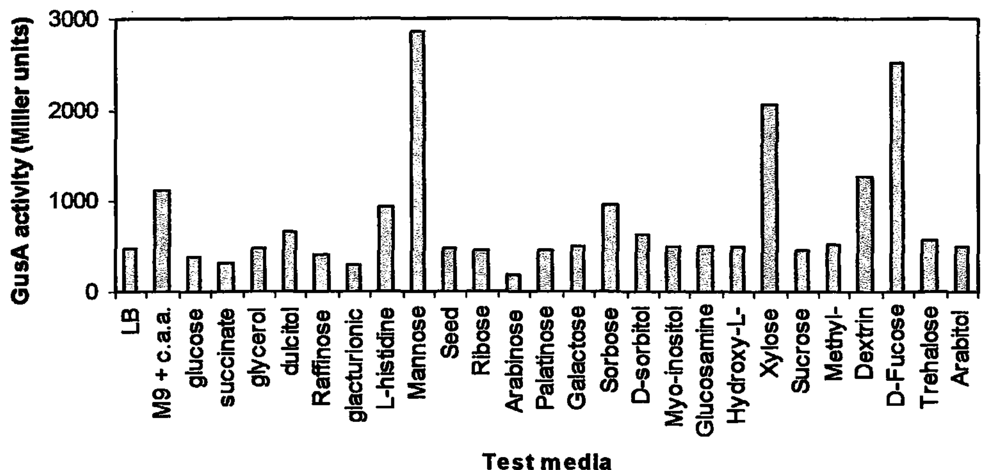

Figure 3-13. ß-glucuronidase assay of RmP226 (SMb20428::*gusA*) grown in M9 minimal media with various test carbon sources.

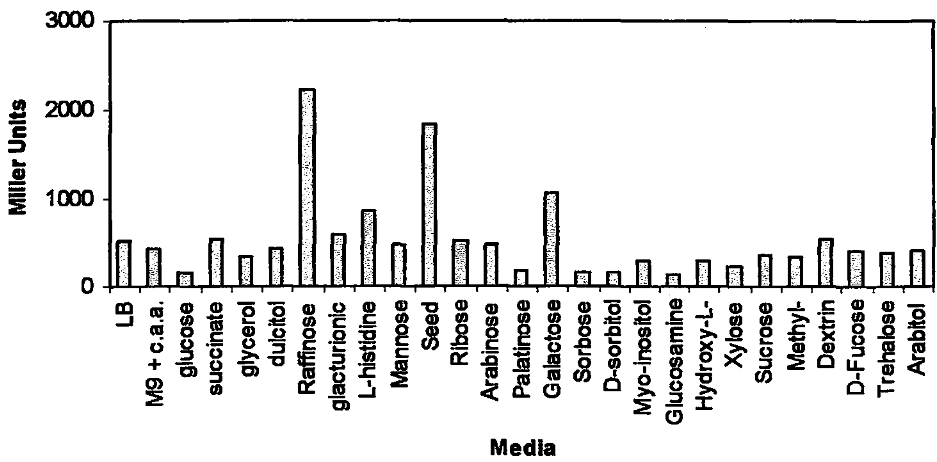

Figure 3-14. ß-glucuronidase assay of RmP203 (SMb21342::*gusA*) grown in M9 minimal media with various test carbon sources.

For thirty of the fusion strains no inducing substrate could be found. An example of which is seen in the below figure depicting the results of RmP196 (SMb20235::*gusA*), a component of a sugar ABC transport system which was not induced by any of the test substrates.

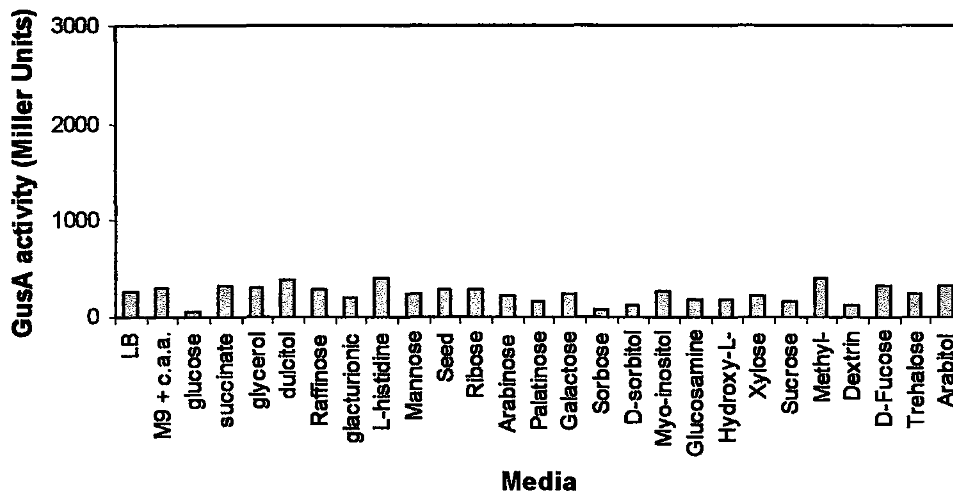

Figure 3-15. ß-Glucuronidase assay of RmP190 (SMb20235::*gusA*) grown in M9 minimal media with various test carbon sources.

As the high-throughput screen produces crude data with only one replicate, it was necessary to repeat the assays, in more detail using the classical method for ß-glucuronidase assays (Reeve et al. 1998). The results obtained in the initial screening for RmP206 (SMb20904::*gusA*) were repeated in triplicate as seen in Figure 3-15 to confirm the original high-throughput screening results. RmP234 (SMb20902::*gusA*) a knock-out mutant of this ABC transport operon was tested for the ability to use glucose, mannose and xylose as the sole carbon sources on M9 minimal media plates and was found to grow normally on all of it's inducing sugars.

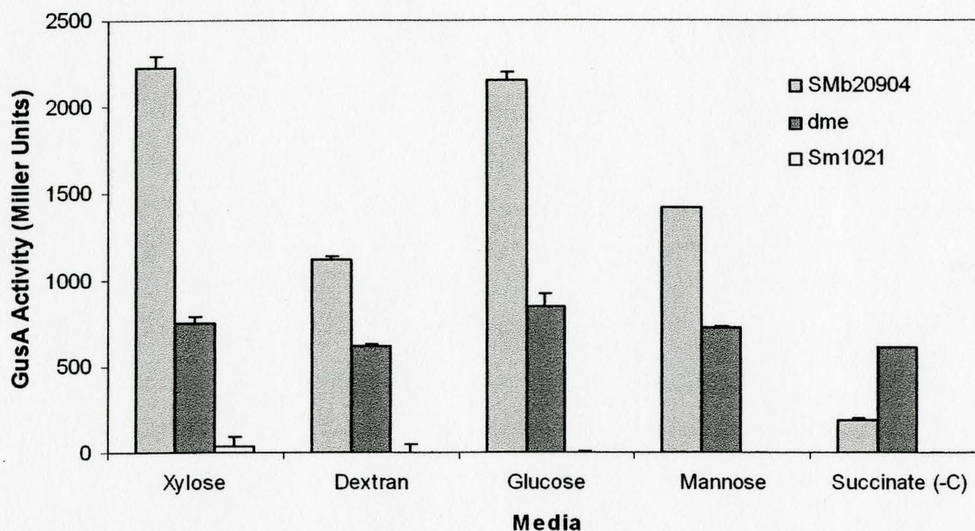

Figure 3-16. ß-glucuronidase assay of RmP206 (SMb20904::*gusA*) in order to confirm results obtained in the high-throughput method of screening with Rm1021 as a negative control and RmP32 (*dme*::*gusA*) as a positive control.

To confirm the high-throughput screening data, conventional, large volume β-glucuronidase assays were carried out in triplicate on cells grown with various carbon sources (Figure 3-15 to 3-22). Rm1021 was used as a negative control and RmP32 (*dme*::*gusA*) as a positive control in each assay. RmP198 (SMb20979::*gusA*) was found to induced by fucose (Figure 3-16), RmP266 (SMb20428::*gusA*) to be induced by mannose and to a lesser extent fucose (Figure 3-17) and RmP192 (SMb21707::*gusA*) was induced by glucosamine, galactosamine and N-acetyl-D-glucosamine (Figure 3-18). RmP192 (SMb21707::*gusA*) is also induced by galacturonic acid. RmP202 (SMb21342::*gusA*) was found to be induced by sorbose (Figure 3-19) and RmP222 (SMb21587::*gusA*) was found to be induced by arabinose and galactose (Figure 3-20).

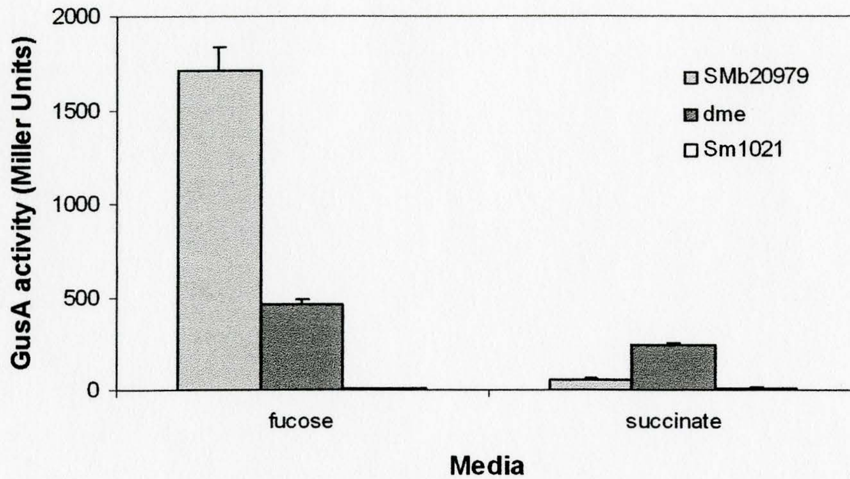

Figure 3-17. β-glucuronidase assay of RmP198 (SMb20979::*gusA*) with Rm1021 as a negative control and RmP32 (*dme*::*gusA*) as a positive control grown in M9 minimal media with fucose and succinate as the sole carbon sources.

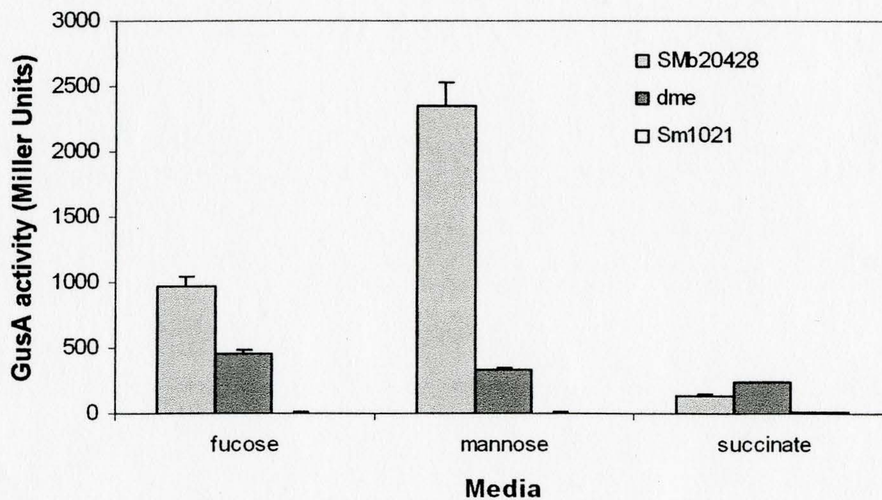

Figure 3-18. β-glucuronidase assay of RmP266 (SMb20428::*gusA*) with Rm1021 as a negative control and RmP32 (*dme*::*gusA*) as a positive control grown in M9 minimal media with fucose, mannose and succinate as the sole carbon sources.

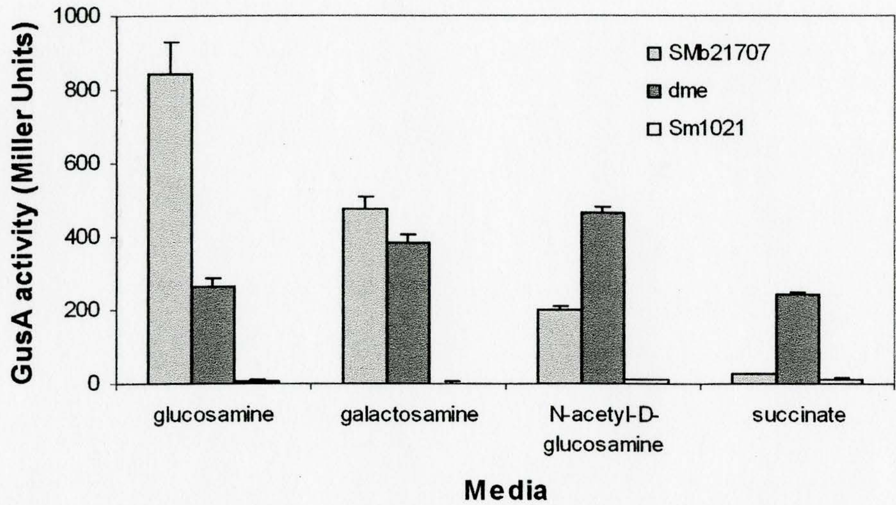

Figure 3-19. β-glucuronidase assay of RmP192 (SMb21707::*gusA*) with Rm1021 as a negative control and RmP32 (*dme*::*gusA*) as a positive control grown in M9 minimal media with glucosamine, galactosamine and N-acetyl-D-glucosamine as the sole carbon sources.

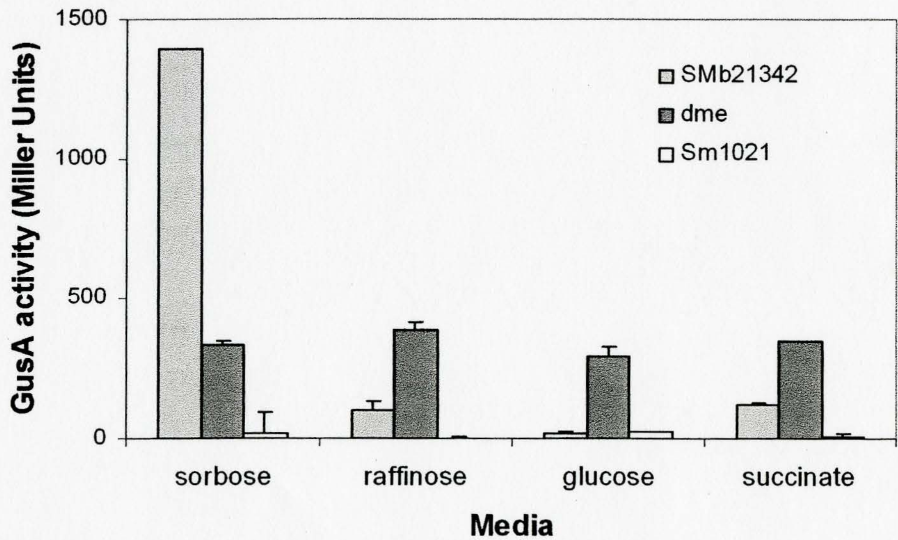

Figure 3-20. β-glucuronidase assay of RmP202 (SMb21342::*gusA*) with Rm1021 as a negative control and RmP32 (*dme*::*gusA*) as a positive control grown in M9 minimal media sorbose, raffinose, glucose and succinate as the sole carbon sources.

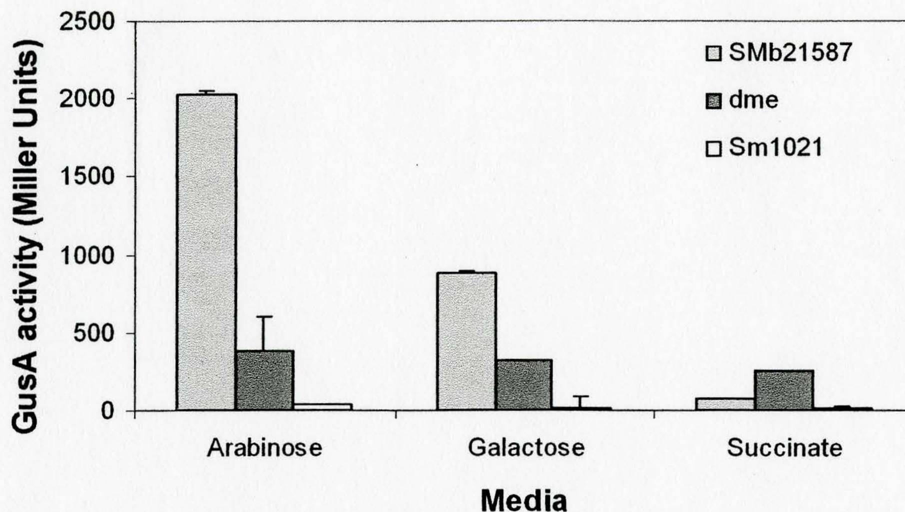

Figure 3-21. β-glucuronidase assay of RmP222 (SMb21587::*gusA*) with Rm1021 as a negative control and RmP32 (*dme*::*gusA*) as a positive control grown in M9 minimal media arabinose, galactose and succinate as the sole carbon sources.

The following two figures (3-21 and 3-22) show strains that were induced by protocatechuate and p-hydroxybenzoate. RmP188 (SMb20784::*gusA*) was found to be induced by protocatechuate and p-hydroxybenzoate, but induction by protocatechuate was about three times higher than seen for p-hydroxybenzoate. RmP230 (SMb20027::*gusA*) was also induced by these two compounds and after growth in protocatechuate the GusA activity was approximately 1.4 fold than seen for growth in p-hydroxybenzoate.

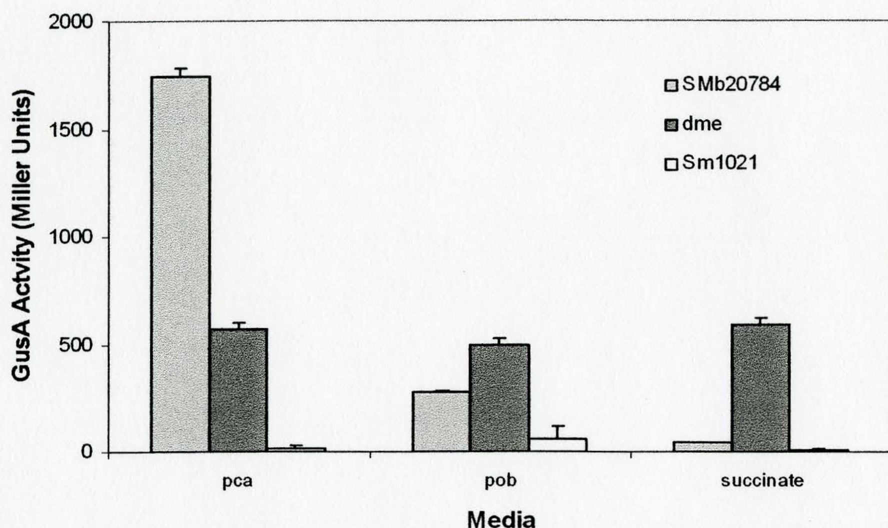

Figure 3-22. ß-glucuronidase assay of RmP188 (SMb20784::*gusA*) grown in M9 minimal media with protocatechuate (pca) and p-hydroxybenzoate (pob) as the sole carbon sources.

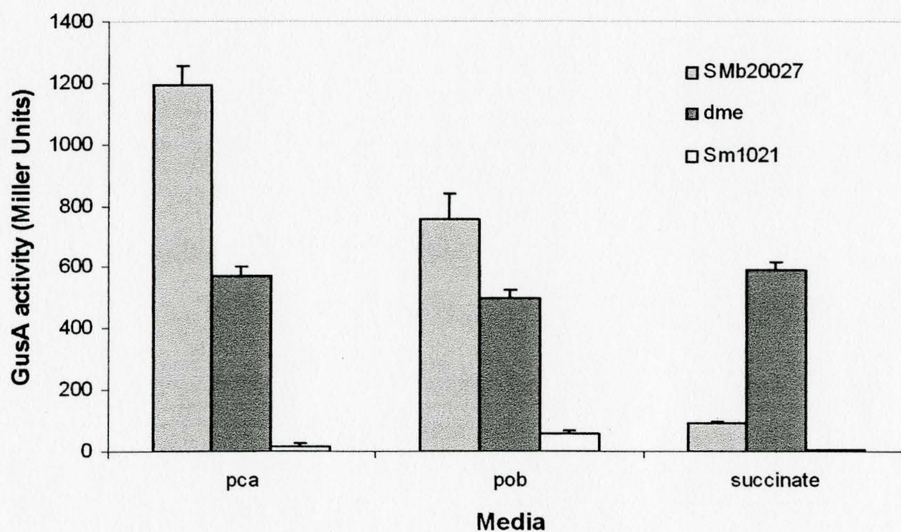

Figure 3-23. ß-glucuronidase assay of RmP230 (SMb20027::*gusA*) grown in M9 minimal media with protocatechuate (pca) and p-hydroxybenzoate (pob) as the sole carbon sources.

3.6. Legume seed and root exudates

In order to find biologically relevant inducers of ABC transporters, seed and root exudates from alafalfa, pea, sweet clover, lentil and bean were used as growth media for the reporter strains. RmP194 (SMb21216::*gusA*) and RmP193 (SMb21138::*gusA*) were both found to be highly induced in pea seed and root exudates as seen in the following figures.

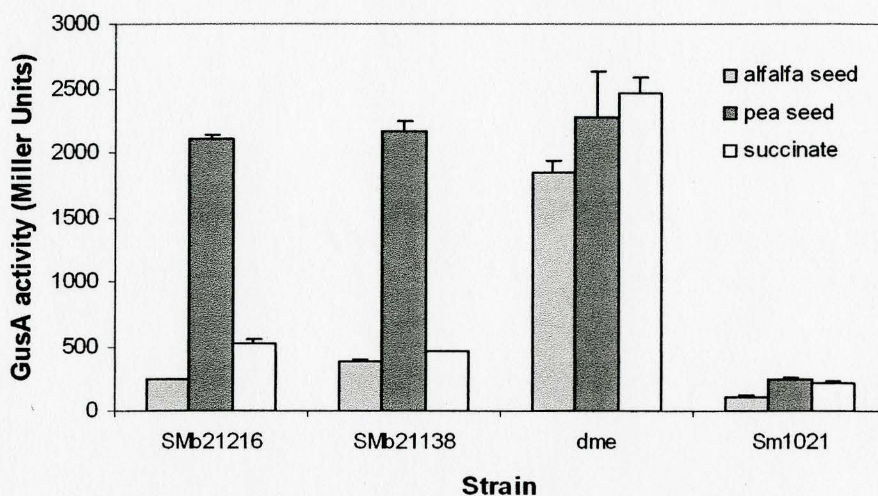

Figure 3-24. ß-glucuronidase assay of RmP193 (SMb21138::*gusA*) and RmP194 (SMb21216::*gusA*) grown in alfalfa and pea seed exudates.

Although no specific inducer(s) of RmP194 (SMb21216::*gusA*) have been found RmP193 (SMb21138::*gusA*) was found to be induced by glucosamine and galactosamine in addition to pea exudates but was not found to be induced by N-acetyl-glucosamine (Figure 3-24).

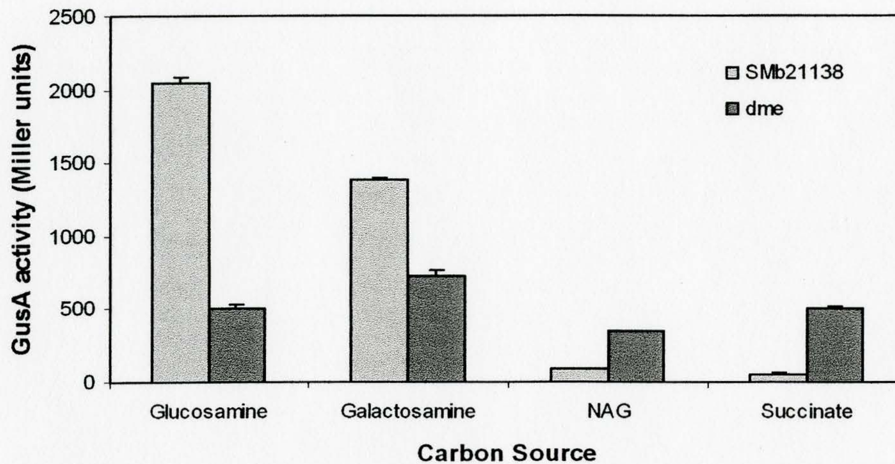

Figure 3-25. ß-glucuronidase assay of RmP193 (SMb21138::*gusA*) grown in various legume exudates with RmP32 (dme::*gusA*) as a positive control.

3.7. Plant screening

To obtain a more complete picture of the role of ABC transporters in the competition and symbiosis of *S. meliloti*, strains carrying the various gene fusions were assayed for activity *in plantar*. Alfalfa plants were inoculated with *S. meliloti* strains containing *gusA* reporter gene fusions and ß-glucuronidase activities were determined in extracts from nodules (see methods). The concentration of protein in each extract sample was determined and the ß-glucuronidase activities were expressed as nmoles formed/min./mg of nodule protein. Values for the control *dme*::*gusA*-gene fusion were included as NAD-malic enzyme (*dme*) is known to be induced in nodules (Driscoll and Finan 1993), in addition a *nifH*:: *gusA* fusion strain *nifH* is known to be induced during symbiosis (Ruvkun et al. 1982). As seen in Figure 3-25 there are several gene fusions showing increased gusA activity in the nodule extracts. SMb20442, SMb20428, SMb21130, SMb20263, SMb02793, SMb21373, SMb20904 (putative sugar uptake ABC

transporter ATP-binding protein) were all found to have elevated GusA activity in the nodules.

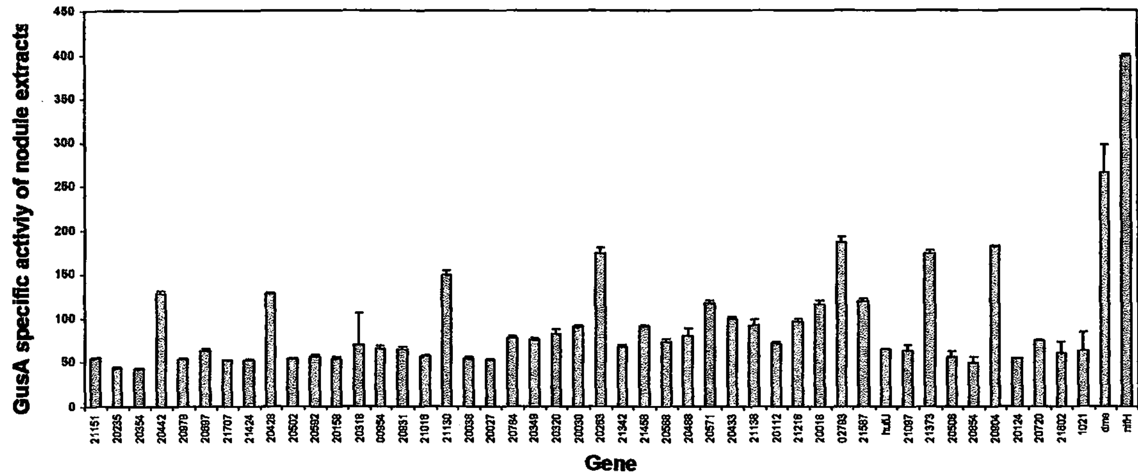

Figure 3-26. ß-glucuronidase assay of nodules extracts from alfalfa plants inoculated with *S. meliloti* strains containing GusA fusions

Figure 3-26 shows the results of ß-glucuronidase assays of all the fusion strains grown in LBmc media are shown as a comparison to the ß-glucuronidase activity obtained from the nodule extracts, *nifH::gusA* excepted. The order the strains appear in Figures 3-25 and 3-26 are the same.

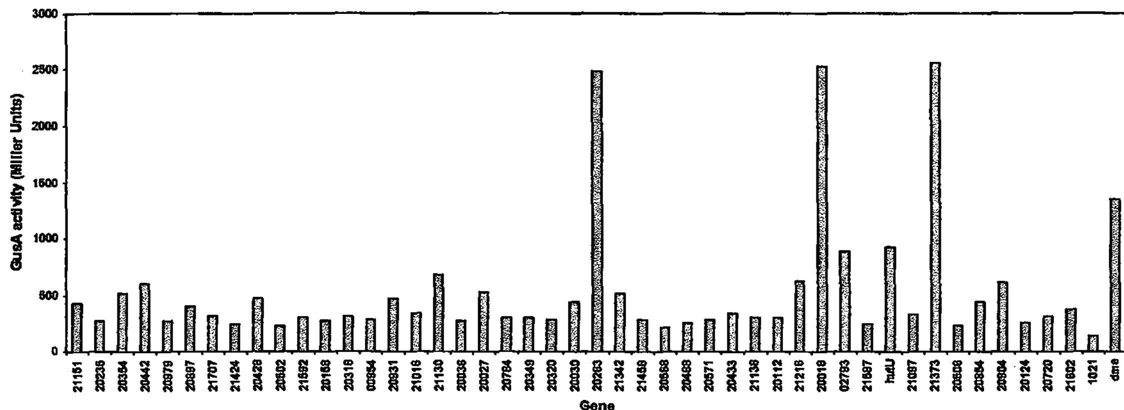

Figure 3-27. ß-glucuronidase assay *gusA* fusion strains grown in LBmc media

CHAPTER 4. GALACTITOL AND TAGATOSE

4.1. Introduction

Galactitol (formerly named dulcitol) is a sugar alcohol found in the plant families Celastraceae and Rosaceae and therefore is a potential soil carbon source that *S. meliloti* would encounter in its natural environment (Plouvier 1972). D-tagatose is a sugar generating much interest as an artificial sweetener due to its low caloric content (less than half that of sucrose) and because it does not cause cavities. The galactitol (*gat*) pathway has been extensively studied in *E. coli* and other enteric bacteria. In this system galactitol is taken up by the cell via a phosphotransferase transport system (PTS) in which the sugar is phosphorylated upon entry to the cell and further metabolized by the *gat* genes (Nobelmann and Lengeler 1996). In various strains of *E. coli* and other enteric bacteria the catabolic enzymes for metabolism of galactitol, tagatose, D-galactosamine and N-acetyl-galactosamine are very similar or even identical and share several intermediates (Shakeri-Garakani et al. 2004). In *E. coli*, *gatABCYZ* were found to among the most highly transcribed genes in M9 minimal media with glucose as the sole carbon source but this is in a strain of *E. coli* in which a mutation in *gatR* results in a truncated and non-functional protein (Corbin et al. 2003). This high level of expression is consistent with the results obtained for *S. meliloti* where the gene cluster very highly induced compared to the positive control *dme::gusA* and other test fusions (Figure 4-1).

The pSymB deletion mutants were tested on various carbon sources and it was found that RmG470 (a deletion spanning nucleotides 889262 – 1177742 on pSymB) was unable to grow on galactitol or tagatose as the sole carbon sources (see Chapter 3.1,

Table 3-1). GusA reporter fusions using pTH1360 were made to two genes in a putative ABC transport operon and tested under various growth conditions. RmP187 (Smb21373::*gusA*) and RmP195 (SMb21375::*gusA*) were found to be induced 5-fold in M9 minimal media with galactose as the sole carbon source as compared to M9 minimal media containing other sugars as the sole carbon sources (Figures 4-1).

4.2. Results

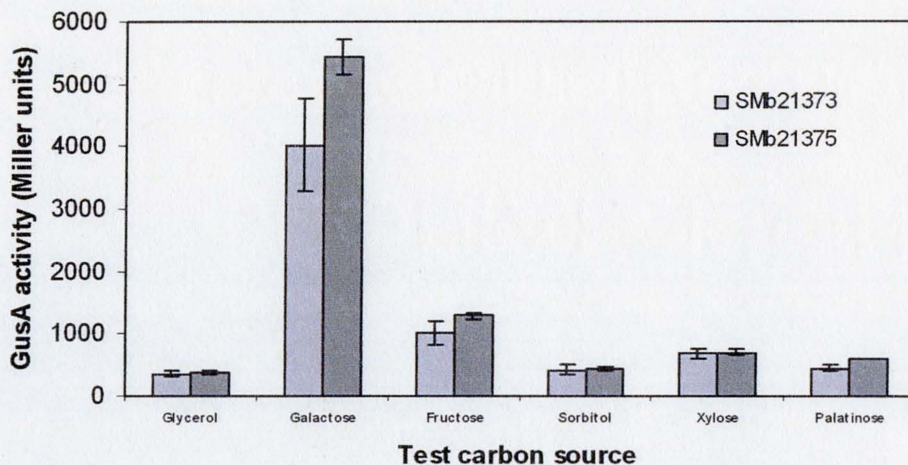

Figure 4-1. ß-glucuronidase assay of RmP187 (SMb21373::*gusA*) and RmP195 (SMb21375::*gusA*) in M9 minimal media with galactose as the sole carbon source.

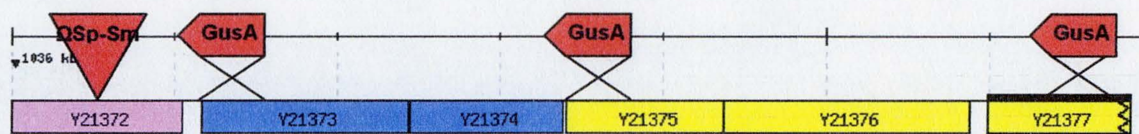

Figure 4-2. Gene map of the operon shown to be induced by galactose, galactitol, tagatose and sorbose and location of *gusA* fusions and ΩSp-Sm cassette insertion in SMb21372 (http://bioinfo.genopole-toulouse.prd.fr/annotation/iANT/bacteria/rhime/).

In a separate experiment galactitol and tagatose were also found to be inducers of this operon as seen in the Figure 4-3 below.

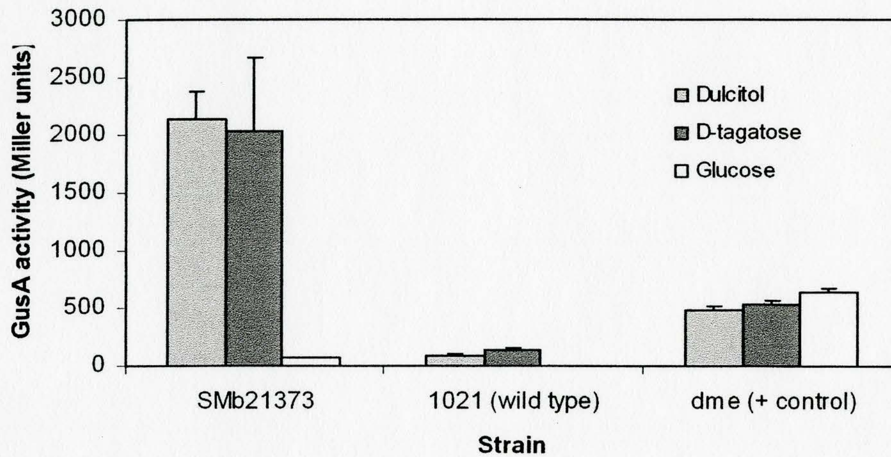

Figure 4-3. ß-glucuronidase assay of RmP187 (SMb21373::*gusA*) and Rm1021 as a negative control and RmP32 (*dme*::*gusA*) as a positive control in M9 minimal media with galactitol and d-tagatose.

The high-throughput screening revealed that this operon was also induced by sorbose and this observation is confirmed below in a separate experiment (Figure 4-5).

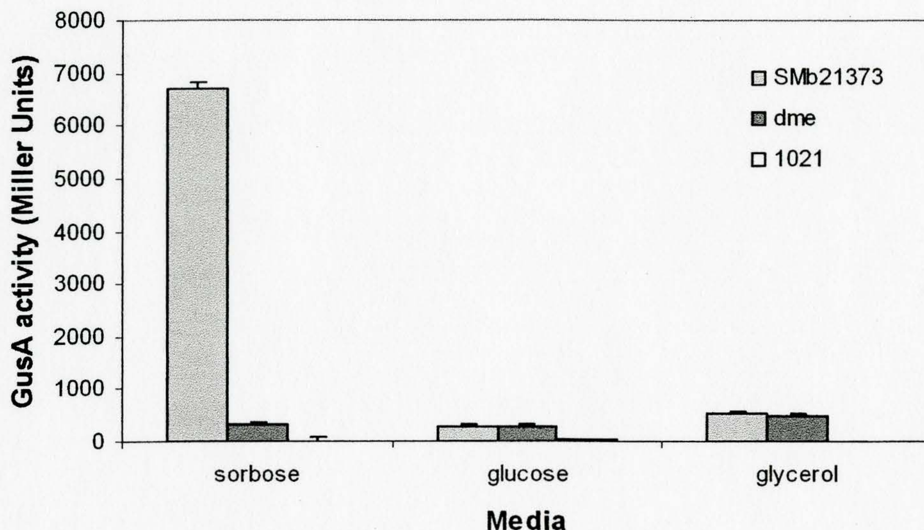

Figure 4-5. ß-glucuronidase assay of RmP187 (SMb21373::*gusA*) with Rm1021 as a negative control and RmP32 (*dme*::*gusA*) as a positive control in M9 minimal media with sorbose as the sole carbon source.

The promoter region of the galactitol operon, the intergenic region upstream of SMb21377 was cloned into a replicating plasmid, pTH1582 that contains a promoterless *gusA* gene and mated into Sm1021 to create the strain RmP244. A ß-glucuronidase assay also demonstrated induction for this strain when grown in galactose, galactitol, tagatose and sorbose (Figure 4-6). The controls used in this experiment were a positive control RmK1001 (*katA*::*gusA*, *phoC-)* and a negative control RmP62 an empty vector both of which were in a Sm1021 background.

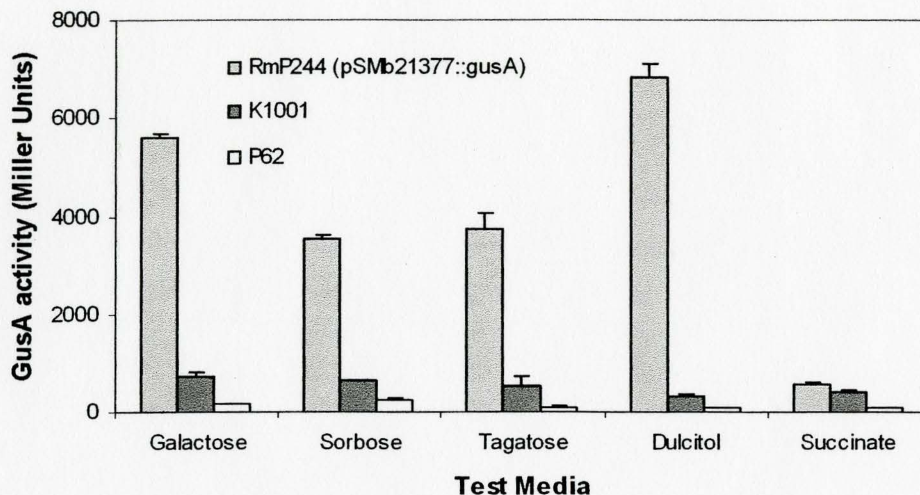

Figure 4-6. ß-Glucuronidase assay RmP244 ($P_{SMb21377::gusA}$) grown in minimal media with dulcitol, sorbose and tagatose.

The possible pathway for galactose metabolism via the intermediates galactitol and tagatose is depicted below when a PTS transporter such as *gatABC* is not the mode of transport into the cell. D-tagatose 1,6- bisphosphate is converted into glycerone phosphate and D-glyceraldehyde 3-phosphate, which then both proceed into glycolysis.

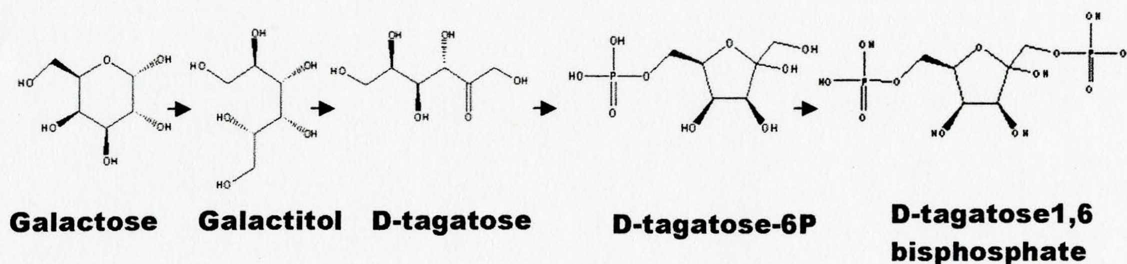

Figure 4-7. Galactose pathway via galactitol and D-tagatose (Kyoto Encyclopaedia of Genes and Genomes (KEGG) website: http://www.genome.jp/kegg/).

Table 4-1 lists the results of growth on M9 minimal media containing various carbon sources using the strains Rm1021, RmG462 (a large deletion spanning the entire

galactitol cluster), RmP233 (SMb21377::*gusA*) a functional knock out of the entire operon and RmP195 (SMb21375::*gusA*), a functional knock-out of the metabolism genes in the operon (Figure 4-3). The growth results indicate that while galactitol and tagatose are likely to be catabolized via the pathway outlined in Figure 4-7, there is a separate pathway for the catabolism of galactose.

Table 4-1. Results of growth on M9 minimal media agar plates containing various carbon sources (all carbon sources present at a concentration of 10 mM).

Strain	Sorbose	Galactose	Galactitol	Tagatose	Glucose (+ cont.)	No C (- cont.)
1021	+	+	+	+	+	–
G462	+	+	–	–	+	–
SMb21377	+	+	–	–	+	–
SMb12375	+	+	–	–	+	–
SMb21373	+	+	–	–	+	–

– = No visible growth
+ = Normal growth (equal to that of Sm1021)

A putative regulator (SMb21372) from the LacI family lies directly downstream of the galactitol gene cluster. To study the connection of this regulator with the expression of the galactitol gene cluster a knock-out mutant was generated. Almost the entire SMb21372 ORF (base 10 – 931) was PCR amplified and cloned into the suicide vector pUCP30T (gentamicin resistant). A ΩSp-Sm resistance cassette from pHP45 (Prentki and Krisch 1984a) was PCR amplified with primers containing *SalI* restriction sites on the ends and cloned into 2 natural *SalI* sites that occur at 466 and 535 bases in SMb21372. This plasmid was mated into RmP233 (SMb21373::*gusA*) and recombinants were identified as Sp^R and Gm^S colonies. When grown in galactitol, RmP233 (SMb21377::*gusA*) and RmP245 (SMb21377::*gusA*, SMb21372ΩSp-Sm) both had the

same level of GusA expression. When grown in succinate RmP245 had a three fold increase relative to growth in galactitol.

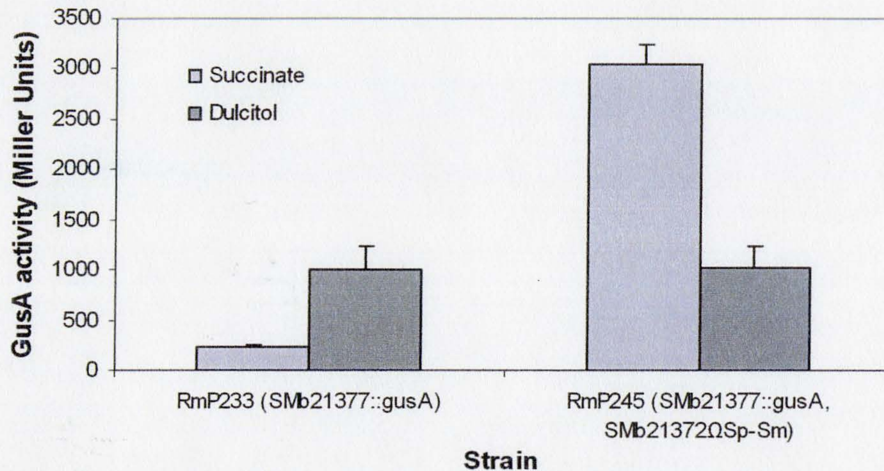

Figure 4-8. ß-Glucuronidase assay of SMb21377::*gusA* with (RmP245) and without (RmP233) a Sp-Sm cassette interrupting SMb21372, a putative regulator grown in M9 minimal media containing galactitol and succinate.

An interesting aspect of galactitol transport and metabolism that has arisen is the presence of reversion mutants. When the galactitol knock-out strains (SMb21377 and G470) are grown on galactitol or tagatose as the sole carbon source there is no growth but after approximately one and a half weeks colonies begin to form which would indicate putative reversion mutants. A frequency of reversion was calculated by comparing the number of reversion mutants on M9 minimal media with galactitol to those that grew on succinate as the sole carbon source. These putative revertants were not characterised further.

Table 4-2. Frequency of reversion mutants of galactitol mutants

Strain	**Frequency of reversion**
G470	3.047 x 10-7
RmP233 (SMb21377::gusA)	2.095 x 10-7

CHAPTER 5. HYDROXYPROLINE

5.1. Introduction

Proline metabolism has been widely studied in enteric and various soil bacteria. In prokaryotes proline dehydrogenase and pyrroline-5-carboxylate dehydrogenase are usually encoded by the same protein (PutA) as was found in *S. meliloti* (Jimenez-Zurdo et al. 1997). Often divergently transcribed away from *putA* is the *putP* gene that encodes for the proline permease which is responsible for transporting proline. Both of these genes are often regulated by PutA (Gu et al. 2004b). In *S. meliloti putA* is located on the chromosome but no proline permease gene has yet been identified. PutA has been shown to be important in the competition of *S. meliloti* in the field and on nodulation efficiency. (Jimenez-Zurdo et al. 1995).

Hydroxyproline is a constituent of collagen and some other mammalian proteins. This imino acid is a proline residue that is modified post-translationally and is therefore not one of the essential amino acids. The hydroxyproline from protein has a configuration where the hydroxyl group and carboxyl are in *trans*. Hydroxyproline aids in stabilizing the triple helix structure of collagen (Fruton JS and Simmonds S 1959). Collagen provides structure to mammalian bodies, protecting and supporting the softer tissues. The collagen peptide most often follows the pattern; Glycine – Proline – Hydroxyproline, although other residues may be present. Ascorbic acid is a co-factor of the mammalian enzyme responsible for the hydroxylation of proline, prolyl hydroxylase (Adams and Frank 1980). A deficiency in Vitamin C slows the production of hydroxyproline and stops the construction of new collagen resulting in scurvy. The symptoms of scurvy (loss

of teeth, easy bruising and lesions) are caused by the lack of collagen in the body (Adams and Frank 1980).

Hydroxyproline is also found in some higher plant glycoproteins. Hydroxyproline arabinosides are present in all photosynthetic plants but hydroxyproline has also been shown to be linked to galactose. Plant proteins containing hydroxyproline have been termed extensins which are a diverse family of cell wall proteins (Adams and Frank 1980). There is a subfamily of glycoproteins, rich in hydroxyproline which has only been found in legumes and are unregulated in nodules relative to root tissues. These matrix glycoproteins have also been localized to the lumen of infection threads and have been termed "root nodule extensions" (Rathbun et al. 2002).

Although extensively studied in mammalian and plant cells there has been very little focus on hydroxyproline in bacteria. The hydroxylation of proline in mammalian cells is carried out by prolyl hydroxylase and this enzyme has not been found in bacteria except in *Streptomyces griseoviridus* (Adams and Frank 1980). Cell extracts of *S. griseoviridus* were shown to hydroxylate L-proline to L-hydroxyproline in a reaction that needs ascorbic acid, iron and α-ketoglutarate. These are the same co-factors needed by prolyl hydroxylase in mammalian cells (Onishi et al. 1984). In *S. griseoviridus* free hydroxyproline is an intermediate synthesized from proline during formation of the allohydroxy-D-proline residue in the antibiotic etamycin using the enzyme proline 4-hydroxylase (Katz et al. 1979). An environmentally isolated strain, *Streptomyces sp.* TH1 was found to contain the enzyme proline 3-hydrolylase (Mori et al. 1997). Figure 5-1 shows the reaction of proline to hydroxyproline catalyzed by prolyl hydroxylase.

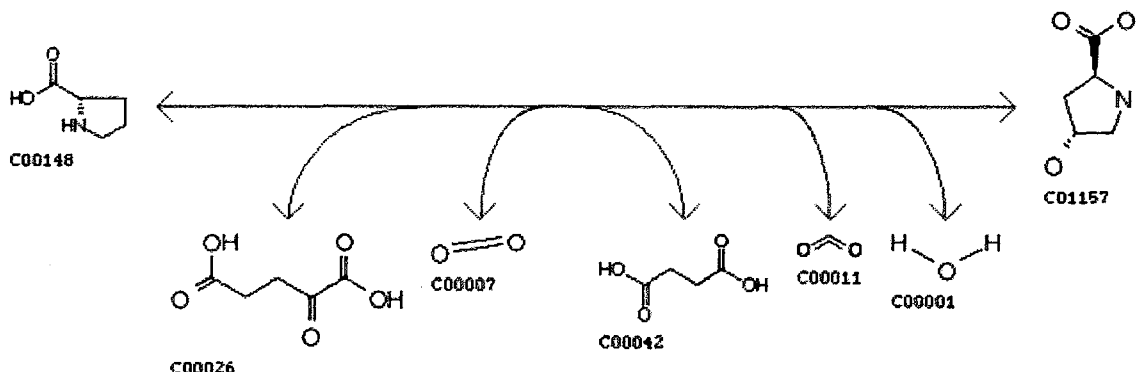

Figure 5-1. Reaction of proline to hydroxyproline catalyzed by prolyl 4-hydroxylase. (Kyoto Encyclopedia of Genes and Genomes (KEGG) website: http://www.genome.jp/kegg/)

Using the results of a BLASTp search it was observed that the only other bacteria containing a similar enzyme to that of *Streptomyces. sp.* TH1 are *S. meliloti* and *Mesorhizobium loti* MAFF303099 with E values of 4e-34 and 2e-35 respectively (results from a NCBI BLASTp query). In S. *meliloti* a gene annotated as L-proline 3-hydroxylase was shown to be FixLJ regulated (SMc03253), as well as ProB2 a glutamate-5-kinase involved in proline biosynthesis. These two genes were determined to not be necessary for symbiosis (Ferrieres et al. 2004). Direct experimental evidence that the SMc03253 protein catalyzes the reaction remains to be carried out.

In depth genetic characterization of the genes required for hydroxyproline metabolism has not been reported. The only genetic study was carried out in *Pseudomonas aeruginosa* POA in which the hydroxyproline degradative pathway was found to be clustered together on the chromosome at around 19 minutes. By using transductional linkage using selectable markers and interrupted mating experiments it

was possible to position the *hyp* genes relative to other amino acid degradative genes (Manoharan and Jayaraman 1979).

Although this metabolic pathway has not been characterized genetically, much about the biochemistry is known due to extensive study in *Pseudomonas putida*. Elijah Adams and his co-workers have identified the catabolic pathway and characterized the biochemical properties of its enzymes in the bacterium *P. putida*.

Bacteria able to catabolize hydroxyproline had not been discovered until a strain of *Pseudomonas putida* isolated from soil was found in the 1950's (Adams 1959). Although both the bacterial and mammalian catabolic pathways convert hydroxyproline to the final product of glutamate, these two metabolic pathways differ in many ways. The first step in the conversion of hydroxyproline is an epimerization reaction where L-hydroxyproline (the most abundant form found in nature) is converted to D-allohydroxyproline by hydroxyproline-2-epimerase. Although all four isomers of hydroxyproline could be oxidized by cells, L-hydroxyproline or D-allohydroxyproline had to be present for the other two forms to be utilized as these were the only two forms that had inducing ability of the enzymes needed for metabolism (Adams 1959). Only after the epimerization of L-hydroxyproline to D-allohydroxyproline could the further metabolism of hydroxyproline be carried out. This was shown in a mutant defective for the epimerase enzyme. The mutant was unable to grow on L-hydroxyproline but could still grow on D-allohydroxyproline as the sole carbon and nitrogen source. This demonstrated that the D epimer was needed before oxidation could occur (Gryder and Adams 1969). Allohydroxy-D-proline oxidase carries out a dehydrogenation reaction

similar to other D-amino acid oxidases in which Δ^1-pyrroline-4-hydroxy-2-carboxylic acid is produced. The next step in the pathway, catalyzed by Δ^1-pyrroline-4-hydroxy-2-carboxylate deaminase was a non-oxidative ring-opening reaction and deamination which was hypothesised to be similar to that of serine dehydrase to produce α-ketoglutarate semialdehyde (KGSA) (Singh and Adams 1965). KGSA dehydrogenase was hypothesized to be the enzyme catalyzing the final step to α-ketoglutarate. When trying to purify this enzyme it became apparent that there was another KGSA dehydrogenase with a different molecular weight, electrophoretic behavior and substrate specificity which was giving a basal level of KGSA activity. This enzyme was induced by lysine and is the first enzyme in the catabolism of L-lysine. It was shown in other *Pseudomonas* strains that glucarate, galactarate and arabonate all induced a KGSA dehydrogenase and the stain of *Pseudomonas putida* being used in the hydroxyproline studies could grow on glucarate (Adams 1973). It seemed as though these two enzymes each induced by different pathways were really one common enzyme in two converging metabolic pathways as in protocatechuate and catechol metabolism. After a more in depth study of the biochemical properties of the enzyme it was clear that there were two distinct enzymes although their Km's and aldehyde substrates were the same (Koo and Adams 1974). Various compounds were investigated for catabolite repression of hydroxyproline induction including glucose, glutamate, succinate and α-ketoglutarate but there was no such effects observed. One of the two hydroxyproline mutants isolated was thought to be a regulatory mutant due the inability of hydroxyproline to induce the catabolic enzymes of the pathway and also the inability of this strain to transport hydroxyproline. In addition

to inducing the metabolic enzymes of the pathway, both epimers of hydroxyproline were shown to induce the transport system. The mutant lacking hydroxyproline-2-epimerase which causes a blockage in the first step of the pathway was used to study the hydroxyproline transport system. This mutant would transport and accumulate hydroxyproline. In the epimerase mutant, uptake of D-allohydroxyproline was inhibited by the presence of L-hydroxyproline (Gryder and Adams 1969). Further uptake experiments showed that the K_m for D-allohydroxyproline was 300 times that of L-hydroxyproline. Although many amino acids interfered with the uptake of hydroxyproline the two with the greatest effect were L-proline and L-alanine (Gryder and Adams 1970).

5.2. Results

It was first noted that genes essential to the transport and/or metabolism of hydroxyproline were present on pSymB as RmF909 a mutant strain of Sm1021 containing a large deletion spanning nucleotides 106128 - 735511 on pSymB was unable to grow on hydroxyproline (5 mM) as the sole carbon source. GusA activity in two reporter fusion strains RmP215 (Smb20263::*gusA*) and RmP217 (SMb20320::*gusA*) gene clusters were found to be induced by L-hydroxyproline after growth in M9 minimal media with hydroxyproline as the sole carbon source (Figures 5-2 and 5-3).

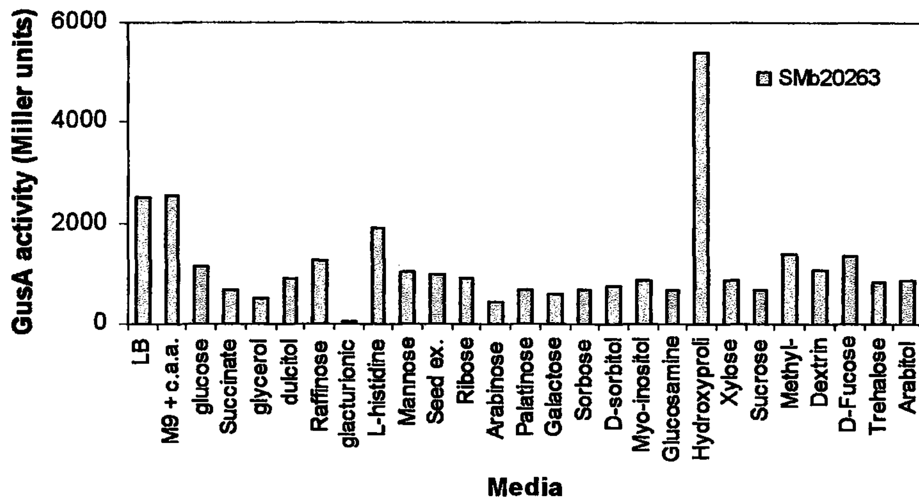

Figure 5-2. ß-Glucuronidase high-throughput assay of RmP215 (SMb20263::*gusA*) showing induction by hydroxyproline

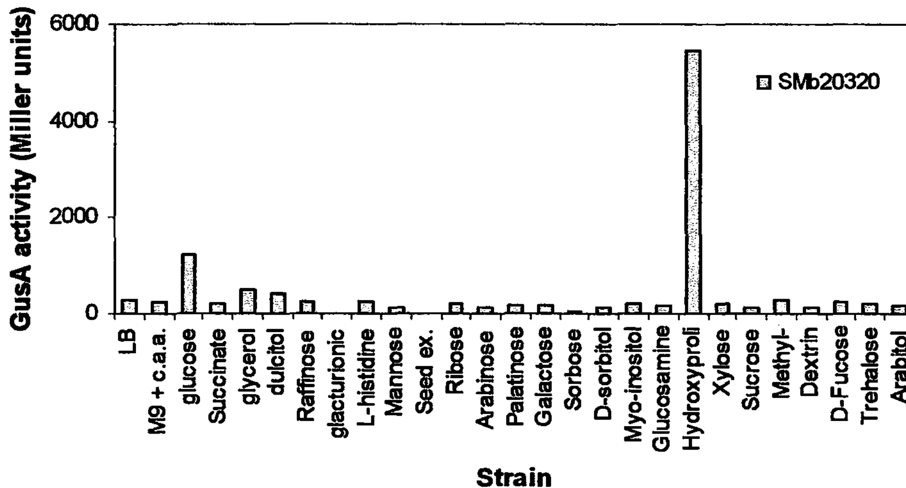

Figure 5-3. ß-Glucuronidase high-throughput assay of RmP217 (SMb20320::*gusA*) showing induction by hydroxyproline

The following three Figures show growth curves of RmP215 (SMb20263::*gusA*) and RmP238 (SMb20322::*gusA*) functional transport knock-out mutants grown in L-hydroxyproline, D-allohydroxyproline and succinate. The negative control RmF909 has the region spanning both transporters deleted and Sm1021 is used as a positive control.

The fusion strain SMb20263 and RmF909 were unable to grow with hydroxyproline or allohydroxyproline as the sole sources of carbon. SMb20322 had a longer lag phase when grown in both epimers of hydroxyproline and reached stationary phase at a lower cell density than Sm1021. When grown in succinate all of the strains had similar growth curves. In a previous growth experiment RmP215 (SMb20263::*gusA*) and RmP238 (SMb20322::*gusA*) again displayed no growth and decreased growth, respectively when grown in M9 minimal media with L-hydroxyproline as the sole carbon source.

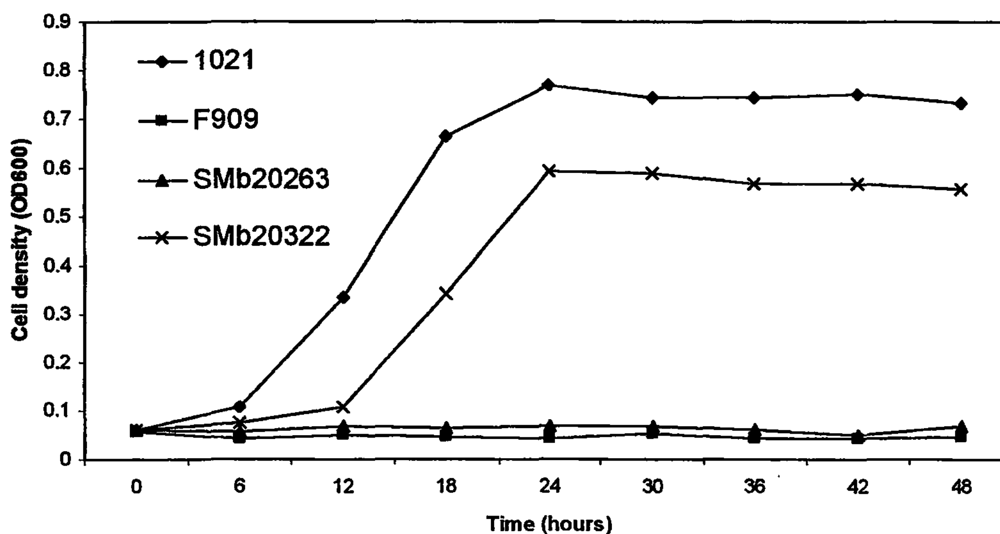

Figure 5-4. Growth curve of Sm1021, RmP215 (Smb20263::*gusA*) and RmP238 (SMb20322::*gusA* and RmF909 (a large deletion spanning the both hydroxyproline transport operons) grown in M9 minimal media with hydroxyproline (5mM) as the sole carbon source.

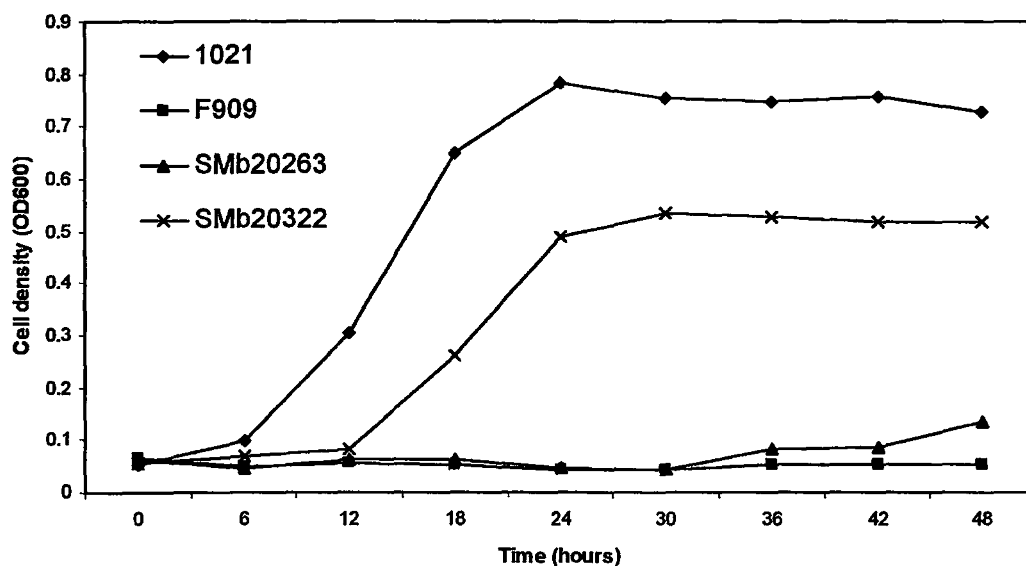

Figure 5-5. Growth curve of Sm1021, RmP215 (Smb20263::*gusA*) and RmP238 (SMb20322::*gusA* and RmF909 (a large deletion spanning the both hydroxyproline transport operons) grown in M9 minimal media with allohydroxyproline (5mM) as the sole carbon source.

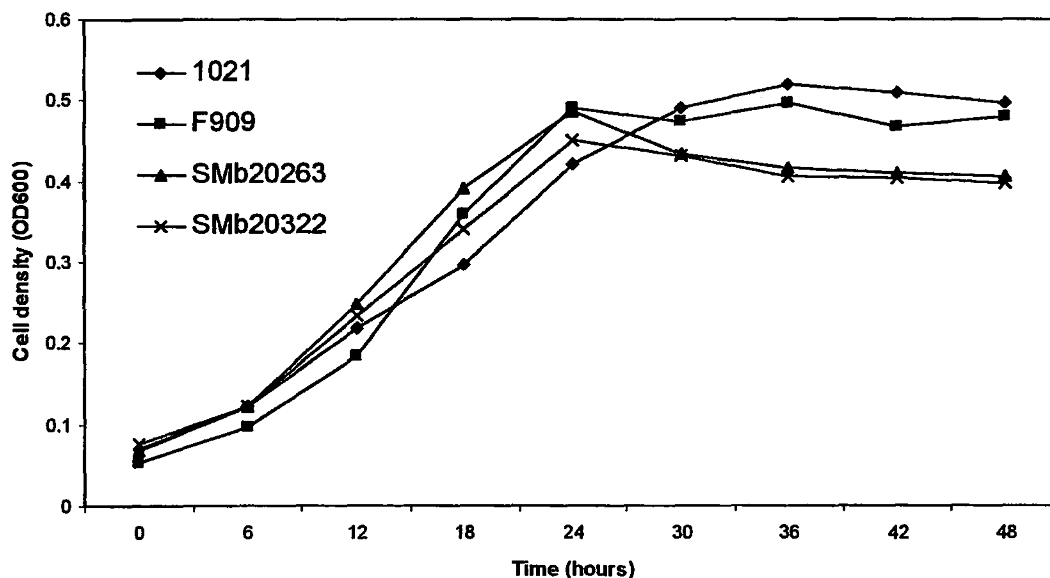

Figure 5-6. Growth curve of Sm1021, RmP215 (Smb20263::*gusA*) and RmP238 (SMb20322::*gusA* and RmF909 (a large deletion spanning the both hydroxyproline transport operons) grown in M9 minimal media with succinate (10 mM) as the sole carbon source.

RmP215 (Smb20263::*gusA*) and RmP217 (SMb20320::*gusA*) were grown in hydroxyproline or proline as the sole carbon source, the sole nitrogen source and as the carbon and nitrogen source to investigate the conditions for induction. The results of the β-glucuronidase assay were that hydroxyproline was an inducer as the sole carbon or nitrogen source or both and that proline was not an inducer of either of these operons. Glucose was also tested as a negative control and it was observed that a higher level of induction was found than in the presence of proline (Figure 5-7).

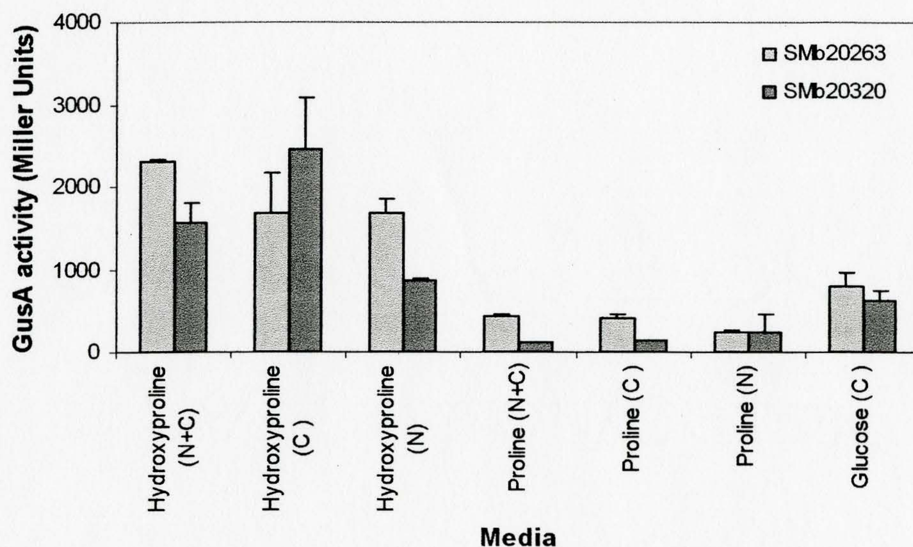

Figure 5-7. β-glucuronidase assay of RmP215 (Smb20263::*gusA*) and RmP217 (SMb20320::*gusA*) grown in M9 minimal media with hydroxyproline or proline as the sole nitrogen or carbon source or as the carbon and nitrogen source.

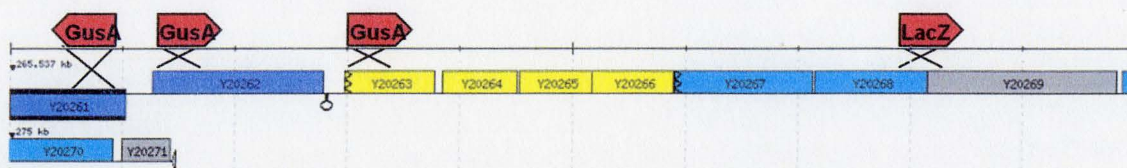

Figure 5-8. Gene map of an operon on pSymB induced by hydroxyproline and location of gene fusions (http://bioinfo.genopole-toulouse.prd.fr/annotation/iANT/bacteria/rhime/).

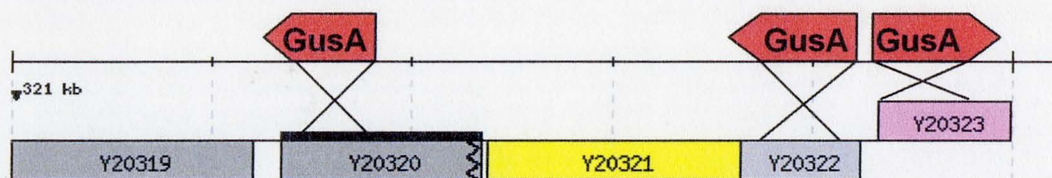

Figure 5-9. Gene map of an operon on pSymB induced by hydroxyproline and location of gene fusions (http://bioinfo.genopole-toulouse.prd.fr/annotation/iANT/bacteria/rhime/).

The SMb20263 cluster is organized with the ABC transport genes preceding 5 other genes. Upstream of the transport genes lies what is annotated as a semialdehyde dehydrogenase (SMb20262) and malate dehydrogenase (SMb20261) divergently transcribed. The promoter regions of SMb20261, SMb20262 and SMb20263 were cloned into a replicating, GusA reporter vector (pTH1582) which contains the *par* genes to confer plasmid stability. The resulting plasmids pTH1800, pTH1801 and pTH1802 were mated into Sm1021 to create the strains RmP241, RmP242 and RmP243, respectively in order to confirm the induction by hydroxyproline. The results with these constructs were the same as when the *gusA* gene was under the control of the native promoter. The results showed a low level of constitutive gene expression and an increase when grown in the presence of hydroxyproline (Figure 5-10). The controls used in this experiment were a

positive control RmK1001 (*katA*::gusA, *phoC-)* and a negative control RmP62 (an empty vector) both of which were in a Sm1021 background.

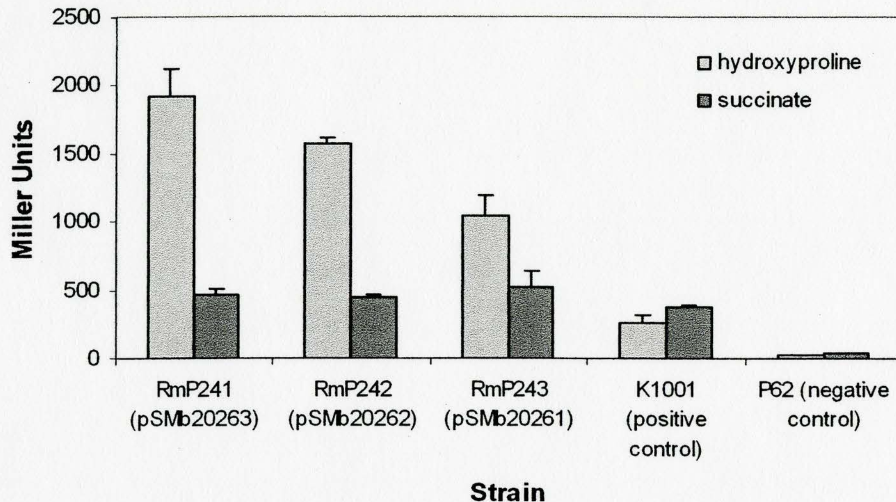

Figure 5-10. ß-Glucuronidase assay of intergenic sequences of SMb20261 (RmP243), SMb20262 (RmP242) and SMb20263 (RmP241) in a replicating GusA reporter vector grown in M9 minimal media with succinate and hydroxyproline.

A *lacZ* reporter strain RmPL43 from a random library of clones that contains the end of SMb20267 and beginning of SMb20268 in pTH1522, recombined into the genome through a single cross-over was tested for induction by hydroxyproline (Alison Cowie, unpublished). This strain had the same activity as the other clones of a basal level of induction in succinate and high induction in hydroxyproline (Figure 5-11). K991 (positive control) is the *pckA* promoter cloned into the reporter suicide vector also in the *lacZ* orientation that RmPL43 is also cloned into. *pckA* has been shown to be induced in succinate, it was also observed that *pckA* was also induced in hydroxyproline.

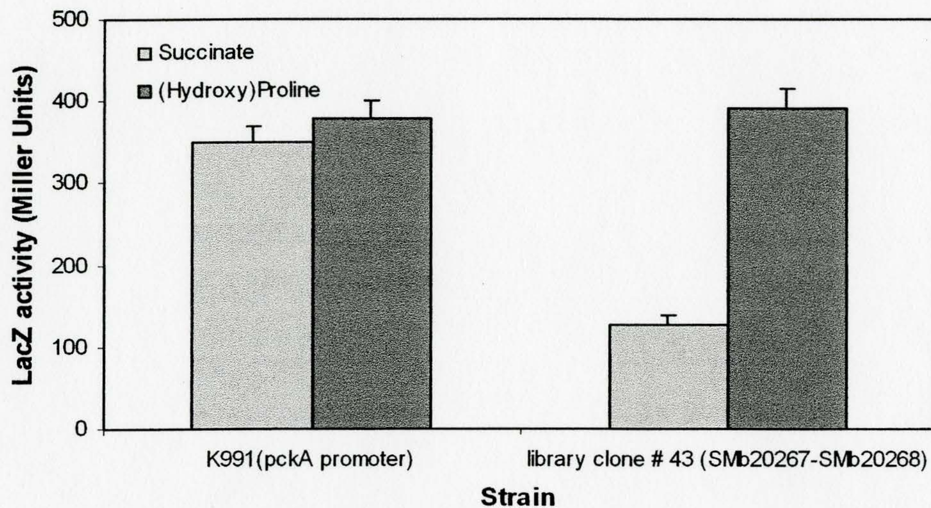

Figure 5-11. ß-galactosidase assay of library clone RmPL43 (a reporter strain of SMb20268).

Since various transport mutants were unable to grow in both epimers of hydroxyproline, allohydroxyproline was also tested as an inducer of the hydroxyproline induced genes. The strains containing a *gusA* reporter fusion to SMb20261, SMb20262, SMb20263 and SMb20322 (RmP239, RmP240, RmP215 and RmP238, respectively) were grown in M9 minimal media containing allohydroxyproline and succinate as the sole carbon source and tested for ß-glucuronidase activity (Figure 5-12). The results showed that allohydroxyproline also demonstrated inducer activity on all the hydroxyproline induced genes and that all the genes except for SMb20322 had a basal level of expression in the absence of either epimer of hydroxyproline, as seen in the previous graphs. The pattern of expression seen in Figure 5-12 has been repeated in different experiments. When grown in a variety of conditions SMb20261, SMb20262 and SMb20263 appear to have a low level of constitutive expression until induced by hydroxyproline or allohydroxyproline when SMb20261 is only slightly increase by two-

fold, SMb20262 three-fold and SMb20263 about four or five fold. As in Figure 5-12 SMb20322 displays the highest level of expression when induced by either hydroxyproline or allohydroxyproline but does not have the constant constitutive level that the other hydroxyproline induced genes show. As seen in Figure 5-2 there is a constant high level of gene expression of SMb20263 in almost all the various media.

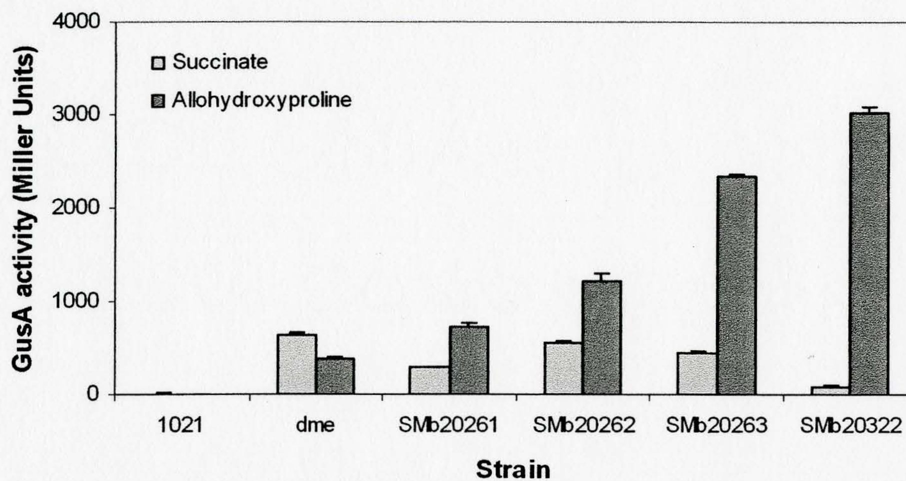

Figure 5-12. ß-glucuronidase assay of RmP239 (SMb20261::*gusA*), RmP240 (SMb20262::*gusA*), RmP215 (SMb20263::*gusA*) and RmP238 (SMb20322::*gusA*) grown in M9 minimal media with allohydroxyproline and succinate as the carbon sources.

SMb20262 (containing proline dehydrogenase and delta-1-pyrroline-5-carboxylate dehydrogenase domains) was tested as a possible regulator of the hydroxyproline induced genes to based on previous studies of proline metabolism in which the bifunctional *putA* gene encoding proline dehydrogenase and delta-1-pyrroline-5-carboxylate dehydrogenase was found to regulate proline metabolism and transport genes (Gu et al. 2004a). A knock-out of SMb20262 was made by recombination of a suicide plasmid pUCP30T within the ORF of the gene. The replicating plasmids

containing the promoter fragments of SMb20261, SMb20262 and SMb20263 cloned in front of a promoterless *gusA* and were mated into this SMb20262 knock-out strain giving strains RmP311, RmP312 and RmP313. ß-glucuronidase activity was assayed in order to determine the regulatory effect of SMb20262 on any of the hydroxyproline induced promoters. As seen in the figure below the expression in a wild type (Sm1021) background were identical to that in the SMb20262 knock-out background.

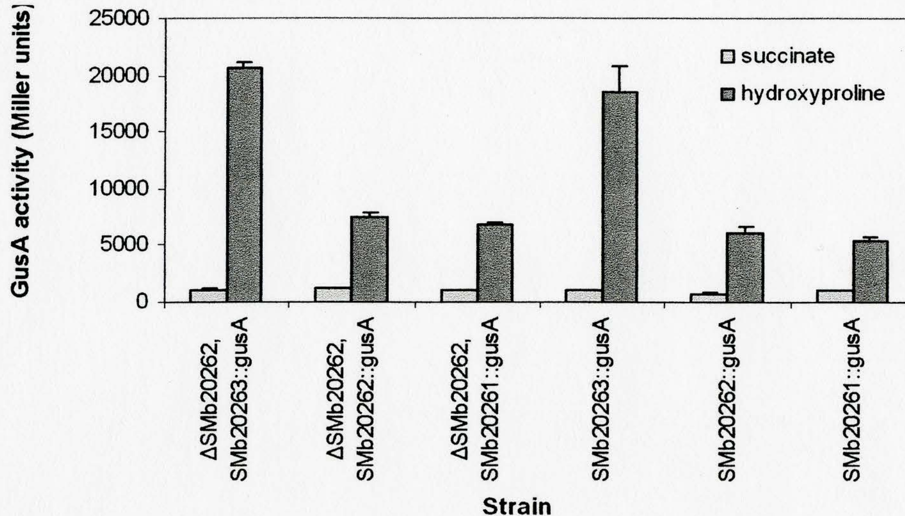

Figure 5-13. ß-glucuronidase assay of the promoter regions of SMb20261, SMb20262 and SMb20263 in a SMb20262 functional knock-out background (RmP311, RmP312 and RmP313) versus a wild-type background (RmP241, RmP242, RmP243).

In some organisms proline transport and metabolism is *ntrC* regulated, although not the case with *S. meliloti* (Suhr and Kleiner 1993). In order to investigate a connection between *ntrC* and the hydroxyproline induced genes RmP242, strains containing the promoter sequence of SMb20262 in a replicating *gusA* reporter vector in a Sm1021 background and RmP315 containing the same plasmid in a *ntrC* – background (RmF932) were tested. These two strains were grown using hydroxyproline as the carbon source,

and succinate as the carbon source with and without a nitrogen source (NH_4Cl). The results show no connection between SMb20262 and nitrogen limiting conditions or *ntrC* regulation.

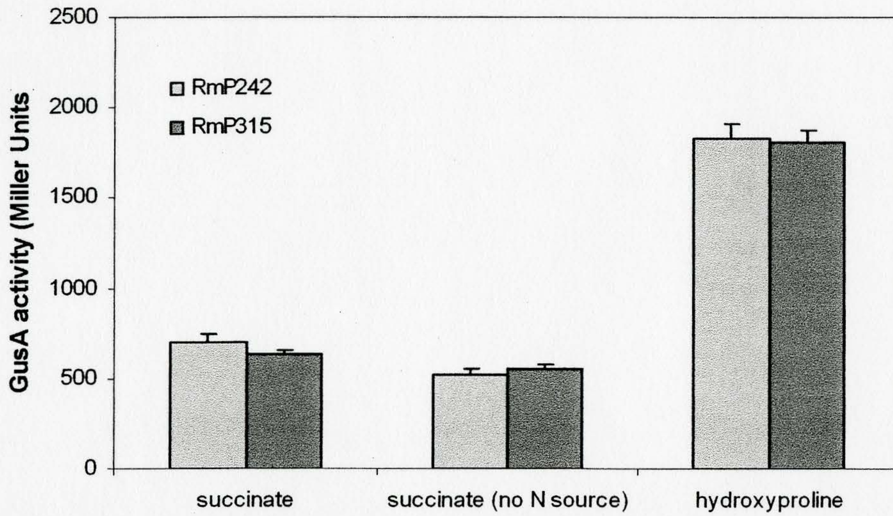

Figure 5-14. ß-glucuronidase assay of the promoter region of SMb20262 in a replicating gusA reporter plasmid (pTH1801) in a wild type background (RmP242) and a ntrC- background (RmP315) grown in M9 minimal media with hydroxyproline as the carbon source and succinate as the carbon source with and without nitrogen (NH_4Cl).

Catabolite repression of transport and metabolism genes by glucose and succinate is a common occurrence in gram negative bacteria such as *E. coli* and *S. meliloti* (Gage and Long 1998). These two carbon sources were tested for possible catabolite repression of the hydroxyproline induced genes. RmP215 (SMb20263::gusA) and RmP238 (SMb20322::*gusA*) were tested in glucose, succinate, hydroxyproline and also hydroxyproline with glucose or succinate as the carbon source. There were no effects on the gene expression of the SMb20263 operon but the presence of glucose and succinate decreased the gene expression of the SMb20322 cluster by about half.

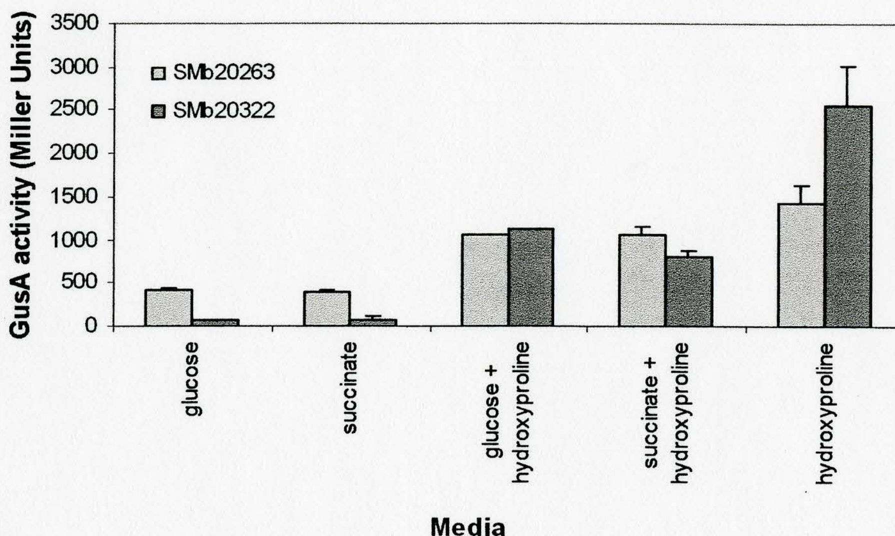

Figure 5-15. β-glucuronidase assay of RmP215 (SMb20263::*gusA*) and RmP238 (SMb20322::*gusA*) grown in M9 minimal media containing hydroxyproline and glucose or succinate to investigate the catabolic repression of these two compound on hydroxyproline induction

It was found during the high-throughput screening with seed and root exudates that RmP215 (SMb20263::gusA) was induced approximately two fold by alfalfa root exudates.

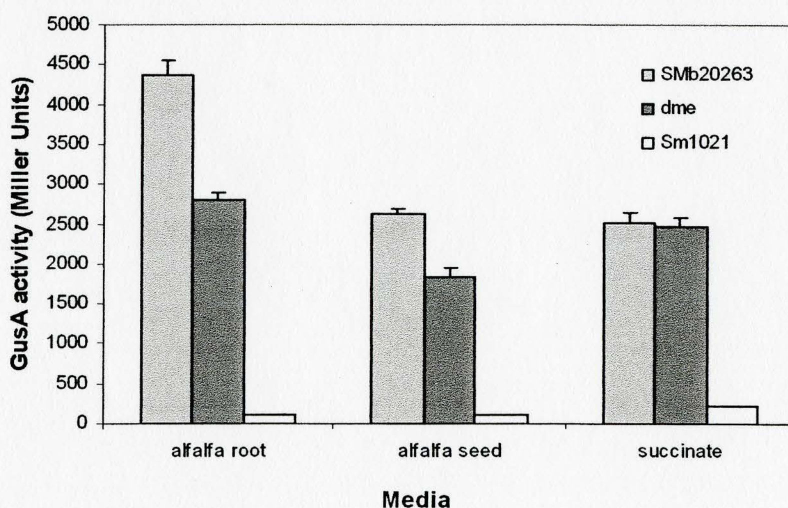

Figure 5-16. β-glucuronidase assay of RmP215 (SMb20263::gusA) assay grown in alfalfa root and seed exudates.

In order to determine the structure of the targeted gene region in the resulting recombinants for the hydroxyproline induced stains, total DNA from RmP239 (SMb20261::*gusA*) (Figure 5-17), RmP240 (SMb20262::*gusA*) (Figure 5-18) and RmP215 (SMb20263::*gusA*) (Figure 5-19) was analysed by southern hybridization. In each case the respective PCR product used to originally clone the gene into pTH1360 was used as the probe. Correct band sizes were obtained for the three strains tested, indicating that only one copy of pTH1360 had recombined. The restriction enzymes sites for the various southern hybridizations vary and are indicated in each Figure. As seen in Figure 5-17 only the restriction map for the *EcoRV* digest is shown due to a problem in the *BamHI* digest.

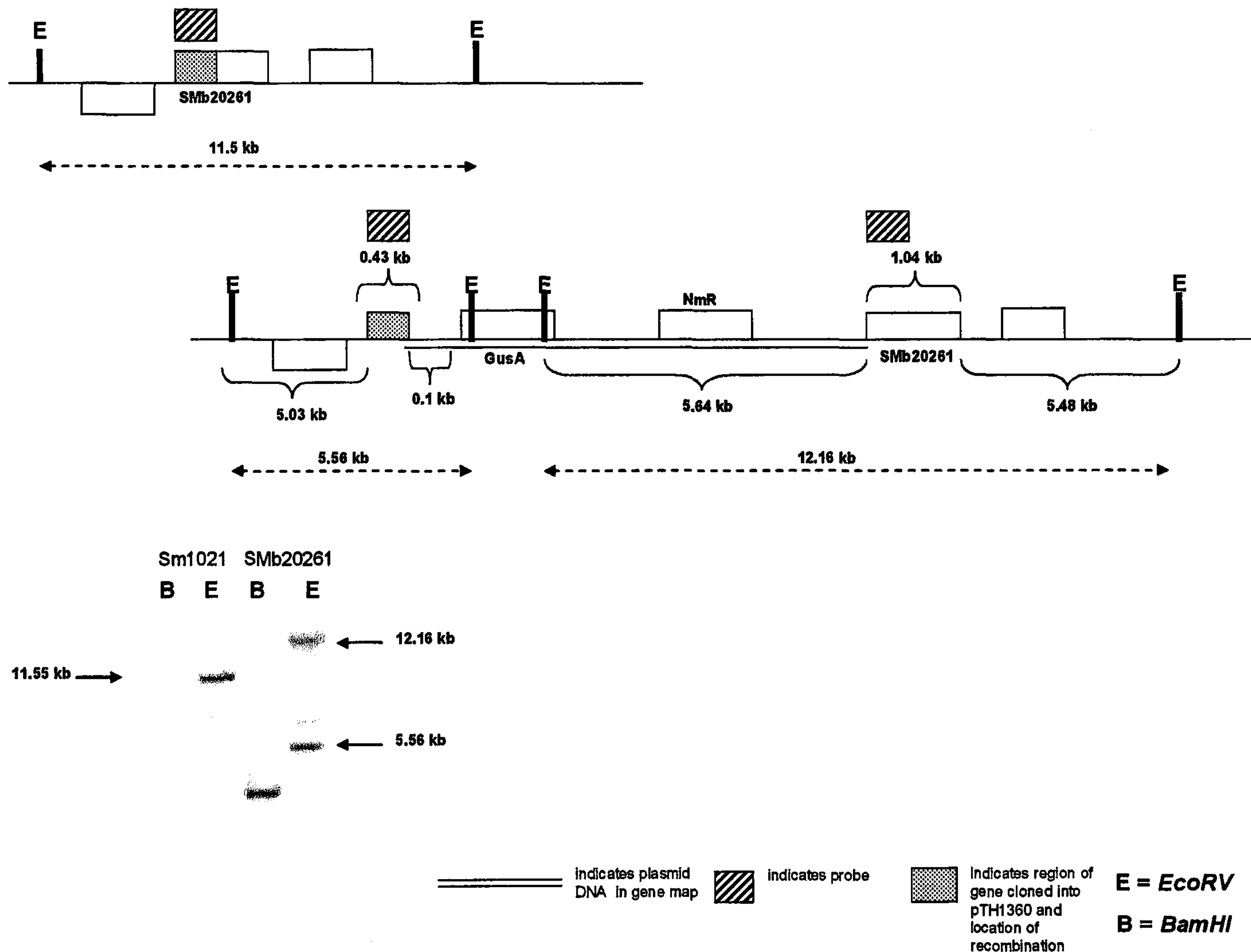

Figure 5-17. Predicted restriction maps of target regions after recombination of the plasmids used in the construction of strain RmP239 (SMb20261::*gusA*) for an *EcoRV* digest.

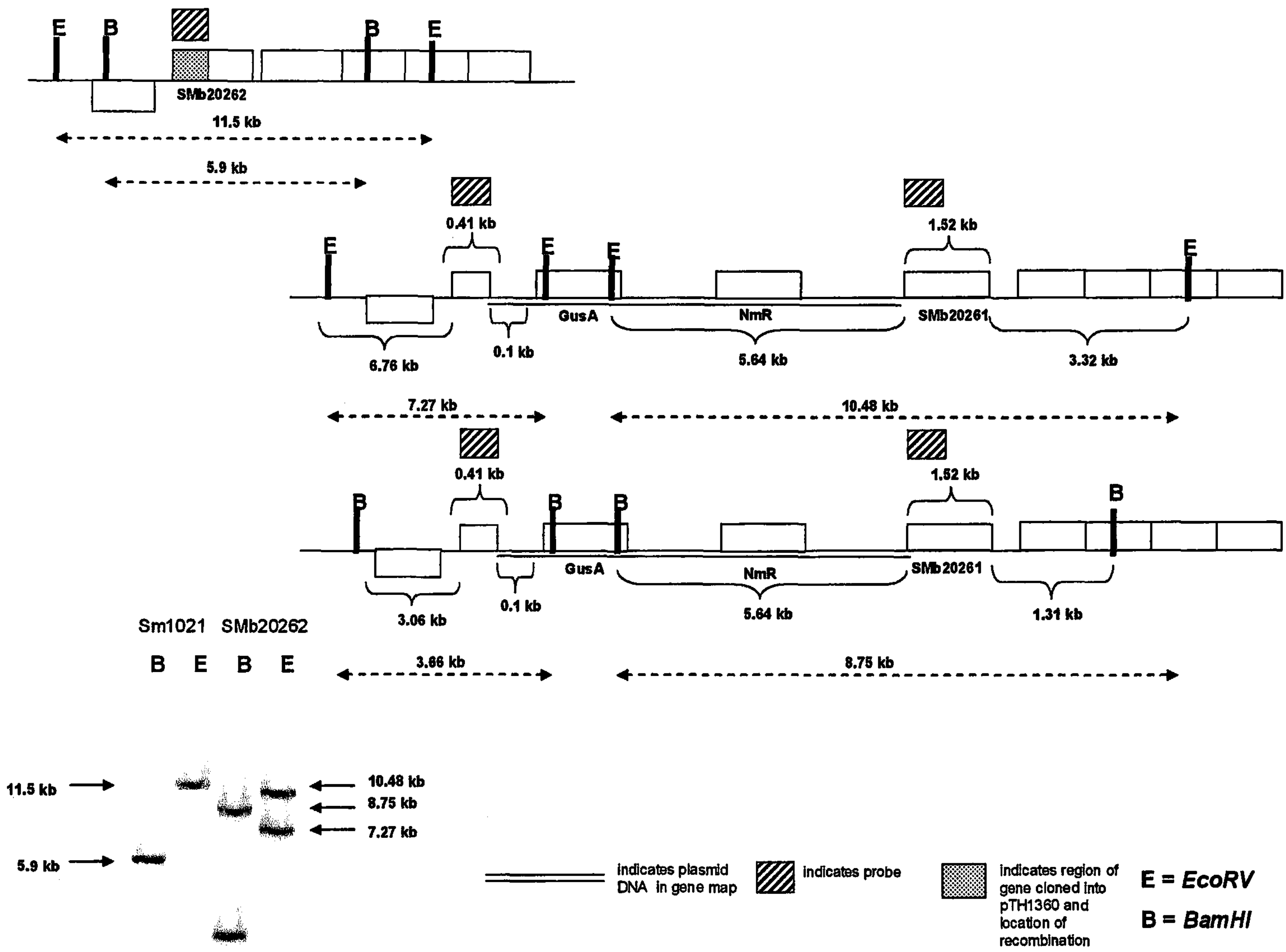

Figure 5-18. Predicted restriction maps of target regions after recombination of the plasmids used in the construction of strain RmP240 (SMb20262::*gusA*) for an *EcoRV* digest and *BamHI* digest.

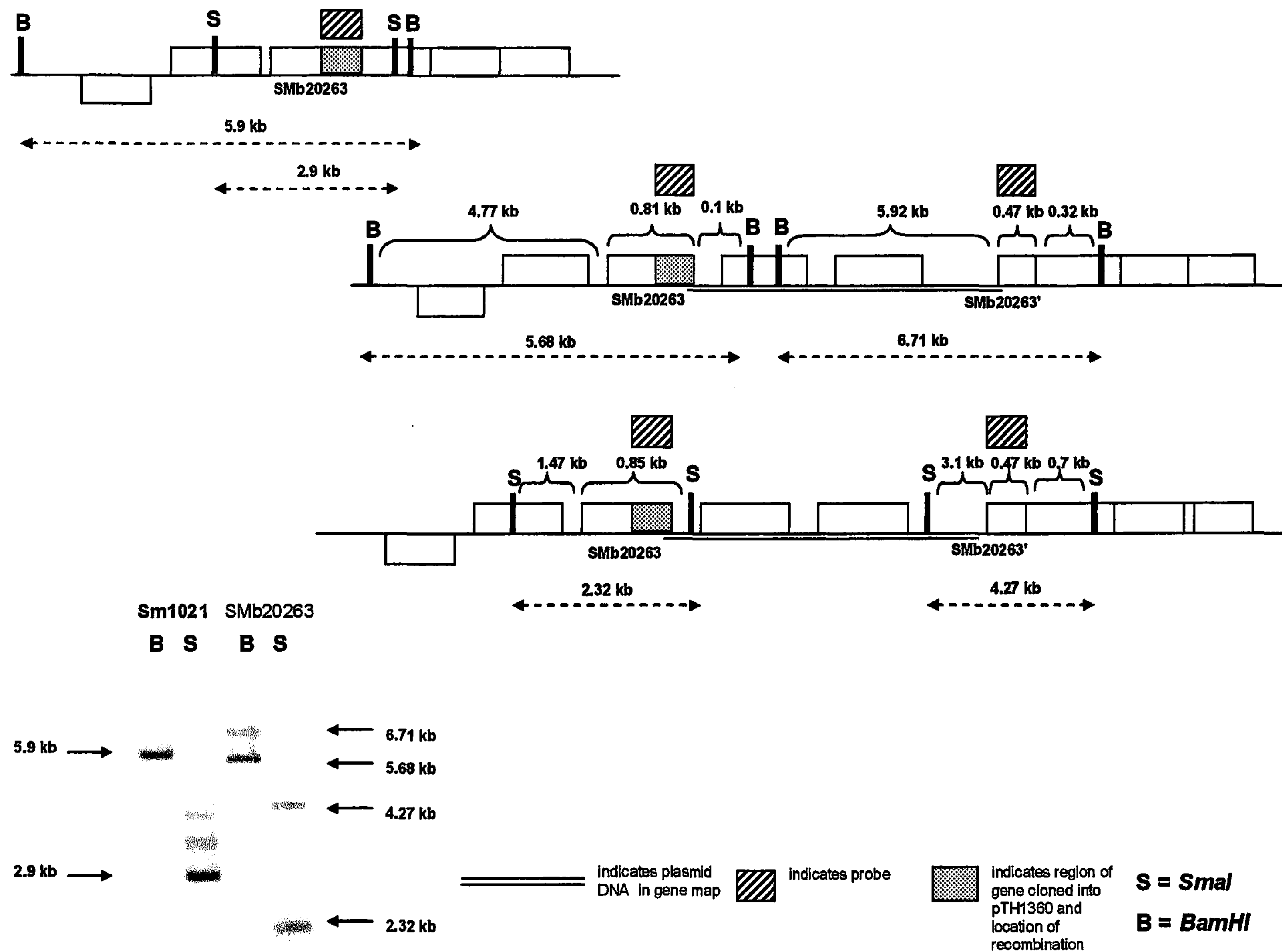

Figure 5-19. Predicted restriction maps of target regions after recombination of the plasmids used in the construction of strain RmP215 (SMb20263::*gusA*) for a *BamHI and SmaI* digest.

CHAPTER 6 - DISCUSSION

6.1. Biolog and deletion mutant screening

The results of the Biolog PMs screening were useful for gathering information on the diversity of compounds utilized by the wild type strain Sm1021 and also for noting the differences between the previously created deletion mutants and the wild type to indicate where on pSymB the genes necessary for the catabolism of specific compound lie (Charles and Finan 1991). In order to confirm the Biolog PMs results, the deletion mutants were screened on M9 minimal media containing some of the substrates found in the Biolog PMs as the sole carbon sources. In all cases the results matched up with the Biolog PMs data. Some of these results could be correlated to data previously known about *S. meliloti*, for example the deletion strain RmG373 (missing nucleotides 1452882 – 63194) was unable to utilize succinic acid, fumaric acid and malic acid, as indicated by screening on M9 minimal media containing these carbon sources and by the Biolog PMs. Previously Tn5 mutants with insertions falling in this area were shown to be unable to grow on these three compounds as the sole carbon sources due to abolishment of the C4-dicarboxylate transport system (*dctABD*) (Finan et al. 1981). The deletion mutant RmG373 (missing nucleotides 1452882 – 63194) was unable to utilize D-melibiose and D-raffinose as determined by the Biolog PMs. Within this deleted region is the alpha-galactoside ABC transporter periplasmic solute-binding protein (*agpA*) which has been shown to abolish growth on raffinose and melibiose as the sole carbon sources (Gage and Long 1998). This same deletion mutant was shown by the Biolog PMs to be unable to

utilize lactose and lactulose and within this coding region lies the probable *lacEFGZ1K1* cluster which may encode the lactose ABC transport and metabolism genes.

The compounds whose catabolic genes were determined to be on pSymB were a starting point for some of the substrates used in the high-throughput screening. It was likely that some of these genes would be ABC transporters due to the abundance of these on the megaplasmid. The diversity of mutant phenotypes found through the deletion mutant screening shows the variety of substrates *S. meliloti* is capable of using as carbon and nitrogen sources. That these genes are found on pSymB indicate that the role of pSymB is instrumental in survival and competition through shifting environmental conditions. The ability to transport and metabolize compounds found in the soil will help *S. meliloti* survive the harsh nutrient deprived bulk soil so that it may have the chance to colonize alfalfa and engage in symbiosis.

It is interesting to note that RmF909 with a deletion spanning over 600 000 bp of pSymB (almost half of pSymB) was only unable to use four carbon sources out of all those tested. These were D-glucosaminic acid, 2-deoxy adenosine, palatinose, sorbose and hydroxy-L-proline. This is further evidence to support the hypothesis that pSymB is not essential to the cell but important for the metabolic diversity of *S. meliloti* in its natural habitat. Due to the abundance of transporters on pSymB it is plausible that some of these phenotypes are due to the deletion of ABC transporters. The small number of phenotypes identified in strain RmF909 is interesting considering the large number of transporters that have been deleted in this region, approximately 25 transport systems. One explanation for this could be due to overlapping substrate specificity in transporters.

This redundancy in transporter specificity has been previously found in the *thu* and *agl* genes of *S. meliloti* (Jensen et al. 2002).

6.2. Reporter fusions

The suicide, GusA reporter vector pTH1360 used in this study was used to facilitate single recombination of the vector into the genome so that one copy of *gusA* would be under control of the native promoter. The majority of fusions were built to the 3' ends of the last gene in the ABC transport cluster in order to keep all the genes functional. One of the useful features of this plasmid is the transcriptional terminator located before the multiple cloning site, thus all *gusA* transcription occurs from the promoter within the cloned fragment (Figure 3-6). Southern blot hybridization was carried out on a number of the fusion strains to ensure that the recombination of the plasmid was in the correct location.

6.3. Development of a high-throughput β-glucuronidase assay

The β-glucuronidase assay adapted to *S. meliloti* by Reeve et al. (Reeve et al. 1998), as with most other reporter enzyme assays, needed a variety of time consuming steps to carry out the assay accurately. In order to assay the ABC transporter – *gusA* fusions following growth in 36 different types of media it was necessary to develop a high-throughput assay which could be carried out in 96-well microtitre plate format. Before the assay could be developed the properties of β-glucuronidase assay were studied. It was found to be unnecessary to wash and resuspend the cell cultures prior to the assay

and that the addition of straight cell culture yielded the same results. The vortexing step after addition of the permeabilizing agent was not needed because the same results were found with no physical agitation. Although carrying out the assay in a 37°C water bath yielded higher GusA activity the approximate fold difference was the same when the assay was carried out at room temperature. Keeping cells on ice for up to 24 hours did not change the results of the assay and 0.01% SDS was found to be a suitable replacement for toluene to permeabilize the cells. All these factors were important in the development of the high-throughput assay to be carried out by the PerkinElmer MultiprobeII liquid handling system as seen in Figure 6-1.

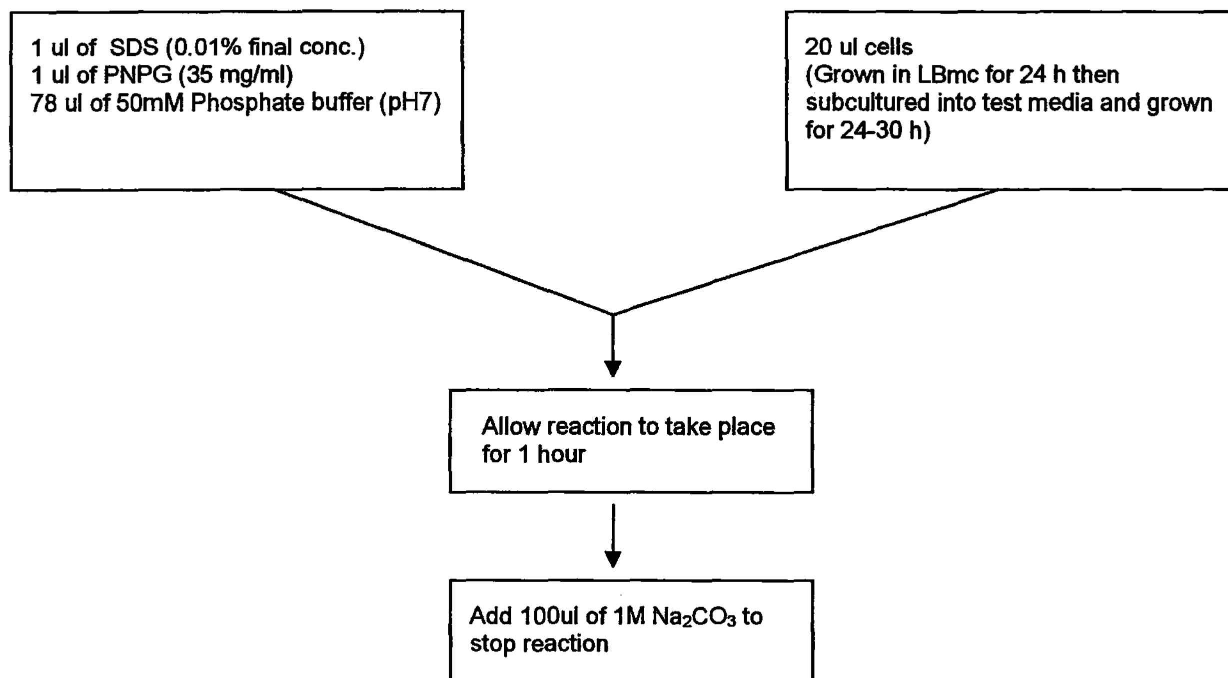

Figure 6-1. High-throughput β-glucuronidase assay developed for screening in 96-well microtitre plates

6.4. Screening Results

The screening protocol was carried out easily and efficiently with the PerkinElmer MultiprobeII liquid handling system using the high-throughput version of the β-glucuronidase assay. Due to the scavenging nature of ABC transporters, it makes sense that when a substrate is present the gene expression of the transporter will be very highly upregulated in order to take up as much of the compound as needed. Also in the absence of the substrate there is no need for the machinery specific to that one compound so the genes are down-regulated. In general when a transporter was induced it would be to very high levels ranging from 10 to 100 times over that of the background level depending on the gene, as seen in Figures 3-10 and 3-11. Many transport systems are not induced until the transport ligand is present in the media. An example of this is the alpha-galactoside (*agp*) uptake system where the permease *agpA* is induced by α-galactosides like raffinose and melibiose and a mutant of the permease is unable to use or transport these sugars. The enzyme assay used to study the gene expression of *agpA* was found to have specific activities ranging from 600 to 2000 with background levels around 50 (Gage and Long 1998). This tight regulation of ABC transporters made identification of inducing compounds relatively easy.

RmP206 (SMb20904::*gusA*) was found to be induced by a multitude of sugars: xylose, glucose, mannose, glucosamine, dextran (a polymer of glucose) and pea root and seed exudates and lentil seed exudates, as see in Figure 3-15. Glucosamine (results not shown) and dextran could be inducers of this gene because they are made up of glucose

and not necessarily specific for this transport system. SMb20904 lies next to the pSymB arg-tRNA gene and close to the *chvE* gene (Figure 6-2).

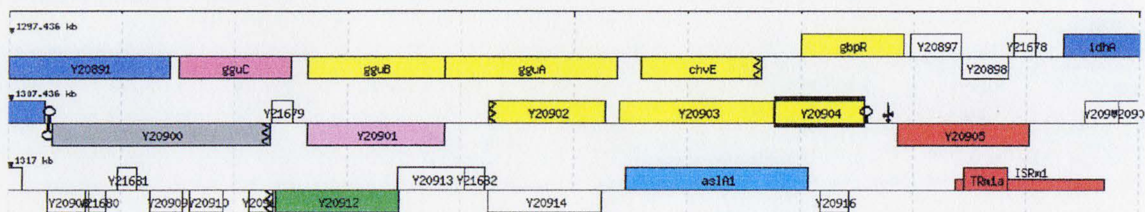

Figure 6-2. Gene map of SMb20904, part of an ABC transport system found to be induced by glucose, mannose and xylose (http://bioinfo.genopole-toulouse.prd.fr/annotation/iANT/bacteria/rhime/).

chvE is the PBP of an ABC transport system involved in virulence, chemotaxis and sugar transport in *A. tumefaciens* (Kemner et al. 1997). Some of the sugars thought to be transported by the *ggu* transport system overlap with the inducers of SMb20904. A knock-out mutant of the operon containing genes SMb20902, SMb20903 and SMb20904, RmP234 (SMb20902::*gusA*) was tested for growth on the inducing sugars mannose, xylose and glucose as the sole carbon sources but normal growth was observed in every case. Some fusions tested in this study were found to also be induced by mannose and xylose so the lack of phenotype of the knock-out mutant of the SMb20904 operon could be due to overlapping substrate specificity in transporters. Redundancy in transporters has been previously documented in sugar and amino acids in rhizobia. The alpha-glucoside *agl* and trehalose *thu* transporters are redundant because they both transport trehalose (Jensen et al. 2002). In *R. leguminosarum* there are general amino acid transporters that show broad specificity. *bra*$_{RI}$, a branched-chain amino acid permease transports acidic and basic polar amino acids (L-glutamate, L-arginine, and L-histidine) and neutral amino

acids (L-alanine and L-leucine). A mutant of this ABC transporter grows the same as wild type on all amino acids except has reduced growth on glutamine and is not able to grow on serine as the sole carbon and nitrogen sources, which indicates other modes of transport for these amino acids (Hosie et al. 2002).

RmP222 (SMb21587::*gusA*) was found to be induced by galactose and arabinose (Figure 3-20). A deletion mutant RmG462 whose deleted region encompasses this transporter system can grow on M9 minimal media agar plates with arabinose and galactose as the sole carbon sources. This would indicate that there are other transporters for these two sugars. There has not yet been an ABC transporter reported in the literature specific to only arabinose and galactose. There are multiple sugar transporters like *chvE* which transports other sugars in addition to arabinose and galactose. In *E. coli* there is an arabinose transporter *araFGH* that is only found to be induced by arabinose (Hendrickson et al. 1990).

RmP203 (SMb21342::*gusA*) was found to be induced by sorbose (Figure 3-19). It is annotated as an ABC sugar transport system.

RmP198 (SMb20979::*gusA*) was found to be induced by fucose (Figure 3-16). It is annotated as a C4-dicarboxylate ABC uptake system. Another strain RmP266 (SMb20428::*gusA*) was found to be induced by fucose and mannose (Figure 3-17). A large deletion encompassing this transport operon, RmF909 was found from the Biolog PMs survey results to have decreased ability to use fucose as the sole carbon source.

The ABC transport operon SMb21707 was found to be induced by galactosamine and glucosamine and to a lesser extent N-acetyl-D-glucosamine (Figure 3-18). These

transport genes are annotated as the components of putative urea/short-chain amide or branched-chain amino acid uptake ABC transporter. A deletion mutant spanning the region of this transporter (nucleotides 1452882 to 63194) was found to be unable to use a variety of nitrogen sources including urea, ammonia, nitrite and nitrate although it was able to use glucosamine and galactosamine (Biolog PMs data, see Appendix A). This fusion has not yet been assayed for induction by those compounds. In a study done to isolate carbon and nitrogen deprivation induced loci, this cluster was found to be induced during nitrogen deprivation (Milcamps et al. 1998). A possible function for this transporter would be to scavenge the environment for substrates containing nitrogen under nitrogen limiting conditions.

RmP194 (SMb21216::*gusA*) and RmP193 (SMb21138::*gusA*) were both found to be induced in pea seed exudates (Figure 3-23). Surprisingly the pea seed exudates and none of the others exudates were inducers. This could be due to the larger size of peas which could perhaps exude a higher level of compounds. The SMb21216 transporter is annotated as a sugar transporter and its operon has two metabolism genes, one sugar kinase and one amino transferase. SMb21138 is the operon located directly upstream from SMb21216 and contains no metabolism genes. This cluster is annotated as an amino acid transporter and SMb21138 has been shown to be induced by glucosamine and galactosamine. A deletion mutant RmF117 in which the region spanning nucleotides 770089 to 889262 of pSymB is removed could not grow on galactosamine, glucosamine and N-acetyl D-galactosamine as the sole carbon sources, but could use these amino sugars as the sole nitrogen source (Biolog data, see Appendix A). N-acetyl D-

glucosamine did not induce this gene fusion but N-acetyl D-galactosamine was not tested. A knock-out of this cluster would be useful to determine whether this is the main transporter for these amino sugars. It is possible that there are other transporters for these sugars because SMb21707 was also found to be induced by galactos- and glucosamine but perhaps for use as a nitrogen source and not as a carbon source such as the SMb21707 discussed above.

SMb20784 appears to be part of an annotated branched-chain amino acid ABC transporter and SMb20027 may be part of a tripartite tricarboxylate transporters (TTT) uptake system. Both of these genes were found to be induced by protocatechuate and hydroxybenzoate (Figures 3-21 and 3-22). The TTT family is another class of transporter which makes use of an extracytoplasmic solute receptor (Winnen et al. 2003). SMb20025 is similar to a *Bordetella* uptake gene (*bug*), which is homologous to a *Salmonella typhimurium* TTT gene *tctC*, which encodes a periplasmic citrate-binding protein. These *bug* genes are found in many soil bacteria, especially in the β-proteobacteria (Antoine et al. 2003). *Agrobacterium tumefaciens* and *Bradyrhizobium japonicum* contained genes highly similar to SMb20025. In *Pseudomonas putida* and *Acinetobacter* sp. hydroxybenzoate and protocatechuate are transported by a MFS *pcaK*, against a concentration gradient. Active transport would be important for the transport of these aromatic compounds because they would be present in the soil in very low concentrations (Nichols and Harwood 1997; D'Argenio et al. 1999). In *S. meliloti* a transporter for protocatechuate and hydroxybenzoate has yet to be found, so there is some significance in finding two transport gene clusters induced by these compounds. When the location of

these two gene clusters are taken into consideration it is even more probable that these transporters are involved in protocatechuate and hydroxybenzoate transport. The SMb20784 is located upstream of the protocatechuate catabolism genes *pcaBGHCD,* as seen in Figure 6-3. The SMb20027 transporter is located about 10 genes upstream of a shikimate 5-dehydrogenase metabolism gene, which is a precursor to protocatechuate.

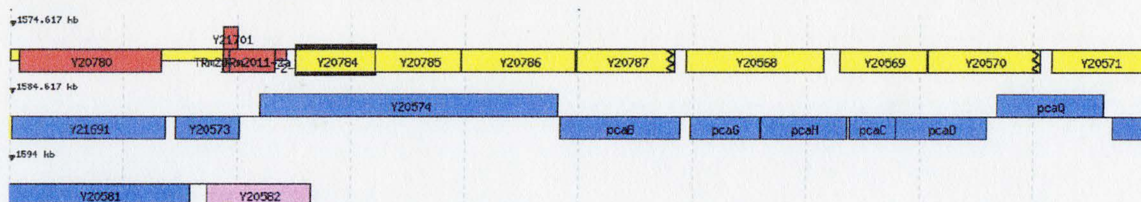

Figure 6-3. Gene map of SMb20784, part of a transport operon induced by pca and pob located upstream of the *pcaBGHCD* (the protocatechuate catabolic genes) (http://bioinfo.genopole-toulouse.prd.fr/annotation/iANT/bacteria/rhime/).

6.5. Plant Screening

The 45 reporter fusions were tested for GusA expression in the nodule and seven were found to have increased activity. There is difficulty in determining whether or not the elevated expression in these seven strains is important. Although elevated gene expression was detected in these nodule extracts it was not as high as that of RmP32 (*dme::gusA*). Gene expression levels in media containing an inducing substrate were in every case higher that those of RmP32 (*dme::gusA*), see Figure 3-25 for an example. An explanation for the lower levels of expression could be that although the inducing compound was present in the nodule, its levels were much lower in the nodule than when used as the sole carbon source in minimal media. Another possibility is that the lower levels of gene expression were due to the effects of catabolite repression by another

compound such as succinate which is known to be present in the nodule. In a study of the proteome of nodule bacteria, the ABC transporters were the largest protein group identified in the study with 84 different transporters detected in the culture grown bacteria. Thirteen ABC transporters were found in nodule bacteria that were not present in culture grown bacteria, suggestive of a very specialized nutrient exchange happening between *S. meliloti* and alfalfa during symbiosis. The types of ABC transporters found only in the nodule bacteria included those annotated as amino acid, iron and phosphate transporters (Djordjevic 2004). In a microarray study concerning changes in gene expression under microoxic and symbiotic conditions 114 genes encoding transport systems were found to be differentially expressed in the bacteroids, with the majority of these being phosphate and iron related transporters (Becker et al. 2004).

SMb20428 is a putative ABC transporter amino acid-binding protein found to be induced in the nodule (Figure 3-25). It is located in the same operon, upstream of three genes of unknown function, two of which are annotated as being involved in amino acid biosynthesis, which is odd considering it has been shown to be induced by mannose and fucose (Figure 3-17). This would indicate that the operon to which SMb20428 belongs is possibly not an amino acid transporter but a sugar transporter. In a proteomics study where the proteome of culture grown bacteria was compared to that of nodule bacteria, SMb20428 was also found only in nodule grown bacteria, although it is possible that this protein was not detected in the culture grown bacteria (Djordjevic 2004). This transport system would be of interest to study since it appears to be present in the nodule, to date it is not clear if the higher level of expression is due to the inducing compounds being

present within the nodule or if also it is something specific in symbiosis that is also turning this gene on.

SMb20442 is a putative periplasmic solute-binding protein which was shown to be induced in nodule extracts and has also been shown to have increased expression when grown in the presence of alfalfa seed exudates.

SMb21130 is a putative sulphate uptake ABC transporter and has not been shown to be induced by any other compounds tested to this date. This gene fusion was found to be induced in nodule extracts. Lying just downstream of this transporter are *nodP2Q2* which encode the two (putative) subunits of the sulphate adenylyltransferase enzyme which is responsible for providing the activated sulphate (3'-phosphoadenosine-5'-phosphosulfate) needed for the biosynthesis of sulphated Nod factors. Perhaps this transporter aids in accumulating the intracellular sulphate needed for this process (Schwedock and Long 1992).

SMb02793 is annotated as a conserved hypothetical protein and is located directly upstream to the origin of replication on the *S. meliloti* chromosome. This gene appears to always be turned on so its induction inside the nodule is not an implication of a role in symbiosis but rather due to the fact that it is highly constitutively expressed in all media (data not shown).

SMb20904 is a putative sugar uptake ABC transporter ATP-binding protein which has been shown to be induced by a variety of sugars including xylose, glucose mannose, glucosamine and various legume exudates (Figure 3-15). It is possible that one

or more of these sugars is present in the nodule to cause the increased expression of this gene during symbiosis.

SMb20263 is an ABC transporter periplasmic amino acid-binding protein which was found to be induced in nodule extracts and by hydroxyproline, allohydroxyproline and alfalfa root exudates. SMb20262 annotated as a putative semialdehyde dehydrogenase protein shown to be induced by hydroxyproline and allohydroxyproline. Although SMb20262 is induced by the same substrates as SMb20263 both have been shown to be under control of their own promoters (Figure 5-10). The expression of SMb20262 within the nodule was not tested in this study but in a proteomics study comparing nodule bacteria to culture grown bacteria SMb20262 was found only in nodule bacteria (Djordjevic 2004). It is probable that SMb20262 is part of the catabolic pathway so it would make sense that both the transport and metabolism genes would be induced under the same conditions. As discussed in Chapter 5 there is a putative L-proline 3-hydroxylase (SMc03253) which has been shown to be FixLJ regulated (Ferrieres et al. 2004). In a microarray study of nodule grown bacteria compared to culture grown bacteria, SMc03253 was found to be upregulated in nodule bacteria. Although the function of this enzyme has not been determined experimentally, the FixLJ regulation indicates that there is the potential for large amounts of hydroxyproline to be made in the nodule. If hydroxyproline is being made within the nodule this could be reason for the elevated gene expression of SMb20263 (a hydroxyproline transport gene) found in this study and for the detection of the α-ketoglutarate semialdehyde dehydrogenase (SMb20262) in the proteome of nodule grown bacteria in the proteomics

study by Djordjevic (Djordjevic 2004). Hydroxyproline is also found in many legume glycoproteins which have been shown be present in the nodule (Rathbun et al. 2002). In Figure 5-2 the gene expression of SMb20263 is compared after growth in 27 different types of media. Although the induction when hydroxyproline is used as the sole carbon source is clearly seen there is also a high background level of gene expression. This elevated, constitutive gene expression observed in SMb20263 under all conditions could be responsible for the apparent gene expression found in nodule extracts.

SMb21373 is a putative sugar kinase located in an operon with the components of a putative sugar uptake ABC transporter periplasmic solute-binding protein precursor. This transporter is involved in galactose metabolism via the galactitol pathway (Figure 4-6 and 4-7). It is induced by galactose and some of its metabolites (galactitol and tagatose) and a knock-out mutant of this operon is unable to grow on galactitol or tagatose as the sole carbon sources. It is possible that one of these sugars is present in the nodule but as with SMb20263 gene fusions which cause a knock-out phenotype result in a constitutive level of gene expression. Although SMb21373 and SMb20263 were shown to be highly induced by their respective ligands Figure 3-26 displays the constitutive level of expression that is observed after growth in LB.

6.6. Galactitol and Tagatose

A deletion mutant RmG470 was found unable to use dulcitol as the sole carbon source (Charles and Finan 1991). Further screening of this deletion using the Biolog PMs and on various carbon sources in this study found that this deletion was also unable to use

tagatose as the sole carbon source. The transporter beginning with SMb21377 (a putative sugar uptake ABC transporter periplasmic solute-binding protein) was found to be induced by galactose, galactitol, tagatose and sorbose so it was hypothesized that this transport and metabolism operon was responsible for the transport and metabolism of galactitol and tagatose (Figure 4-6). A knock-out of this transporter created by dissociating the promoter from the operon was unable to use galactitol and tagatose as the sole carbon sources but could grow on galactose and tagatose. Galactose is the precursor of galactitol so it was understandable that this was also an inducer but it was not clear why sorbose would induce this operon. In *Aspergillus nigulans* an alternate route of galactose metabolism via the LeLoir pathway was found by detecting growth using galactose as the sole carbon source in a *galE* mutant. This alternate pathway of galactose catabolism was found to involve the reduction of the galactose to galactitol the oxidation of galactitol by L-arabitol dehydrogenase to sorbose by L-arabitol dehydrogenase (Fekete et al. 2004).

It appears that *S. meliloti* possesses one pathway for the metabolism of both galactitol and tagatose from the results obtained using a knock-out of this operon (Table 4-1). Lying directly downstream of the ABC transport gene are 2 metabolism genes. SMb21374 is annotated as a putative sugar kinase of the phosophofructokinase family (*pfkB*) and SMb21375 as a putative sugar kinase with similarity to tagatose 6-phosphate kinase. Tagatose and galactitol metabolism has only been studied in *E. coli* and *Klebsiella oxytoca* that transport these two compounds by a PTS transporter. This would initiate a different metabolic pathway for these compounds due to the phosphorylation of

galactitol upon entry into the cell (Shakeri-Garakani et al. 2004). It appears that galactitol dehydrogenase (galactitol → D-tagatose) and tagatose kinase (D-tagatose → D-tagatose-6P) are needed for this metabolic pathway to be complete There is not much in the literature on these two enzymes due to the fact that this pathway has not been studied in a bacterium that does not use a PTS system for uptake. Tagatose kinase has been purified from *Mycobacterium butyricum* from cells grown in the presence of galactitol (Szumioo 1981).

Lying directly downstream of the operon is a putative transcription regulator of the LacI family, SMb21372. A mutant of this regulator was made by inserting a Sp-Sm resistance cassette into the middle of the gene. The SMb21372ΩSm-Sp knock-out was built in a RmP233 (SMb21377::gusA) background in order to investigate the expression in the regulator mutant, RmP245 (SMb21377::gusA, SMb21372ΩSpSm). There was increased β-glucuronidase activity in the regulator mutant strain as compared to RmP233 when grown in M9 minimal media with succinate which would indicate a repressor role for this regulator (Figure 4-8).

An interesting occurrence was observed with the deletion mutants of this operon when grown on galactitol or tagatose. Although a "no growth phenotype" would be observed upon incubation of about one week revertant colonies would appear (Table 4-2). This would indicate a second site mutation restoring the galactitol- , tagatose- phenotypes. Since these revertant colonies would appear in the deletion mutant G470 (with a deletion spanning nucleotides 972983 to 1084094 of pSymB) and RmP233 it would appear that

the second site reversion was occurring some distance from this galactitol transport and metabolism cluster.

6.7. Hydroxyproline

The screening of the large deletion mutants in the Biolog PMs system (Appendix A) and again by screening the mutants on M9 minimal media with various carbon sources (Table 3-1) first demonstrated that wild type *S. meliloti* could use hydroxyproline as a carbon source and that the genes necessary for its transport and metabolism were located within the region of pSymB deleted in strain RmF909. This compound was used as a substrate for the high-throughput screening of ABC transporters and two different transport operons, SMb20263 (part of a putative ABC transport cluster) and SMb20322 (part of a putative TRAP-T transport cluster) were found to be induced by this non-essential amino acid. Both of these genes were shown to be uninduced and perhaps repressed when grown in proline (Figure 5-7) and both were also found to be induced by hydroxyproline and its epimer allohydroxyproline (Figure 5-12). It was expected that both epimers would cause induction as previous biochemical studies on hydroxyproline metabolism in *P. putida* showed that both these epimers induced the metabolism genes and the first step in the metabolic pathway is the epimerization of hydroxyproline to allohydroxyproline (Adams 1959). We have noticed that uninduced SMb20322 has virtually no gene expression whereas SMb202623 is expressed at a low constitutive level in all types of growth media. When induced by hydroxyproline, SMb20322 displays a higher level of gene expression than SMb20263. Two metabolism genes lie upstream of

SMb20263 that are divergently transcribed, SMb20261 is annotated as a malate dehydrogenase and SMb20262 is annotated as a semialdehyde dehydrogenase. Theses genes display the same expression patterns as the transporter in that they are always turned on at a low level and induced to a greater level in the presence of either epimer of hydroxyproline (Figure 5-12). Each of these genes has its own promoter and the levels of induction by hydroxyproline vary slightly (Figure 5-10). When grown in the presence of hydroxyproline, SMb20261 doubles its expression level, SMb20262 increases three-fold and SMb20263 increases by four or five fold the constitutive levels.

A knock-out of the operon to which SMb20263 belongs was unable to grow on hydroxyproline as the sole carbon source on M9 minimal media plates while a knock-out of the operon to which SMb20322 belongs appeared to grow normally. Growth curves of both these strains in hydroxyproline revealed a subtle phenotype in the second transporter, SMb20322 where the growth was slightly slower than wild type and reached stationary phase at a lower cell density than wild type. This would indicate that the SMb20263 ABC transporter was the major transporter for hydroxyproline and since this operon probably contains the catabolic genes for hydroxyproline a knock-out of the SMb20263 operon is not able to grow on hydroxyproline as the sole carbon source.

Hydroxyproline metabolism has not been characterized genetically so it is of interest to compare the biochemical knowledge of the pathway to the annotated gene sequence of *S. meliloti*. As seen in Figure 5-8 the SMb20263 operon not only has two hydroxyproline induced genes upstream but 5 or 6 metabolism genes downstream in what appear to be one operon. An attempt to correlate the information present in the annotated

gene region to the characterized biochemical activities involved in hydroxyproline catabolism (Adams 1959; Adams 1973) is shown in Figure 6-3. The annotated gene sequence matches up to the enzymes necessary for the catabolism of hydroxyproline. One aspect that isn't clear would be why there are two proline racemases when the metabolism only requires one to convert hydroxyproline to allohydroxyproline because it has previously been shown in *P. putida* that only hydroxyproline and allohydroxyproline are inducers of the hydroxyproline transport and metabolism genes (Gryder and Adams 1970). In the metabolism cluster, there is one hypothetical protein without a specific annotation but also one gap in the pathway so perhaps Y20269 is the 1-pyrroline-4-hydroxy-2-carboxylate deaminase. There are two possibilities for the end product of hydroxyproline metabolism. 2-oxoglutarate could be converted into glutamate or the pathway could lead into the TCA cycle.

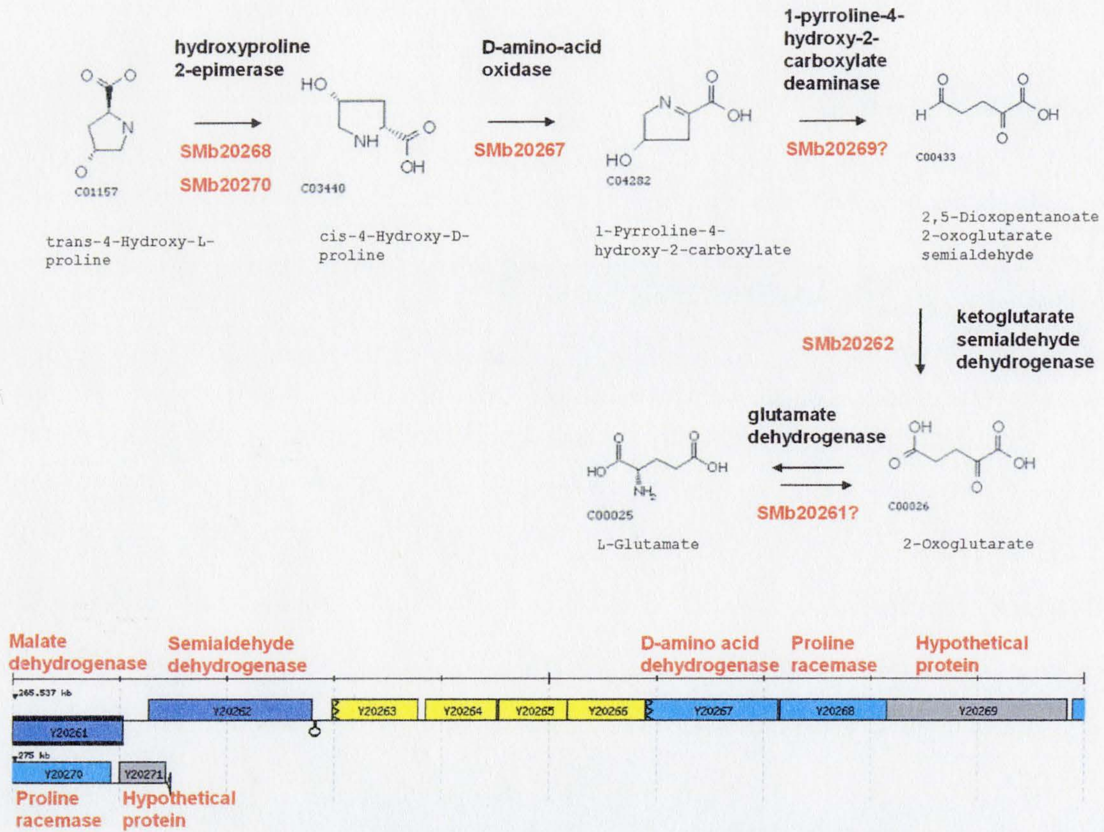

Figure 6-3. Hydroxyproline metabolic pathway and annotation of operon with hypothesized gene function.

As discussed in the above paragraph the hydroxyproline transporter (SMb20263) was shown to be induced in nodule extracts and in a proteomics study SMb20262 encoding the putative semialdehyde dehydrogenase was to be present in bacteroids and not culture grown bacteria (Djordjevic 2004). This indicates hydroxyproline may be present within the nodule. Hydroxyproline rich glycoproteins have been found within the infection threads of legumes and have also been shown to be upregulated in pea root nodules (Rathbun et al. 2002). From growth studies of *S. meliloti* on hydroxyproline as the sole carbon source it appears that this carbon source supports robust growth, even more so than succinate. Sm1021 grown in M9 minimal media with succinate as the sole

carbon source has a doubling time of six hours, whereas in M9 minimal media with hydroxyproline support a tripling of growth in six hours which demonstrates some importance of this imino acid to the bacterium.

6.8. Conclusions

This study has demonstrated the effectiveness of employing gene fusions to identify possible roles of genes of unknown function and the efficiency of generating expression data through the screening of reporter fusions. Inducers for half of the gene-fusions studied were found by screening only a limited group of carbon sources and exudates. A similar screen with more amino acids should identify more induced genes. Some interesting transport and metabolism operons have been located which definitely warrant more study due to their link with the competition of *S. meliloti* and due to their lack of characterization in the α-proteobacteria and other organisms.

The variety of substrates inducing the transporters indicates a scavenging function for this megaplasmid which would be important to *S. meliloti* in the nutrient deprived bulk soil. There is a diversity of transporters which have been linked to environmentally relevant compounds, such as galactitol, hydroxyproline, protocatechuate, p-hydroxybenzoate and various components and sugars of legumes seed and root exudates. This diversity and abundance of ABC transporters in *S. meliloti* indicate its need for scavenging in the environment. PTS systems can only transport sugars and their derivatives while ABC transporters can transport a wide range of substrates including peptides, iron, sugars and amino acids.

When comparing the number of transport genes in bacterial genomes it is clear that soil organisms have an abundance of these. 15 % of the *A. tumefaciens* genome is made up of transport genes with more than half of these encoding the various components of ABC transporters (60%) with 153 complete systems (Wood et al. 2001). This abundance is comparable to the *S. meliloti* genome in which 12% are transport genes with most of these being ABC transporters (Finan et al. 2001), (Galibert et al. 2001) and *Mesorhizobium loti* transport genes make up 12% of the genome (Van Sluys et al. 2002). The *Mycobacterium tuberculosis* genome is comprised of only 2.5% ABC transporters (Braibant et al. 2000) and *E. coli* has approximately 5% of it's genome annotated as ABC transporters (Linton and Higgins 1998). For organisms that live in environments less prone to fluctuations in nutrient availability who have less of a need to scavenge, the role for ABC transporters is not as important. A pathogen like *M. tuberculosis* that has little need for ABC transporters since it can easily obtain nutrients from the host it infects so it is understandable why this bacterium has a lower proportion of its genome dedicated to ABC transporters as compared to soil bacteria.

The diversity and abundance of ABC transporters, especially those that are environmentally relevant allow *S. meliloti* to compete and proliferate in the soil so that it may survive the harsh conditions and constantly changing conditions and successfully colonize its host plant alfalfa. With almost one fifth of the genes on pSymB being made up of transporters this indicates pSymB is important in giving *S. meliloti* the competitive edge it need to persist in the nutrient deprivation conditions of its natural environment.

Reference List

Adams,E. 1959. Hydroxyproline metabolism. I. Conversion to alpha-ketoglutarate by extracts of Pseudomonas. J. Biol Chem. **234:** 2073-2084.

Adams,E. 1973. The metabolism of hydroxyproline. Mol. Cell Biochem. **2:** 109-119.

Adams,E. and Frank,L. 1980. Metabolism of proline and the hydroxyprolines. Annu. Rev. Biochem. **49:** 1005-1061.

Antoine,R., Jacob-Dubuisson,F., Drobecq,H., Willery,E., Lesjean,S., and Locht,C. 2003. Overrepresentation of a gene family encoding extracytoplasmic solute receptors in Bordetella. J. Bacteriol. **185:** 1470-1474.

Becker,A., Berges,H., Krol,E., Bruand,C., Ruberg,S., Capela,D., Lauber,E., Meilhoc,E., Ampe,F., de Bruijn,F.J., Fourment,J., Francez-Charlot,A., Kahn,D., Kuster,H., Liebe,C., Puhler,A., Weidner,S., and Batut,J. 2004. Global changes in gene expression in Sinorhizobium meliloti 1021 under microoxic and symbiotic conditions. Mol. Plant Microbe Interact. **17:** 292-303.

Bochner,B.R., Gadzinski,P., and Panomitros,E. 2001. Phenotype microarrays for high-throughput phenotypic testing and assay of gene function. Genome Res. **11:** 1246-1255.

Bradford,M.M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72:** 248-254.

Braibant,M., Gilot,P., and Content,J. 2000. The ATP binding cassette (ABC) transport systems of Mycobacterium tuberculosis. FEMS Microbiol. Rev. **24:** 449-467.

Charles,T.C. and Finan,T.M. 1991. Analysis of a 1600-kilobase Rhizobium meliloti megaplasmid using defined deletions generated in vivo. Genetics **127:** 5-20.

Corbin,R.W., Paliy,O., Yang,F., Shabanowitz,J., Platt,M., Lyons,C.E., Jr., Root,K., McAuliffe,J., Jordan,M.I., Kustu,S., Soupene,E., and Hunt,D.F. 2003. Toward a protein profile of Escherichia coli: comparison to its transcription profile. Proc. Natl. Acad. Sci. U. S. A **100:** 9232-9237.

D'Argenio,D.A., Segura,A., Coco,W.M., Bunz,P.V., and Ornston,L.N. 1999. The physiological contribution of Acinetobacter PcaK, a transport system that acts upon protocatechuate, can be masked by the overlapping specificity of VanK. J. Bacteriol. **181:** 3505-3515.

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

Driscoll,B.T. and Finan,T.M. 1993. NAD(+)-dependent malic enzyme of Rhizobium meliloti is required for symbiotic nitrogen fixation. Mol. Microbiol. **7:** 865-873.

Fekete,E., Karaffa,L., Sandor,E., Banyai,I., Seiboth,B., Gyemant,G., Sepsi,A., Szentirmai,A., and Kubicek,C.P. 2004. The alternative D-galactose degrading pathway of Aspergillus nidulans proceeds via L-sorbose. Arch. Microbiol. **181:** 35-44.

Ferrieres,L., Francez-Charlot,A., Gouzy,J., Rouille,S., and Kahn,D. 2004. FixJ-regulated genes evolved through promoter duplication in Sinorhizobium meliloti. Microbiology **150:** 2335-2345.

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

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

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

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

Fruton JS and Simmonds S 1959. General Biochemistry Second Edition.

Gage,D.J. and Long,S.R. 1998. alpha-Galactoside uptake in Rhizobium meliloti: isolation and characterization of agpA, a gene encoding a periplasmic binding protein required for melibiose and raffinose utilization. J. Bacteriol. **180:** 5739-5748.

Galibert,F., Finan,T.M., Long,S.R., Puhler,A., Abola,P., Ampe,F., Barloy-Hubler,F., Barnett,M.J., Becker,A., Boistard,P., Bothe,G., Boutry,M., Bowser,L., Buhrmester,J., Cadieu,E., Capela,D., Chain,P., Cowie,A., Davis,R.W., Dreano,S., Federspiel,N.A., Fisher,R.F., Gloux,S., Godrie,T., Goffeau,A., Golding,B., Gouzy,J., Gurjal,M., Hernandez-Lucas,I., Hong,A., Huizar,L., Hyman,R.W., Jones,T., Kahn,D., Kahn,M.L., Kalman,S., Keating,D.H., Kiss,E., Komp,C., Lelaure,V., Masuy,D., Palm,C., Peck,M.C., Pohl,T.M., Portetelle,D., Purnelle,B., Ramsperger,U., Surzycki,R., Thebault,P., Vandenbol,M., Vorholter,F.J., Weidner,S., Wells,D.H., Wong,K., Yeh,K.C., and Batut,J. 2001. The composite genome of the legume symbiont Sinorhizobium meliloti. Science **293:** 668-672.

Gryder,R.M. and Adams,E. 1969. Inducible degradation of hydroxyproline in Pseudomonas putida: pathway regulation and hydroxyproline uptake. J. Bacteriol. **97:** 292-306.

Gryder,R.M. and Adams,E. 1970. Properties of the inducible hydroxyproline transport system of Pseudomonas putida. J. Bacteriol. **101:** 948-958.

Gu,D., Zhou,Y., Kallhoff,V., Baban,B., Tanner,J.J., and Becker,D.F. 2004a. Identification and characterization of the DNA-binding domain of the multifunctional PutA flavoenzyme. J. Biol Chem. **279:** 31171-31176.

Gu,D., Zhou,Y., Kallhoff,V., Baban,B., Tanner,J.J., and Becker,D.F. 2004b. Identification and characterization of the DNA-binding domain of the multifunctional PutA flavoenzyme. J. Biol Chem. **279:** 31171-31176.

Hendrickson,W., Stoner,C., and Schleif,R. 1990. Characterization of the Escherichia coli araFGH and araJ promoters. J. Mol. Biol **215:** 497-510.

Higgins,C.F. 1992. ABC transporters: from microorganisms to man. Annu. Rev. Cell Biol **8:** 67-113.

Higgins,C.F. 2001. ABC transporters: physiology, structure and mechanism--an overview. Res. Microbiol. **152:** 205-210.

Hosie,A.H., Allaway,D., Galloway,C.S., Dunsby,H.A., and Poole,P.S. 2002. Rhizobium leguminosarum has a second general amino acid permease with unusually broad substrate specificity and high similarity to branched-chain amino acid transporters (Bra/LIV) of the ABC family. J. Bacteriol. **184:** 4071-4080.

Jefferson,R.A., Burgess,S.M., and Hirsh,D. 1986. beta-Glucuronidase from Escherichia coli as a gene-fusion marker. Proc. Natl. Acad. Sci. U. S. A **83:** 8447-8451.

Jefferson,R.A., Kavanagh,T.A., and Bevan,M.W. 1987b. GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J. **6:** 3901-3907.

Jefferson,R.A., Kavanagh,T.A., and Bevan,M.W. 1987a. GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J. **6:** 3901-3907.

Jensen,J.B., Peters,N.K., and Bhuvaneswari,T.V. 2002. Redundancy in periplasmic binding protein-dependent transport systems for trehalose, sucrose, and maltose in Sinorhizobium meliloti. J. Bacteriol. **184:** 2978-2986.

Jimenez-Zurdo,J.I., Garcia-Rodriguez,F.M., and Toro,N. 1997. The Rhizobium meliloti putA gene: its role in the establishment of the symbiotic interaction with alfalfa. Mol. Microbiol. **23:** 85-93.

Jimenez-Zurdo,J.I., van Dillewijn,P., Soto,M.J., de Felipe,M.R., Olivares,J., and Toro,N. 1995. Characterization of a Rhizobium meliloti proline dehydrogenase mutant altered in nodulation efficiency and competitiveness on alfalfa roots. Mol. Plant Microbe Interact. **8:** 492-498.

Jones,P.M. and George,A.M. 1999. Subunit interactions in ABC transporters: towards a functional architecture. FEMS Microbiol. Lett. **179:** 187-202.

Katz,E., Kamal,F., and Mason,K. 1979. Biosynthesis of trans-4-hydroxy-L-proline by Streptomyces griseoviridus. J. Biol Chem. **254:** 6684-6690.

Kemner,J.M., Liang,X., and Nester,E.W. 1997. The Agrobacterium tumefaciens virulence gene chvE is part of a putative ABC-type sugar transport operon. J. Bacteriol. **179:** 2452-2458.

Koo,P.H. and Adams,E. 1974. Alpha-ketoglutaric semialdehyde dehydrogenase of Pseudomonas. Properties of the separately induced isoenzymes. J. Biol Chem. **249:** 1704-1716.

Lambert,A., Osteras,M., Mandon,K., Poggi,M.C., and Le Rudulier,D. 2001. Fructose uptake in Sinorhizobium meliloti is mediated by a high-affinity ATP-binding cassette transport system. J. Bacteriol. **183:** 4709-4717.

Linton,K.J. and Higgins,C.F. 1998. The Escherichia coli ATP-binding cassette (ABC) proteins. Mol. Microbiol. **28:** 5-13.

Manoharan,H.T. and Jayaraman,K. 1979. Mapping of the loci involved in the catabolic oxidation of L-hydroxyproline in Pseudomonas aeruginosa PAO. Mol. Gen. Genet. **172:** 99-105.

Meade,H.M., Long,S.R., Ruvkun,G.B., Brown,S.E., and Ausubel,F.M. 1982. Physical and genetic characterization of symbiotic and auxotrophic mutants of Rhizobium meliloti induced by transposon Tn5 mutagenesis. J. Bacteriol. **149:** 114-122.

Meighen,E.A. 1993. Bacterial bioluminescence: organization, regulation, and application of the lux genes. FASEB J. **7:** 1016-1022.

Milcamps,A., Ragatz,D.M., Lim,P., Berger,K.A., and de Bruijn,F.J. 1998. Isolation of carbon- and nitrogen-deprivation-induced loci of Sinorhizobium meliloti 1021 by Tn5-luxAB mutagenesis. Microbiology **144 (Pt 11):** 3205-3218.

Miller,J.F. 1972. Experiments in molecular genetics. Cold Spring Harbour Laboratory Press, Cold Spring Harbour, N. Y.

Mori,H., Shibasaki,T., Yano,K., and Ozaki,A. 1997. Purification and cloning of a proline 3-hydroxylase, a novel enzyme which hydroxylates free L-proline to cis-3-hydroxy-L-proline. J. Bacteriol. **179:** 5677-5683.

Nichols,N.N. and Harwood,C.S. 1997. PcaK, a high-affinity permease for the aromatic compounds 4-hydroxybenzoate and protocatechuate from Pseudomonas putida. J. Bacteriol. **179:** 5056-5061.

Nobelmann,B. and Lengeler,J.W. 1996. Molecular analysis of the gat genes from Escherichia coli and of their roles in galactitol transport and metabolism. J. Bacteriol. **178:** 6790-6795.

Oke,V. and Long,S.R. 1999. Bacterial genes induced within the nodule during the Rhizobium-legume symbiosis. Mol. Microbiol. **32:** 837-849.

Onishi,M., Okumura,Y., Okamoto,R., and Ishikura,T. 1984. Proline hydroxylation by cell free extract of a streptomycete. Biochem. Biophys. Res. Commun. **120:** 45-51.

Plouvier,V. 1972. [New study of scyllitol, myo-inositol and dulcitol in some botanical groups]. C. R. Acad. Sci. Hebd. Seances Acad. Sci. D. **274:** 1735-1738.

Prell,J., Boesten,B., Poole,P., and Priefer,U.B. 2002. The Rhizobium leguminosarum bv. viciae VF39 gamma-aminobutyrate (GABA) aminotransferase gene (gabT) is induced by GABA and highly expressed in bacteroids. Microbiology **148:** 615-623.

Prentki,P. and Krisch,H.M. 1984c. In vitro insertional mutagenesis with a selectable DNA fragment. Gene **29:** 303-313.

Prentki,P. and Krisch,H.M. 1984b. In vitro insertional mutagenesis with a selectable DNA fragment. Gene **29:** 303-313.

Prentki,P. and Krisch,H.M. 1984a. In vitro insertional mutagenesis with a selectable DNA fragment. Gene **29:** 303-313.

Quentin,Y., Fichant,G., and Denizot,F. 1999. Inventory, assembly and analysis of Bacillus subtilis ABC transport systems. J. Mol. Biol **287:** 467-484.

Rathbun,E.A., Naldrett,M.J., and Brewin,N.J. 2002. Identification of a family of extensin-like glycoproteins in the lumen of rhizobium-induced infection threads in pea root nodules. Mol. Plant Microbe Interact. **15:** 350-359.

Reeve,W.G., Tiwari,R.P., Wong,C.M., Dilworth,M.J., and Glenn,A.R. 1998. The transcriptional regulator gene phrR in Sinorhizobium meliloti WSM419 is regulated by low pH and other stresses. Microbiology **144 (Pt 12):** 3335-3342.

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

Ruvkun,G.B., Sundaresan,V., and Ausubel,F.M. 1982. Directed transposon Tn5 mutagenesis and complementation analysis of Rhizobium meliloti symbiotic nitrogen fixation genes
2. Cell **29:** 551-559.

Sambrook J and Russell DW 2001. Molecular Cloning: A Laboratory Manual Third Edition.

Schwedock,J.S. and Long,S.R. 1992. Rhizobium meliloti genes involved in sulfate activation: the two copies of nodPQ and a new locus, saa. Genetics **132:** 899-909.

Schweizer HP, Klassen TR, and Hoang T 1996. Improved methods for gene analysis and expression in Pseudomonads. pp. 229-237.

Shakeri-Garakani,A., Brinkkotter,A., Schmid,K., Turgut,S., and Lengeler,J.W. 2004. The genes and enzymes for the catabolism of galactitol, D-tagatose, and related carbohydrates in Klebsiella oxytoca M5a1 and other enteric bacteria display convergent evolution. Mol. Genet. Genomics **271:** 717-728.

Sharma,S.B. and Signer,E.R. 1990. Temporal and spatial regulation of the symbiotic genes of Rhizobium meliloti in planta revealed by transposon Tn5-gusA. Genes Dev. **4:** 344-356.

Singh,R.M. and Adams,E. 1965. Enzymatic deamination of delta-1-pyrroline-4-hydroxy-2-carboxylate to 2,5-dioxovalerate (alpha-ketoglutaric semialdehyde). J. Biol Chem. **240:** 4344-4351.

Suhr,M. and Kleiner,D. 1993. Genetic analysis of the regulatory putP region (coding for proline permease) in Klebsiella pneumoniae M5a1: evidence for regulation by the nac system
1. FEMS Microbiol. Lett. **114:** 191-194.

Szumioo,T. 1981. A novel enzyme, tagatose kinase, from Mycobacterium butyricum. Biochim. Biophys. Acta **660:** 366-370.

Tam,R. and Saier,M.H., Jr. 1993. Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. Microbiol. Rev. **57:** 320-346.

Tomii,K. and Kanehisa,M. 1998. A comparative analysis of ABC transporters in complete microbial genomes. Genome Res. **8:** 1048-1059.

Tsien,R.Y. 1998. The green fluorescent protein. Annu. Rev. Biochem. **67:** 509-544.

van Rhijn,P. and Vanderleyden,J. 1995. The Rhizobium-plant symbiosis. Microbiol. Rev. **59:** 124-142.

Van Sluys,M.A., Monteiro-Vitorello,C.B., Camargo,L.E., Menck,C.F., da Silva,A.C., Ferro,J.A., Oliveira,M.C., Setubal,J.C., Kitajima,J.P., and Simpson,A.J. 2002. Comparative genomic analysis of plant-associated bacteria. Annu. Rev. Phytopathol. **40:** 169-189.

Willis,L.B. and Walker,G.C. 1999. A novel Sinorhizobium meliloti operon encodes an alpha-glucosidase and a periplasmic-binding-protein-dependent transport system for alpha-glucosides. J. Bacteriol. **181:** 4176-4184.

Winnen,B., Hvorup,R.N., and Saier,M.H., Jr. 2003. The tripartite tricarboxylate transporter (TTT) family
1. Res. Microbiol. **154:** 457-465.

Wood,D.W., Setubal,J.C., Kaul,R., Monks,D.E., Kitajima,J.P., Okura,V.K., Zhou,Y., Chen,L., Wood,G.E., Almeida,N.F., Jr., Woo,L., Chen,Y., Paulsen,I.T., Eisen,J.A., Karp,P.D., Bovee,D., Sr., Chapman,P., Clendenning,J., Deatherage,G., Gillet,W., Grant,C., Kutyavin,T., Levy,R., Li,M.J., McClelland,E., Palmieri,A., Raymond,C., Rouse,G., Saenphimmachak,C., Wu,Z., Romero,P., Gordon,D., Zhang,S., Yoo,H., Tao,Y., Biddle,P., Jung,M., Krespan,W., Perry,M., Gordon-Kamm,B., Liao,L., Kim,S., Hendrick,C., Zhao,Z.Y., Dolan,M., Chumley,F., Tingey,S.V., Tomb,J.F., Gordon,M.P., Olson,M.V., and Nester,E.W. 2001. The genome of the natural genetic engineer Agrobacterium tumefaciens C58. Science **294:** 2317-2323.

APPENDIX

The Biolog data shown below was obtained by Punita Anja & Finan TM (unpublished data) and analyzed in this study

PM1: Carbon sources
PM2: Carbon sources
PM3: Nitrogen sources
PM4: Phosphorous & Sulfur sources

RmF263 (5408 Fix+ revertant taken from nodule)
Ω5033Δ5007
889262 – 1084094

		Utilized by F263 and wt	
PM	**Not utilized by F263 (but used by wt)**	**more slowly than wt (or weaker than wt)**	**more quickly than wt (or stronger than wt)**
3 (B10)		D-Glutamic acid	
3 (C7)	D-Lysine		
3 (C8)		D-Serine (both weak)	
3 (H5)		Ala-His	
3 (B3)	L-Histidine		
3 (E12)		N-Acetyl-D-Galactosamine	
4 (A9)			
4 (A10)			
4 (A11)			
4 (C11)			
4 (F10)		L-Cysteic acid	
4 (E6)		Phosphono Acetic Acid	
4 (E7)		2-aminoethyl Phosphonic acid	
4 (H12)		Tetramethylene Sulfone	
4 (E10)			Thymidine- 5'-monophosphate

Rm5416
Ω5007Δ5011
1084094 - 1177742

		Utilized by 5416 and wt	
PM	**Not utilized by 5416 (but used by wt)**	**more slowly than wt (or weaker than wt)**	**more quickly than wt (or stronger than wt)**
3 (D10)	Ethylenediamine		
3 (D12)		Agmatine (weak)	
3 (D1)	N-Acetyl D,L Glutamic acid		
3 (E12)	N-Acetyl-D-Galactosamine		
3 (F11)		Uridine (weak)	
3 (F12)	Inosine		
3 (G1)	Xanthine		

RmG462
PM 3 & 4 missing
Ω5033Δ5011
889262 - 1177742

PM	Not utilized by G462 (but used by wt)	Utilized by G462 and wt	
		more slowly than wt (or weaker than wt)	more quickly than wt (or stronger than wt)
1 (A12)	Dulcitol		
1 (B12)		L-Glutamic acid (weak)	
1 (C12)		Thymidine (weak)	
1 (D1)	L-Asparagine		
1 (D3)	D-Glucosaminic acid		
1 (G3)	L-Serine		
1 (G5)	L-Alanine		
2 (B1)	N-Acetyl-D-Galactosamine		
2 (B9)		2-Deoxy-D-Ribose (weak)	
2 (D6)	D-Tagatose		
2 (D11)	D-Amino Valeric Acid		
2 (G4)	L-Arginine		
2 (G6)	L-Histidine		
2 (G7)	L-Homoserine		
2 (G10)	L-Leucine		
2 (G11)	L-Lysine		
2 (H1)	L-Ornithine		

RmF117
Ω5060Δ5033
770089 - 889262

PM	Not utilized by F117 (but used by wt)	Utilized by F117 and wt: more slowly than wt (or weaker than wt)	more quickly than wt (or stronger than wt)
1 (C12)	Thymidine		
1 (D3)	D-Glucosaminic acid		
1 (E12)	Adenosine		
1 (F6)		Bromo Succinic acid (both weak)	
1 (G10)		Methyl Pyruvate	
2 (B1)	N-Acetyl D-Galactosamine		
2 (B9)			
2 (G2)			L-Alinamide (more strongly)
3 (E9)	D-Galactoamine		
3 (A11)			L-Cystein
3 (A6)			Bluret
3 (G1)		Xanthine	
3 (G2)		Xanthosine	
3 (F12)		Inosin (slow and weak)	
3 (F9)		Thymidine (slow and weak)	
4 (C8)		Cytidine-3'-monophosphate (both weak and slow)	
4 (E8)		Methylene Diphosphonic Acid	
4 (F9)	L-Cysteinyl-Glycine		
4 (F11)	Cysteamine		
4 (E4)		Phosphoryl Choline	
4 (H12)	Tetramethylene Sulfone		

RmG373
Ω5177Δ5069
1452882 - 63194

PM	Not utilized by G373 (but used by wt)	Utilized by G373 and wt: more slowly than wt (or weaker than wt)	Utilized by G373 and wt: more quickly than wt (or stronger than wt)
1 (A5)	Succinic Acid		
1 (A7)	L-Aspartic Acid		
1 (C3)	D,L-Malic Acid		
1 (C8)	Acetic Acid		
1 (C11)	D-Melibiose		
1 (C12)	Thymidine		
1 (D8)	α-Methyl—D-Glucoside		
1 (D9)	α-D-Lactose		
1 (D10)	Lactulose		
1 (F5)	Fumaric Acid		
1 (F6)	Bromo SuccinicAcid		
1 (F7)	Propionic Acid		
1 (G9)	Mono Methyl Succinate		
1 (G12)	L-Malic Acid		
1 (H6)		L-Lyxose (weak)	
1 (H12)		2-Aminoethanol (weak)	
2 (B9)		2-Deoxy-D-Ribose (weak)	
2 (B12)	3-O-β-D-Galacto-pyranosyl-D-Arabinose		
2 (C6)		α-Methyl-D-Galactoside (weak)	
2 (C7)		β- Methyl-D-Galactoside (weak)	
2 (D1)	D-Raffinose		
2 (D11)		D-Amino-ValericAcid (weak)	
2 (E8)	β-Hydroxy Butyric Acid		
2 (F3)	MelibionicAcid		
2 (G7)		L-Homoserine (weak)	
2 (G9)	L-Isoleucine		
2 (G10)		L-Leucine (weak)	
3 (A2)	Ammonia		

3 (A3)	Nitrite		
3 (A4)	Nitrate		
3 (A5)	Urea		
3 (A10)	L-Aspartic Acid		
3 (B2)	Glycine		
3 (B4)	L-Isoleucine		
3 (B5)		L-Leucine (weak)	
3 (C2)	L-Valine		
3 (C3)	D-Alanine		
3 (C7)	D-Lysine		
3 (C10)	L-Citrulline		
3 (C11)		L-Homoserine (weak)	
3 (D2)	N-Phthaloyl-L-Glutamic Acid		
3 (D5)		Methylamine (weak)	
3 (D8)	Ethylamine		
3 (D9)		Ethanolamine	
3 (D12)	Agmatine		
3 (E9)			D-Galactosamine
3 (F2)	Adenine		
3 (F3)		Adenosine (weak)	
3 (F4)		Cytidine (weak)	
3 (F5)	Cytosine		
3 (F7)	Guanosine		
3 (F8)	Thymine		
3 (F9)	Thymidine		
3 (F10)	Uracil		
3 (F12)	Inosine		
3 (G1)	Xanthine		
3 (G2)	Xanthosine		
3 (G5)	Allantoin		
3 (G6)	Parabanic Acid		
3 (G8)		α-Amino-N-Butyric Acid (weak)	
3 (G12)	2-Amino-N Valeric Acid		
3 (H11)	Gly-Met		
3 (H12)	Met-Ala		

RmF909
Ω5085Δ5047
106128 – 735511

PM	Not utilized by G462 (but used by wt)	Utilized by G462 and wt: more slowly than wt (or weaker than wt)	Utilized by G462 and wt: more quickly than wt (or stronger than wt)
1 (D3)	D-Glucosaminic Acid		
1 (E11)	2-Deoxy Adenosine		
1 (H5)		D-Psicose (weak)	
1 (H6)		L-Lyxose (weak)	
2 (B3)		β-D-Allose (weak)	
2 (B9)		2-Deoxy-D-Ribose (weak)	
2 (B11)		D-Fucose (weak)	
2 (C5)		Maltitol (weak)	
2 (C12)	Palatinose		
2 (D4)	L-Sorbose		
2 (G7)		L-Homoserine (weak)	
2 (G8)	Hydroxy-L-Proline		
3 (A11)			
3 (C7)		D-Lysine (weak)	
4 (A11)			
4 (C11)			
4 (D6)	Uridine-2'-monophosphate		
4 (E1)		O-Phospho-D-Tyrosine (weak)	
4 (E11)		Inositol Hexaphosphate (weak)	
4 (F12)			L-Cysteine Sulfinic Acid
4 (G7)		L-Methionine (weak and slow)	
4 (G12)		L-Methionine Sulfone (weak and slow)	