Factors involved in the regulation of purine degradation genes in *Sinorhizobium meliloti*

by

Keith T. Walsh

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Biology

Waterloo, Ontario, Canada, 2010

© Keith T. Walsh 2010

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Genes involved in purine degradation in *Sinorhizobium meliloti* to date remain largely uncharacterized. Analysis of the *bdhAxdhA2xdhB2* operon established a link between the degradation of purines and the carbon storage compound poly-3-hydroxybutyrate (PHB). This operon contains genes (*xdhA2xdhB2*) that encode xanthine oxidase / xanthine dehydrogenase, an enzyme involved in the conversion of hypoxanthine and xanthine to uric acid. The *bdhA* gene located in the same operon encodes 3-hydroxybutyrate dehydrogenase, an enzyme responsible for catalyzing the second step in PHB degradation. This linkage between the degradation of PHB (a carbon source) and purines (a nitrogen source) suggests a possible means by which *Sinorhizobium meliloti* obtains sufficient carbon and nitrogen to allow it to successfully colonize a host plant.

Purine degradation genes in *S. meliloti* have also been studied by the phenotypic characterization of Tn5 mutants unable to utilize hypoxanthine. Mutations resulting in these phenotypes were found in three different genes, SMc03849 (*ccmC*), a cytochrome c biogenesis mutant, SMb20684, a gene coding for a hypothetical protein possibly involved in the utilization of glyoxylate and SMb2192, a gene coding for a membrane spanning protein possibly involved in purine transport. In this study we further characterized these mutants

by examining their ability to establish a symbiosis with *Medicago sativa* (alfalfa) and to fix atmospheric nitrogen. It was demonstrated that in the case of all the mutant strains there was a competitive deficiency in terms of gaining entry to root nodules relative to the wild-type strain. It was shown that this deficiency occurred even in strains capable of fixing atmospheric nitrogen suggesting that the inability to utilize hypoxanthine impairs the ability of *S. meliloti* to colonize the host plant.

Of all of these genes studied thus far only one (SMb21292) is located in the region of the genome containing the greatest number of genes potentially involved in purine degradation. In this study we used transcriptional fusions to confirm the activation of genes in this genomic region when grown in media containing purines as carbon and nitrogen sources. These genes include *xdhA1*, SMb21284 and *guaD1*. Genes from the genome region containing the mixed function operon including *xdhA2* and *guaD2* were also studied. In addition we were able to demonstrate the requirement of *xdhC* in producing a functional oxidase / xanthine dehydrogenase as well as the ability to grow on hypoxanthine or xanthine as a carbon and nitrogen source. This work was also able to demonstrate the critical nature of the LysR transcription regulator (SMb21291) in purine degradation in *S. meliloti*. Mutating this gene resulted in an inability to grow on hypoxanthine or xanthine as well as an alteration in levels of xanthine oxidase / xanthine dehydrogenase activity. By transducing the gene fusions into the LysR mutant background it was demonstrated that the protein coded for by SMb21291 acts to regulate or influence the expression of genes involved in the purine degradation cycle such as SMb21284 and *xdhA1*. In addition we were able to characterize strains with mutations in purine degradation genes in terms of their growth on different purines.

Acknowledgements

I would like to thank all those whose contributions both great and small made this thesis possible.

In particular I would like to thank Dr. Trevor Charles for providing the moral, financial and intellectual support so indispensable to moving this project forward. His enthusiasm for scientific inquiry, willingness to try new approaches, patience and advice were invaluable and greatly appreciated.

I would like thank my committee members, Dr. Brian Ingalls, Dr. Kesen Ma and Dr. Barbara Moffatt for their helpful advice, discussions and ideas. I would also like to thank Dr. Herb Schellhorn for reviewing this thesis.

I would like to thank Dr. Turlough Finan from McMaster University for providing the fusion library and pTH1703 plasmid so integral to this project.

I would like to thank all my colleagues from Charles lab both past and present. I would like to thank Adi Rolider, Anh Ly and Asha Jacob for their friendship, support and scientific advice. I would like to thank David Capstick not only for his friendship but for providing the platform from which this project was launched and his key role in the creation of our fusion strains.

I am particularly grateful to Louise Bélanger who has been there from the beginning to the end of this thesis and has provided continual and unwavering support which has made this process infinitely more enjoyable.

Financial support for this project was provided by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Ontario Graduate Scholarship program. Their support is greatly appreciated.

Table of Contents

List of Tables **xii**

List of Figures **xiii**

Abbreviations **xvii**

Claims of Contributions to Knowledge **xxi**

1 Literature Review **1**

1.1 Introduction 1

1.2 *Sinorhizobium meliloti* 3

1.3 Rhizobia-Legume Symbiosis 4

1.4 Purine Metabolism 6

1.5 Xanthine Oxidase/Dehydrogenase 9

1.6 Poly-beta-hydroxybutyrate (PHB) 11

1.7 PHB Plays a Role in Bacterial Survival 11

1.8 The Role of PHB in Rhizobial Physiology 12

1.9 *S. meliloti* PHB Metabolism 14

1.10 The *S. meliloti bdhA* Gene 15

1.11 Symbiotic and Competitive Phenotypes of PHB Mutants of *S. meliloti* . . 16

1.12 The Relationship between EPS Production and PHB Synthesis 17

1.13 How Does *S. meliloti* Colonize the Rhizosphere of Alfalfa? A Proposed Model 18

1.14 The Genetic Toolbox of *S. meliloti* 19

1.15 LysR Transcriptional Regulators 20

1.16 Review of LysR Transcriptional Regulators in Rhizobia 24

1.16.1 *nodD* and *syr* genes 24

1.16.2 Other LysR proteins found in Rhizobia 31

1.17 This Work 35

2 Materials and Methods **37**

2.1 Bacterial Culture and Microbiological Techniques 37

2.1.1 Bacterial Strains, Plasmids, Transposons and Phage 37

2.1.2 Bacterial Growth Media 38

2.2 Bacterial Genetic Techniques 47

2.2.1 Bacterial Conjugation 47

2.2.2 Single-Crossover Mutagenesis 47

2.2.3 Preparation of ϕM12 Transducing Lysate 48

2.2.4 Transductions using ϕM12 Phage 48

2.2.5 Creation of Fusion Strains using pTH1703 48

2.2.6 Complementation of Mutant Strains 49

2.3 Molecular Biology Techniques 50

2.3.1 DNA Isolation and Purification 50

2.3.2 Genomic DNA Isolation 50

2.3.3 Preparation of Competent *E. coli* cells 51

2.3.4 Transformation of *E. coli* Competent cells 52

2.3.5 DNA Manipulation 52

2.4 Biochemical Methods 54

2.4.1 Analysis of Fusion Strains 54

2.4.2 Cell-Free Extracts 55

2.4.3 Enzyme Assay and Activity 56

2.5 Plant Assays 57

2.5.1 Inoculation of Plants with Bacterial Strains 57

2.5.2 Statistical Analysis 59

3 Phenotypic Analysis of Mutants Unable to utilize Hypoxanthine 60

3.1 Introduction 60

3.2 Results 62

3.2.1 Nodulation Phenotype 62

3.2.2 Shoot Dry Weights 64

3.2.3 Nodulation Kinetics 64

3.2.4 Nodulation Competitiveness 67

3.2.5 Acetylene Reduction Assays 67

3.3 Discussion 70

4 Purine Mutants 77

4.1 Introduction 77

4.2 Materials and Methods 78

4.3 Results 79

4.3.1 Sequence Analysis 79

4.3.2 Organization of *xdh* Genes in Species Closely Related to *S. meliloti* 81

4.3.3 Strain RmF117 84

4.3.4 Expression of *lacZ* Fusion Strains 84

4.3.5 *xdhA1* Fusion 84

4.3.6 SMb21285 Fusion 87

4.3.7 *guaD1* Fusion 89

4.3.8 SMb21284 Fusion 89

4.3.9 SMb21281 Fusion 92

4.4 Analysis of *lacZ* Fusions to Purine Degradation Genes Found Elsewhere in the *S. meliloti* Genome 92

4.4.1 *xdhA2* Fusion 92

4.4.2 *guaD2* Fusion 95

4.4.3 Phenotypic Analysis of Mutations in *xdhA1*, *xdhC* and SMb21291 . 95

4.4.4 *xdhC* Mutant 99

4.4.5 SMb20847 Mutant 103

4.4.6 Analysis of the Impact of the SMb21291 Mutation on *lacZ* Expression in Fusion Strains . . . 104

4.5 Growth and Expression of *lacZ* Fusion Strains on Different Carbon Sources 114

4.6 Discussion . . . 117

4.6.1 The *S. meliloti* RmF117 Strain . . . 119

4.6.2 Analysis of *lacZ* Fusions to Putative Purine Degradation Genes . . 120

4.6.3 Analysis of *xdhC* . . . 129

4.6.4 The Impact of Mutating SMb21291, a Putative LysR Transcriptional Regulator . . . 132

4.6.5 Growth and Expression of *lacZ* Fusion Strains on Carbon Sources . 148

5 Attempts to express the SMb21291 LysR protein 149

6 General Discussion 155

References 163

List of Tables

2.1 Bacterial strains used in this study. 39

2.2 Plasmids used in this study. 42

2.3 Primers used in this study. 44

3.1 Nodulation competitiveness of *S. meliloti* mutants co-inoculated with Rm1201 on *Medicago sativa* cv *Iroquois* plants. 69

4.1 Growth of *lacZ* fusion strains on M9 minimal media with various purines as sole carbon sources or with purines supplemented with glucose and ammonium chloride. 115

4.2 Carbon sources on which *S. meliloti lacZ* fusions strains were grown and assayed for *lacZ* expression which showed no change in expression relative to SmP110. 116

List of Figures

1.1 Pathway for purine degradation 7

1.2 Model of LysR activation of gene transcription. In the absence of inducer molecules (top) the LysR protein binds to DNA as a tetramer causing a bend in the DNA. 28

3.1 Phenotype of nodules from *Medicago sativa* cv *Iroquois* plants inoculated with *S. meliloti* Rm1021 (wild type), Rm11421 (*ccmC* mutant), Rm11424 (SMb21292 mutant), Rm11425 (*ccmC* mutant), Rm11428 (*ccmC* mutant) and Rm11429 (SMb20684 mutant). 63

3.2 Shoot dry weight of *Medicago sativa* cv. *Iroquois* plants inoculated with *S. meliloti* Rm1021 (wild type), Rm11421 (*ccmC* mutant), Rm11424 (SMb21292 mutant), Rm11425 (*ccmC* mutant), Rm11428 (*ccmC* mutant) and Rm11429 (SMb20284 mutant). 65

3.3 Nodulation kinetics for the percentage of *Medicago sativa* cv *Iroquois* plants nodulated by *S. meliloti* Rm1021 66

3.4 Acetylene reduction assay of *S. meliloti* hypoxanthine utilization mutant strains inoculated onto *Medicago sativa* cv *Iroquois* plants. 68

4.1 Organization of *xdh* genes in bacterial species closely related to *S. meliloti*. 82

4.2 Relative location of putative purine degradation genes deleted in *S. meliloti* RmF117 strain. 85

4.3 *lacZ* expression for SmRL2522 (*xdhA1lacZ* fusion strain) grown in various M9 minimal media. 86

4.4 *lacZ* expression for Rm11513 (SMb21285*lacZ* fusion strain) grown in various M9 minimal media. 88

4.5 *lacZ* expression for Rm11522 (*guaD1lacZ* fusion strain) grown in various M9 minimal media. 90

4.6 *lacZ* expression for Rm11517 (SMb21284("uricase like protein")*lacZ* fusion strain grown in various M9 minimal media. 91

4.7 *lacZ* expression for Rm11519 (SMb21281*lacZ* fusion strain) grown in various M9 minimal media. 93

4.8 *lacZ* expression for Rm11518 (*xdhA2lacZ* fusion strain) grown in various M9 minimal media. 94

4.9 *lacZ* expression for Rm11518 (*guaD2lacZ* fusion strain) grown in various M9 minimal media. 96

4.10 Growth of *S. meliloti xdhA1* mutant and SmP110 (wild-type) strains on M9 minimal media 98

4.11 Growth of *S. meliloti xdhC* mutant and SmP110 (wild-type) strains on M9 minimal media. 100

4.12 Growth of *S. meliloti* SMb21291 mutant (Rm11507) and SmP110 (wild-type) strains on M9 minimal media. 101

4.13 Staining of *S. meliloti* cell-free extract on a non-denaturing polyacrylamide gel for xanthine oxidase/xanthine dehydrogenase activity. 102

4.14 *lacZ* expression for *S. meliloti xdhA1lacZ* fusion strain transduced into the SMb21291 mutant grown in various M9 minimal media. 105

4.15 *lacZ* expression for *S. meliloti* SMb21285*lacZ* fusion strain transduced into the SMb21291 mutant grown in various M9 minimal media. 107

4.16 *lacZ* expression for *S. meliloti guaD1lacZ* fusion strain transduced into the SMb21291 mutant grown in M9 minimal media. 108

4.17 *lacZ* expression for SMb21284 *S. meliloti lacZ* fusion strain transduced into the SMb21291 mutant grown in various M9 minimal media. 110

4.18 *lacZ* expression for *S. meliloti guaD2lacZ* fusion strain transduced into the SMb21291 mutant grown in various M9 minimal media. 112

4.19 *lacZ* expression for *S. meliloti xdhA2lacZ* fusion strain transduced into the SMb21291 mutant grown in various M9 minimal media. 113

4.20 Distance tree comparing the similarity of the *S. meliloti* SMb21291 sequence with similar sequences in other bacterial genomes 141

4.21 Subset of possible palindromic sequences predicted by EMBOSS for the *S. meliloti* genomic region between SMb21284 and SMb21285. 146

5.1 Sequence of SMb21291 construct cloned into the pSMART vector 152

6.1 Purine degradation pathway showing position of reactions catalyzed or regulated by genes analyzed in this study 158

List of Abbreviations

AA	acetoacetate
AA-CoA	acetoacetyl-CoA
Ade	adenine
Amp	ampicillin
ATP	adenosine-5'-triphosphate
Arg	arginine
BdhA	3-hydroxybutyrate dehydrogenase
BLAST	Basic Local Alignment Search Tool
Cm	chloramphenicol
DMSO	dimethyl sulfoxide
DTT	dithiothreitol
DNA	deoxyribonucleic acid
EDTA	ethylenediaminetetraacetic acid
EPS	exopolysaccharide
GC	gas chromatography
Gm	gentamicin
Glu	glucose
gua	guanine

Gus	Beta-glucuronidase
HB	3-hydroxybutyrate
HTH	helix-turn-helix
HX	hypoxanthine
kb	kilobases
kDa	kiloDalton
Km	kanamycin
LB	Luria-Bertani
LysR	LysR type transcriptional regulator
LTTR	LysR type transcriptional regulator
Lys	Lysine
M9	M9 minimal media
Mb	megabase
NAD	Nicotinamide adenine dinucleotide
NADP	Nicotinamide adenine dinucleotide phosphate
NH_4	ammonium chloride
Nm	neomycin
nM	nanomole
OD	optical density
ONPG	ortho-Nitrophenyl-β-galactoside
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PHA	polyhydroxyalkanoate
PHB	poly-β-hydroxybutyrate
Phe	phenylalanine

PMA	phenazine methosulphate
ppGpp	guanosine tetraphosphate
RBS	ribosome binding site
RNA	ribonucleic acid
rRNA	ribosomal RNA
SDR	short-chain dehydrogenase / reductase
Sm	streptomycin
TCA	tricarboxylic acid
Tc	tetracycline
TAE	Tris-acetate-EDTA
Trp	tryptophan
TY	tryptone-yeast extract
UA	uric acid
UV	ultraviolet
WT	wild-type
X	xanthine
X-Gal	5-bromo-4-chloro-3-indolyl-D-galactopyranoside
XO/XDH	xanthine oxidase/xanthine dehydrogenase

Claims of Contributions to Knowledge

1. I examined the symbiotic phenotype of *Sinorhizobium meliloti* strain Rm11421 (SMb21292, a conserved hypothetical membrane protein, mutant) and found it to be less competitive than the wild-type strain Rm1021 in terms of establishing itself in root nodules on *Medicago sativa*. It was also demonstrated that this reduced competitiveness was not related to the strain's levels of nitrogenase activity (i.e. its ability to fix atmospheric nitrogen).

2. I examined the symbiotic phenotype of *Sinorhizobium meliloti* strain Rm11429 (SMb20284, a putative ABC transporter, mutant) and found it to be less competitive than the wild-type strain Rm1021 in terms of establishing itself in root nodules on *Medicago sativa*. It was also demonstrated that this reduced competitiveness was not related to the strain's levels of nitrogenase activity (i.e. its ability to fix atmospheric nitrogen).

3. I was able to add an additional piece of evidence in the form of an acetylene reduction assay to confirm that mutations in the *S. meliloti* SMc03849 gene (*ccmC*) result in the loss of nitrogen fixing ability.

4. I was able to demonstrate that the *S. meliloti xdhA1* gene shows significantly elevated

levels of expression when grown on hypoxanthine or xanthine. I was also able to demonstrate that expression levels were reduced when *S. meliloti* was grown with hypoxanthine or xanthine plus an additional nitrogen and carbon source source.

5. I was able to demonstrate that the *S. meliloti* SMb21285 gene (the first gene in the operon predicted to also include *xdhA1*) shows significantly elevated levels of expression when grown on hypoxanthine or xanthine. I was also able to demonstrate that expression levels were reduced when *S. meliloti* was grown with hypoxanthine or xanthine plus an additional nitrogen and carbon source source.

6. I was able to demonstrate that the *S. meliloti guaD1* gene shows significantly elevated levels of expression when grown in media containing guanine. I was also able to demonstrate that the high level of expression is reduced when an alternative carbon and nitrogen source was provided. It was also shown that *guaD1* showed a smaller though significant increase in expression when grown in media containing xanthine or hypoxanthine.

7. I was able to demonstrate that the *S. meliloti* SMb21284 (a gene encoding a predicted uricase like protein) showed significantly elevated levels of expression when grown in media containing uric acid and to a lesser extent hypoxanthine or xanthine. The addition of easier to utilize carbon and nitrgen sources reduced this expression to background levels even when uric acid was present in the media.

8. I was able to demonstrate that the *S. meliloti guaD2* gene shows significantly elevated levels of expression when grown in media containing guanine. I was also able to demonstrate that the high level of expression is reduced when an alternative carbon and nitrogen source was provided. It was also shown that *guaD2* showed a smaller though significant increase in expression when grown in media containing xanthine

or hypoxanthine.

9. I created a *S. meliloti* strain in which *xdhA1* was mutated. I was able to show that this strain was unable to grow on hypoxanthine or xanthine as a sole carbon and nitrogen source. I was able to show that this strain grows normally when glucose and ammonium chloride were used as carbon and nitrogen sources. I was also able to show that the mutation in this gene resulted in a loss of activity in xanthine oxidase / xanthine dehydrogenase.

10. I created a *S. meliloti* strain in which *xdhC* was mutated. I was able to show that this strain was unable to grow on hypoxanthine or xanthine as a sole carbon and nitrogen source. I was able to show that this strain grows normally when glucose and ammonium chloride were used as carbon and nitrogen sources. I was also able to show that the mutation in this gene resulted in a loss of activity in xanthine oxidase / xanthine dehydrogenase.

11. I also demonstrated that mutating the putative LysR transcriptional regulator SMb21291 in *S. meliloti* resulted in a loss of ability to grow on hypoxanthine or xanthine as sole carbon and nitrogen sources. I was also able to demonstrate that mutating this gene results in an altered pattern of xanthine oxidase / xanthine dehydrogenase activity in native protein gels assays. The altered pattern shows a reduction in levels of staining (a measure of activity) for the enzyme composed of subunits encoded by *xdhA1* and *xdhB1* and a concurrent increase in activity for the enzyme formed from units encoded by *xdhA2* and *xdhB2*

12. I was able to transduce the *xdhA1* reporter fusion into the SMb21291 (LysR) mutant strain allowing me to measure the impact of the LysR mutation on *xdhA1* expression. As a result of the LysR mutation *xdhA1* expression was demonstrated to decline

significantly. I was able to show that complementation of the LysR mutation restored normal levels of *xdhA1* expression.

13. I was able to transduce the SMb21285 reporter fusion into the SMb21291 (LysR) mutant strain allowing me to measure the impact of the LysR mutation on SMb21285 expression. As a result of the LysR mutation SMb21285, expression was demonstrated to decline significantly. I was able to show that complementation of the LysR mutation restored normal levels of SMb21285 expression.

14. I was able to transduce the SMb21284 (uricase like protein) reporter fusion into the SMb21291 (LysR) mutant strain allowing me to measure the impact of the LysR mutation on expression of SMb21284. As a result of the LysR mutation, SMb21284 expression was demonstrated to declined to background levels. I was able to show that complementation of the LysR mutation restored normal levels of SMb21284 expression.

15. I was able to transduce the *xdhA2* reporter fusion into the SMb21291 (LysR) mutant strain allowing me to measure the impact of the LysR mutation on *xdhA2* expression. As a result of the LysR mutation, *xdhA2* expression was demonstrated to increase significantly. Interestingly the level of expression of *xdhA2* was not elevated in response to the presence of hypoxanthine or xanthine in the growth media.

16. I was able to test the growth phenotype/ *lacZ* expression levels of the purine degradation gene mutant and / or fusion strains on different purines.

17. I was able to test the above mentioned purine degradation genes for elevated expression relative to the wild-type strain on representative compounds from a variety of classes of biomolecules. I was thus able to demonstrate that these genes only show elevated gene expression when grown in media containing purines.

Chapter 1

Literature Review

1.1 Introduction

Climate change is widely acknowledged to be one of the greatest scientific and societal challenges facing humanity in the twenty-first century. The activities contributing to climate change cover a wide range of activities associated with modern human society ranging from transportation to food production. Central to these activities are the ever increasing demands for energy required to support an increasing global population. This energy ranges in form from fossil fuels used for transportation to the food energy required to support an ever increasing global population. In agriculture these issues intersect. Modern agricultural production not only relies upon fossil fuels to power farm machinery but also large energy inputs associated with the production of nitrogen containing fertilizers required to

support the growth of crops (Bernstein *et al.*, 2007). It is this use of commercial fertilizers that impacts climate change on two levels. Not only does the production of such fertilizers require large amounts of energy, the application of these fertilizers in the field leads to increased levels of reactive nitrogen in the environment. Reactive nitrogen, as nitrous oxide for example, is associated with multiple environmental challenges including contamination of groundwater, ozone pollution in the lower atmosphere and an enhanced greenhouse effect (Mosier, 2008).

One approach to reducing the environmental impacts of modern agricultural production could be to take advantage of systems already present in the natural world which solve these same problems with a biological approach refined over time by evolution. One such example is the symbiotic association between legumes and nitrogen-fixing bacteria. Such bacteria are able to supply their host plants with the required amount of nitrogen in a completely biological process seemingly free of the negative impacts caused by the inefficient application of industrial fertilizers. If this biological approach to nitrogen fixation could be more widely used, agricultural production could be placed on a more sustainable path. To accomplish this a greater understanding of the genetic or molecular processes involved in these biological systems is required.

One such aspect of understanding how nitrogen fixation is achieved in nature is to understand how nitrogen-fixing symbiotic bacteria satisfy the nutritional requirements associated with surviving in the soil and colonizing a host plant. Achieving this symbiosis requires the coordinated action of multiple bacterial genes whose activities range from causing morphological changes in the host plant leading to the formation of root nodules to activating metabolic genes associated with obtaining sufficient carbon, nitrogen and other key nutri-

ents. It is this latter area that is the subject of this thesis. In particular the procuring of nitrogen (and carbon) by *Sinorhizobium meliloti* through the degradation of purines.

1.2 *Sinorhizobium meliloti*

Sinorhizobium meliloti is a Gram-negative bacterium of the family Rhizobiaceae. Among this family can also be found *Azorhizobium*, *Allorhizobium*, *Rhizobium* and *Mesorhizobium*. Members of this family are characterized by their ability to fix atmospheric nitrogen (N_2) within nodules located on the roots of leguminous plants. *S. meliloti* perform nitrogen fixation as part of a symbiotic relationship with alfalfa (*Medicago sativa*). This relationship greatly facilitates the growth of alfalfa plants by providing a source of nitrogen without the addition of exogenous nitrogen in the form of fertilizer or mineral nitrogen from the surrounding soil (van Berkum and Eardly, 1998).

The genomic sequence of *S. meliloti* has been determined and published. It has three separate replicons. These consist of a 3.65 Mb chromosome and two large megaplasmids designated pSymA (1.35 Mb) and pSymB (1.68 Mb)(Barnett *et al.*, 2001; Capela *et al.*, 2001; Finan *et al.*, 2001; Galibert *et al.*, 2001). The chromosome has been shown to contain the bulk of genes involved in basic cellular metabolism, the so-called housekeeping genes, such as those involved in the synthesis of nucleotides, amino acids, carbon metabolism (gluconeogenesis and glycolysis) vitamin synthesis and chemotaxis.The pSymA megaplasmid has a specialization in the sense that it houses the genes required for the key symbiotic process of nitrogen fixation. These genes include those that are involved in the process of nodule formation (the Nod factors) which are fundamental to the unique host-plant symbiosis. The pSymA megaplasmid also contains the genes required for nitrogen fixation

(the *nif* and *fix* genes). The pSymB megaplasmid appears to be specialized in allowing the bacterium to transport and utilize many useful/essential compounds from the environment surrounding the bacterium. In addition pSymB carries genes required for the synthesis of polysaccharides such as succinoglycan and galactoglucan (the *exo* and *exp* genes)

1.3 Rhizobia-Legume Symbiosis

Leguminous plants secrete flavonoid molecules into the rhizosphere surrounding their roots. It is these flavonoid molecules which have been shown to cause the induction of the *nod* genes in the bacteria. Upon induction the bacterial *nod* genes result in the production of lipochitooligosaccharides, also known as Nod factors (Lerouge *et al.*, 1990; Schultze and Kondorosi, 1998). It is these Nod factors that cause the curling of the plant root hair and the envelopment and trapping of the bacteria. Receptors for Nod factors have been identified in legumes such as *M. truncatula* (LTK3 and LYK4) and in *Lotus japonicus*. In each case the Nod factors bind via a LysM domain and play a role in controlling the entry of the bacteria into the plant (Ardourel *et al.*, 1994; Limpens *et al.*, 2003; Madsen *et al.*, 2003; Parniske and Downie, 2003; Radutoiu *et al.*, 2003). Following the curling of the root hair and envelopment of the bacteria there is an invagination of the plant cell wall and the resulting formation of an infection thread that propagates between the cells of the root cortex. Cortical plant cells located near the infection site undergo cell division leading to the production of the nodule primordium. The rhizobia move down the infection thread by means of bacterial cell division until they reach the end of the infection thread where they are released into the newly formed root nodule. Within the root nodule the bacteria

are surrounded by the peribacterial membrane and undergo differentiation into mature bacteroids which fix atmospheric nitrogen (Schultze and Kondorosi, 1998).

Fixing atmospheric nitrogen is an extremely energy demanding process requiring the input of 16 ATP molecules for each N_2 molecule that is reduced (Arwas *et al.*, 1985; El-Din, 1992; Engelke *et al.*, 1987; Finan *et al.*, 1983; Humbeck and Werner, 1989; Lafontaine *et al.*, 1989; Ronson *et al.*, 1981). It has been experimentally verified that rhizobial mutants deficient in the ability to transport C4-dicarboxylates such as fumarate, malate and succinate are unable to establish nitrogen fixing nodules. It has been suggested that these plant derived dicarboxylates provide the energy required by the bacterial symbionts for nitrogen fixation. The use of these dicarboxylates by the bacteria does however present an interesting metabolic challenge. It is believed that these carbon sources are oxidized in the tricarboxylic acid (TCA) cycle which is an aerobic metabolic pathway while at the same time the nitrogenase enzyme which catalyzes the reaction converting N_2 to NH_3 undergoes an irreversible denaturation in the presence of molecular oxygen thus these two fundamental metabolic processes within the nodule would seem to be in direct conflict (Driscoll and Finan, 1993; Jackson and Dawes, 1976; Senior and Dawes, 1971). To overcome this physiological contradiction the oxygen levels within the nodule must be controlled in such a way as to create a micro-aerobic environment where low oxygen concentrations are maintained in a manner that allows the TCA cycle to operate yet the oxygen levels are not so high as to inhibit the nitrogenase enzyme. This balance is even more remarkable when one considers that under these microaerobic conditions it would be expected that the NADPH/NADP and NADH/NAD ratios would be increased resulting in an expected inhibition of several steps in the TCA cycle such as those involving the enzymes citrate synthase, isocitrate dehydrogenase, and α-ketoglutarate dehydrogenase (Jackson and Dawes, 1976; Senior and

Dawes, 1971).

Maintaining such a low yet optimal oxygen concentration is achieved through several mechanisms. These include the production of leghaemoglobin within the nodule, a protein that binds O_2 in a reversible manner. In addition the cortex of the nodule limits the rate of oxygen diffusion into the nodule such that the rate at which O_2 enters is balanced by the rate at which it is used in respiration and / or bound by leghaemoglobin. To deal with the potential impact of low oxygen concentration on cellular respiration the Rhizobia have a branch of the respiratory pathway that terminates with a high affinity oxidase. This allows for the continued production of ATP under the micro-aerobic conditions that exist within the nodule (Anraku, 1988; Poole, 1983; Witty *et al.*, 1986).

1.4 Purine Metabolism

Within the cell there is a continuous synthesis and breakdown of purine nucleotides. Purine synthesis is typically accomplished via two methods, biosynthesis by the cell from purine precursor molecules or by the salvage of preformed purine nucleotides and nucleosides from the cell's external environment produced via the breakdown of RNA associated with dead cells or those excreted from live cells. The salvage of these preformed purine nucleotides and nucleosides from the environment would seem to be a more efficient approach since excess cellular levels of these molecules would not be expected to occur (Burton, 1994; Nygaard, 1983).

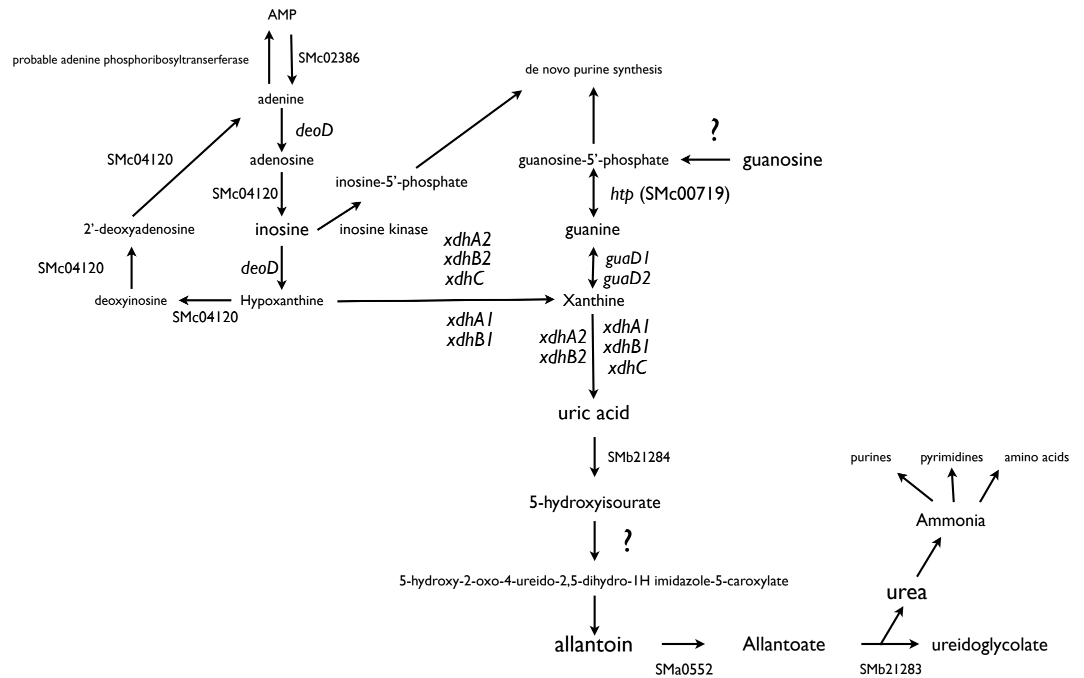

Figure 1.1: Pathway for purine degradation showing genes predicted to code for key pathway enzymes. XO / XOH = xanthine oxidase / xanthine dehydrogenase.

Many bacteria are able to utilize purines for more than encoding genetic information (Vogels and van der Drift, 1976). Purines can be utilized as a source of both nitrogen and carbon. Breaking the glycolytic bond results in the release of both a purine base, a nitrogen/carbon source, and a pentose sugar (a carbon source). The purine base can be further degraded into carbon dioxide and ammonia, the actual nitrogen source derived from purines (Fig. 1). The purine catabolic pathway begins with the deamination of adenine or guanine to hypoxanthine (HX) or xanthine (X) by a deaminase. This is followed by the conversion of HX to xanthine catalyzed by xanthine oxidase/dehydrogenase (XO/XDH) and the subsequent oxidation of xanthine to uric acid which also results in the reduction of NAD. Uric acid is further degraded to urea and ultimately to ammonia and carbon dioxide (Vogels and van der Drift, 1976). A key step in this pathway is the conversion of the purine base to hypoxanthine. It is at this step that the cell can decide to continue the degradation of the base to ammonia and carbon dioxide or utilize the hypoxanthine to produce more base through the salvage pathway depending on the availability of nitrogen within the cell. This situation has been demonstrated in *Bacillus subtilis* in which the purine degradation genes have been shown to be induced when only purines are available as a source of nitrogen. When there is an excess of nitrogen the purine catabolic genes are repressed and the salvage pathway genes activated (Christiansen *et al.*, 1997).

When the cell has an adequate supply of nutrients and is growing normally, the normal source of purine and pyrimidine bases is the relatively unstable mRNA. When nutrient/growth conditions become less favourable however due to a lack of carbon or nitrogen sources more stable RNA molecules such as rRNA can also be utilized as carbon/nitrogen sources (Mandelstam, 1960; Nath and Koch, 1971).It has been shown that under conditions of nutrient limitation*E. coli* will accumulate hypoxanthine as a result of rRNA degradation.

It is believed that the accumulated hypoxanthine store acts as a source of nitrogen that can be utilized under nitrogen limiting conditions (Ben-Hamida and Schlessinger, 1966; Dawes and Ribbons, 1965; Kaplan and Apirion, 1975; Mandelstam and Halvorson, 1960; Okamura *et al.*, 1973).

1.5 Xanthine Oxidase/Dehydrogenase

A critical enzyme in both the purine salvage and purine catabolic pathways is xanthine oxidase/xanthine dehydrogenase (EC 1.1.3.22/1.1.1.204). The functional form of this protein consists of a heterotetramer composed of two alpha and two beta subunits and is a molybdo/iron sulfur flavoprotein (Xi *et al.*, 2000). This enzyme is responsible for catalyzing the conversion of both hypoxanthine to xanthine as well as the oxidation of xanthine to uric acid. It is this second conversion that commits the cell to the degradation of the hypoxanthine molecule as it removes the possibility of the hypoxanthine being salvaged. XO/XDH is initially synthesized in the NAD-dependent dehydrogenase form and can be easily and reversibly converted to an oxidase form by the oxidation of sulfhydryl residues. It can also be irreversibly converted to the oxidase form by proteolysis (Amaya *et al.*, 1990; Corte and Stirpe, 1968, 1972; Nishino, 1994; Stirpe and Della Corte, 1969). The dehydrogenase form of the enzyme has been shown to be dependent on residues Phe-549, Trp-336, Arg-427 and Arg-335 (Kuwabara *et al.*, 2003).

In the *S. meliloti* genome there are two loci responsible for the synthesis of XO/XDH both of which are located on the pSymB megaplasmid. The *xdhA2* and *xdhB2* genes are

located within a single operon that also encodes the *bdhA* gene which plays a crucial role in the PHB degradation cycle. A second copy of these genes, referred to as *xdhA1* and *xdhB1*, resides on a second region of pSymB in a 20 kb cluster of genes predicted to be involved in purine and pyrimidine catabolism and salvage (Finan *et al.*, 2001). This second *xdh* gene locus also consist of a third gene located in the same predicted operon as *xdhA1* and *xdhB1* (Aneja and Charles, 1999). This gene has been named *xdhC* and is not believed to encode an active *xdh* subunit. Instead it is believed that this gene encodes a chaperone that may play a role involving the insertion of the molybdopterin cofactor into the XO/XDH molecule.

The activity of the XO/XDH genes is regulated by the levels of nitrogen available to the cell. In *B. subtilis* under conditions of excess nitrogen these genes are repressed while they are induced to a high level of activity when there is a limited supply of nitrogen (Schultz *et al.*, 2001). It has also been shown in *B. subtilis* that *glnA* mutants which have a defective glutamine synthase, the XO/XDH activity is high in the presence of glutamine (Leimkuhler *et al.*, 1998). This suggests that XO/XDH can be regulated through catabolite repression and is dependent on GlnA. Experimental evidence from *Rhodobacter capsulatus* has shown that XO/XDH are not regulated by the nitrogen dependent regulators RpoN (an alternative sigma factor) and NtrC (part of the Ntr nitrogen level sensing system) (Magasanik, 1996).

1.6 Poly-beta-hydroxybutyrate (PHB)

Polyhydroxyalkanoates (PHA) such as poly-beta-hydroxybutyrate (PHB) are polyesters that can be biologically degraded. PHB is known to accumulate in a variety of bacteria when there is an excess supply of carbon but a limited supply of O_2, N or P available for growth. PHB accumulates intracellularly and in the rhizobia can account for more than half of the cell dry weight. These large molecules can be utilized by the cell as a source of carbon and energy during periods of carbon limitation (Anderson and Dawes, 1990; Tombolini and Nuti, 1989). PHA polymers can be composed of chains of varying lengths including those categorized as short chain (3 to 5 carbons), medium length (4-14 carbons) or long chain (15 or more carbons). PHB is composed of short chain hydroxyalkanoic acids (homopolymers of 3-hydroxybutyrate) in 23,000 to 25,000 repeating units (Byrom, 1987).

PHA and PHB molecules have been investigated as an alternative to petroleum-based plastics as these bacterially derived polymers are heat resistant and biodegradable. One example of such a product was BIOPOL, a product of Monsanto. This product was composed of poly-(3-hydroxybutyrate-*co*-hydroxyvalerate) (Anderson and Dawes, 1990). All rights to the production of BIOPOL are currently held by Metabolix Inc (Cambridge, MA USA). Metabolix is currently marketing a PHA derived bioplastic under the trade name Mirel.

1.7 PHB Plays a Role in Bacterial Survival

PHB has been demonstrated to play an important role in the survival of bacteria. This has been demonstrated for soil bacteria such as *Azospirillum brasilense*, a soil bacterium

associated with plant roots, and *Pseudomonas* which are known to produce PHAs. It has been shown that *Azospirillum brasilense* mutants lacking the ability to synthesize PHB have a reduced ability to survive environmental stresses as varied as osmotic shock, elevated temperatures, desiccation and UV exposure (Tal and Okon, 1985; Kadouri *et al.*, 2003). It has also been demonstrated in *Pseudomonas oleovorans* that a PHB degradation mutant could not endure heat or ethanol exposure as well as PHB degrading strains (Ruiz *et al.*, 2001). Accumulation of PHA molecules have also been shown to improve the survival of bacteria in aquatic environments (Lopez *et al.*, 1995).

The mechanism by which PHB accumulation enhances bacterial survival is not entirely understood but it has been suggested that PHB has an effect on the stress response of Gram-negative bacteria. In *P. oleovorans* degradation of PHB leads to an increase in cellular levels of both ATP and guanosine tetraphosphate (ppGpp). ppGpp is known to induce the production of the alternative sigma factor RpoS which itself is involved in helping cells to adapt to stationary phase. It is also known that ppGpp can alter the specificity of the RNA polymerase holoenzyme for particular promoter regions by altering the stability of the RNA polymerase-DNA interaction and thus may allow alternative patterns of gene expression under stress conditions.

1.8 The Role of PHB in Rhizobial Physiology

All Rhizobia are capable of accumulating PHB within the cell in the free living part of their life cycle but not all Rhizobia can accumulate PHB as bacteroids. In *S. meliloti* PHB is stored/polymerized in the free living cell, is still observed as the cells multiply down the infection thread but disappears from within the differentiated bacteroid (Hirsch *et al.*,

1983, 1982; Paau *et al.*, 1980). This scenario is not observed with all Rhizobia. For example *R. etli* and *Rhizobium* NGR234 (van Slooten *et al.*, 1990) and *B. japonicum* (McDermott *et al.*, 1989) can amass PHB in the bacteroid state. Several physiological roles have been suggested for PHB. These include acting as a reservoir for reducing power which would allow the cell to continue respiration when oxygen levels are low as reduced oxygen levels would lead to the elevation of NADPH levels which can be used in the synthesis of PHB thus preventing the inhibition of TCA cycle enzymes (Dunn, 1998; Senior *et al.*, 1972; Senior and Dawes, 1971; Walshaw *et al.*, 1997). It is also believed that PHB may act as an energy source for nitrogen fixation during the night when the supply of photosynthate from the plant is reduced (Bergersen *et al.*, 1991; Gerson *et al.*, 1978). In addition the energy stored in PHB may be used by the bacteria following nodule senescence and the release of the bacteroids into the soil.

The amount of carbon and energy that is stored as PHB in the cell may provide the free-living bacteria with a source of carbon and energy required during the process of establishing the symbiosis with the plant. In entering the rhizosphere around the plant roots the bacteria transition from a carbon limited environment in the soil to a carbon rich area in close proximity to the plant roots (van Elsas and van Overbeek, 1993). In the rhizosphere it would be expected that the bacteria take advantage of the opportunity to accumulate and store surplus carbon as PHB. This PHB could either be used for survival if the bacteria reenter a carbon poor area of the soil or as a supply of carbon and energy that could allow the rhizobia to outcompete other non PHB accumulating bacteria in the soil. It is believed that the stored PHB provides an energy source which allows the rhizobia to power its establishment of a symbiosis with the plant, in particular the energy needed for the bacteria to divide down the infection thread to the nodule (Charles *et al.*, 1997).

This theory suggesting that PHB is used to power the establishment of the infection is somewhat undermined by genetic studies using *S. meliloti* where mutants that are unable to synthesize or utilize PHB are still able to establish an infection within the roots of alfalfa (Aneja and Charles, 1999; Povolo *et al.*, 1994). These findings still leave open the theory that PHB helps the rhizobia in the soil environment. In fact it has been established that the production and utilization of PHB gives the rhizobia a competitive advantage in establishing symbiosis (Ratcliff *et al.*, 2008; Trainer and Charles, 2006; Dawes, 1965).

1.9 *S. meliloti* PHB Metabolism

It has been proposed that in *S. meliloti* PHB metabolism follows a metabolic cycle consisting of both the synthesis and degradation of the polymer (Trainer and Charles, 2006; Senior and Dawes., 1973). The process of synthesizing PHB begins with the condensation of two molecules of acetyl-CoA to form acetoacetyl-CoA through the action of a beta-ketothiolase (*phbA*). This step is followed by the reduction of the AA-CoA by an acetyl-CoA reductase (*phbB*) forming 3-hydroxybutyryl-CoA which is then polymerized by PHB synthase (*phbC*) to form PHB. The degradation part of the cycle begins with the depolymerization of PHB to free HB molecules. The HB monomers then undergo oxidation to acetoacetate via the action of acetoacetyl-CoA synthetase (*acsA2*) to acetoacetyl-CoA. The acetoacetyl-CoA can then be split by beta-ketothiolase to produce two molecules of acetyl-CoA which can be further oxidized in the TCA cycle or as a substrate for the re-synthesis of PHB.

1.10 The *S. meliloti bdhA* Gene

The *bdhA* gene codes for the enzyme 3-hydroxybutyrate dehydrogenase (BdhA) (EC 1.1.1.30) which catalyzes the oxidation of hydroxybutyrate to acetoacetate in the PHB cycle (Charles *et al.*, 1997). BdhA is a short-chain dehydrogenase / reductase (SDR) and a member of a superfamily of similar enzymes and found only in PHB accumulating bacteria (Aneja and Charles, 1999). These enzymes are NAD(P)(H)-dependent and act upon a diverse array of organic substrates including aromatic compounds, steroids and carbohydrates. The SDRs have been classified into five families referred to as classical, extended, intermediate, divergent and complex (Persson *et al.*, 2008; Jörnvall *et al.*, 1995; Kallberg *et al.*, 2002). BdhA belongs to the intermediate family. In the *S. meliloti* genome the *bdhA* gene is the first in an operon that also includes the *xdhA2* and *xdhB2* genes which code for two of the active subunits of XO/XDH involved in purine catabolism.

Characterization of *bdhA* by Aneja and Charles (Aneja and Charles, 1999) demonstrated that this gene has two transcriptional start sites designated S1 and S2 but the promoters associated with these sites are not known. A sigma 54 consensus sequence has been identified between S1 and S2 but does not seem to be required for the utilization of HB as an *rpoN S. meliloti* mutant is still able to utilize HB. It has been demonstrated that disrupting the *bdhA* gene has a polar effect on the downstream *xdhA2* and *xdhB2* genes suggesting that these genes all share a common promoter and belong in the same operon.

Expression of *bdhA* has been demonstrated to change with growth phase. Low levels of expression have been demonstrated during the lag phase followed by an increasing rate through log phase and reaching a peak at stationary phase (Aneja and Charles, 1999;

Hofmann *et al.*, 2000a). The level of expression of *bdhA* has been shown to increase in the presence of biotin, a compound known to be exuded by alfalfa and present in the rhizosphere of the plant. Biotin is a vitamin co-factor, involved in the binding of carbon dioxide in carboxylase enzymes. *S. meliloti* does not require biotin for growth but its growth is stimulated when supplemental biotin is provided (Hofmann *et al.*, 2000b; Dunn, 1998; Encarnacion *et al.*, 1998; Sierra *et al.*, 1999; West and Wilson, 1939). Studies with *bdhA-lacZ* fusion strains of *S. meliloti* have demonstrated that as little as 0.4 nM of biotin results in an increase in *bdhA* expression that becomes 4-5 times higher at a biotin concentration of 4 nM (Hofmann *et al.*, 2000a). It is assumed that the presence of biotin in the rhizosphere would result in a significant increase in *bdhA* expression and the resulting PHB degradation would fuel the growth of *S. meliloti* and allow for the colonization of the plant.

1.11 Symbiotic and Competitive Phenotypes of PHB Mutants of *S. meliloti*

The ability of *S. meliloti* to successfully compete in the rhizosphere and colonize a plant can be impacted by the bacteria's ability to synthesize and use PHB reserves. *S. meliloti* with mutated *bdhA* do not appear to be prevented from establishing a symbiosis. Mutations in *S. meliloti* PHB synthesis genes such as *phbC* display a symbiotic defect, visible on the host plant, characterized by reduced nodule count and a temporal lag in the appearance of root nodules compared with alfalfa inoculated with wild-type *S. meliloti* (Aneja *et al.*, 2005). In terms of competitiveness in the rhizosphere plants co-inoculated with wild type *S. meliloti* and *bdhA* mutants (a PHB degradation mutation) or with *phbC* mutants (a PHB synthesis mutation) and wild-type display a clear advantage for the wild-type in

establishing nodulation compared with either mutant type (Willis and Walker, 1998). Of the two mutant types the *phbC* mutants show a reduced ability to compete in establishing symbiosis compared with the *bdhA* mutant as well as resulting in host plants with reduced dry mass. It has also been demonstrated that when grown in batch cultures with varying carbon supply strains with mutations in PHB cycle genes show a similar reduction in competitive ability as that described above with the *phbC* mutants showing the greatest loss of competitiveness (Aneja *et al.*, 2005).

1.12 The Relationship between EPS Production and PHB Synthesis

When there is an excess level of carbon source but a limitation in nitrogen or phosphorous both PHB and the EPS succinoglycan will be synthesized. If the ability to synthesize PHB is lost due to a mutation in *S. meliloti* the ability to synthesize succinoglycan is also eliminated under conditions that would normally favour the synthesis of both (Aneja, 1999). Phenotypic analysis of mutants suggests a common regulatory mechanism for both synthetic pathways as *bdhA* mutants produce mucoid colonies while *phbB* and *phbC* mutants produce non-mucoid colonies on carbon rich media (Aneja, 1999).

There have been several ideas put forward to explain why the inability to synthesize PHB leads to reduced succinoglycan synthesis. It has been suggested that acetyl phosphate acts as a common regulatory element between these two pathways. It has been shown that in *Synechococcus* acetyl phosphate is involved in PHB synthase regulation in a manner that is concentration dependent (Miyake *et al.*, 1997). In addition it was demonstrated in this

study that phosphotransacetylase, which catalyzes the conversion of acetyl-CoA to acetyl phosphate is active only during conditions that favour the accumulation of PHB (Miyake *et al.*, 1997). Much attention has focused on the role of *S. meliloti* ChvI in this system as it has been demonstrated that ChvI autophosphorylates in the presence of acetyl phosphate *in vitro* (Cheng and Walker, 1998). PHB synthesis mutants can excrete excess acetyl-CoA resulting in low levels of acetyl phosphate which may play a role in the regulation of succinoglycan production by ExoS-ChvI (Aneja, 1999).

1.13 How Does *S. meliloti* Colonize the Rhizosphere of Alfalfa? A Proposed Model

Establishment of a symbiosis with alfalfa by *S. meliloti* requires that these bacteria be able to compete effectively with other bacteria surrounding the plant roots. The arrangement of genes in the *bdhAxdhA2xdhB2* operon suggests a linkage of the PHB and purine degradation cycles. It is proposed that this operon allows for enhanced competitiveness of *S. meliloti* under conditions where there may be an excess supply of carbon but a limited amount of usable nitrogen. Under these conditions it would be expected that some of the excess carbon will be stored as PHB. In this situation bacteria would also be experiencing growth limitations due to nitrogen deficiency and would be expected to degrade rRNA/RNA to produce purines such as hypoxanthine which would then accumulate and be fed into the purine degradation pathway (in which *xdhA* and *xdhB* play a key role) resulting in usable nitrogen being made available to the cell. It is also interesting to note that this mixed function operon is induced by biotin, a molecule known to be exuded from alfalfa roots suggesting a model whereby the plant acts to induce PHB and purine degra-

dation in the bacteria helping to fuel the early stages of the bacteria's colonization efforts (Charles, 2002).

1.14 The Genetic Toolbox of *S. meliloti*

S. meliloti has a variety of well developed genetic tools which facilitate the study of gene function in this bacterium. One such tool is the now completely sequenced *S. meliloti* genome (Galibert *et al.*, 2001). In addition this genetic toolbox includes transposons such as Tn5 which can be used to create mutations via insertions in genes as well as reporter fusions linked to the transposon. Transposons can also be used to transfer mutations from one *S. meliloti* strain to another to create new strains with more than one mutation (Finan *et al.*, 1984; Charles and Finan, 1990). Phage such as ϕM12 can also be used to transduce mutations between *S. meliloti* strains. There are plasmids that act as suicide vectors, such as pK19*mob*, that lack the ability to replicate in *S. meliloti* but can be used to create targeted mutations when carrying fragments of genes. Once in the cell the plasmid can undergo recombination thus disrupting and mutating the target gene. Plasmids have been developed that allow for the creation of fusions to the promoter region of a gene while simultaneously creating a mutation in that gene (Cowie *et al.*, 2006). There are helper and mobilizing strains that allow the transfer of plasmids into *S. meliloti* by triparental mating that have antibiotic resistances distinct from *S. meliloti* that allow the new *S. meliloti* strain to be relatively easily differentiated from the *E. coli* helper strain (Finan *et al.*, 1986).

1.15 LysR Transcriptional Regulators

LysR transcriptional regulators constitute a well characterized family of highly conserved and widespread transcriptional regulators. These regulators can be found not only in bacteria but in the archaea and eukaryotes as well (Pérez-Rueda and Collado-Vides, 2001; Sun and Klein, 2004; Stec *et al.*, 2006). From an initial discovery of some nine similar regulatory proteins in *Salmonella enterica* serovar Typhimurium, *Escherichia coli*, *Rhizobium* spp and *Enterobacter cloacae* the family has been expanded through the identification of amino acid and "dot matrix" comparisons to identify many putative LysR type regulators (LTTRs) (Henikoff *et al.*, 1988). The initial grouping of these proteins into a common family was done based on sequence similarity and the conservation of DNA-binding domains. The best characterized member of the LTTR family is the *LysR* gene responsible for regulating *lysA*, a gene coding for diaminopimelate decarboxylase, the enzyme that catalyzes the decarboxylation of diaminopimelate to lysine (Stragier and Patte, 1983; Stragier *et al.*, 1983).

Early study of LTTRs suggested that these proteins behave as transcriptional regulators of a single gene usually transcribed in the opposite orientation as the LTTR and acting as a negative regulator of the target gene (Parsek *et al.*, 1994; Schell, 1993; Lindquist *et al.*, 1989). Subsequent research has lead to a reevaluation of this simple concept to one where LTTRs are regarded as global regulators, that is a single LTTR is capable of regulating numerous operons. These LTTRs can act as activators or repressors. In addition it has been shown that they can act in the regulation of operons as well as single genes. It has been demonstrated that in addition to acting on genes in a divergent manner they can also be located on a different part of the chromosome from their target gene or operon (Hernández-Lucas *et al.*, 2008; Heroven and Dersch, 2006). It has also been discovered

that these proteins require a co-inducer to act as an activator or a co-repressor to act as a repressor (Picossi *et al.*, 2007; van Keulen *et al.*, 2003; Celis, 1999).

The fact that LTTRs are so conserved in such a diverse number of bacteria suggests that they play a role in the regulation of genes with a highly diverse set of functions. These include cell division, quorum sensing, motility, nitrogen fixation, stress responses, and metabolism among others. It is believed that the LTTRs were acquired in such a diverse range of bacteria via horizontal gene transfer and their common ancestry is suggested by their conserved sequence similarity which gives rise to similar functional domains (Sperandio *et al.*, 2007; Lu *et al.*, 2007; Byrne *et al.*, 2007; Russell *et al.*, 2004; Kim *et al.*, 2004; Cao *et al.*, 2001; Deghmane *et al.*, 2002, 2000; Kovacikova and Skorupski, 1999). One notable example of structural conservation in this protein family is the presence of the DNA binding helix-turn-helix (HTH) motif near the N-terminus which is typically about 330 amino acids in size. This structure appears to be a universal characteristic of LTTRs. The position of the HTH domain exclusively at the N-terminus of these proteins is in itself interesting in that usually in proteins with HTH residues near the N-terminus the proteins act as repressors while in the case of the LTTRs the proteins can act as either activators or repressors (Aravind *et al.*, 2005; Huffman and Brennan, 2002; Pérez-Rueda and Collado-Vides, 2001).

A characteristic of genes encoding LTTRs is a high G+C content with a characteristic Lys to Arg ratio in the coded protein. It has also been noted that many LysR encoding genes are located on transmissible regions of bacterial genomes lending further support to the idea that the appearance of these proteins in a wide variety of bacteria has occurred due to horizontal gene transfer (Viale *et al.*, 1991; Henikoff *et al.*, 1988).

The highly conserved nature of the amino acid composition and resulting secondary structure have played a key role in helping to identify members of the LTTR family. Among the most highly conserved region of these proteins is the approximately 60 residues that make up the region of these proteins that interacts with the major groove of DNA, the HTH domain. When one looks at the opposite end of the protein however a different picture emerges. In this region (the C-terminal end) there is little amino acid sequence conservation (Stec *et al.*, 2006). What can be seen though is a common general structure consisting of two domains (referred to as alpha and beta) that are interconnected in such a manner that they form a hinge like structure believed to allow for the binding of a co-inducer molecule. This hinge region joins to the HTH DNA binding domain via another hinge like structure. Mutagenesis studies have shown that changes in the sequence of the C-terminal result in loss of function for LTTRs. The binding of a co-inducer or co-repressor to the C-terminal region is believed to cause a change in the tertiary structure of the protein that alters the affinity of the protein (presumably the HTH region) for the target DNA. Mutational analysis has also shown that changes in an additional C-terminal region can also result in changes to DNA binding behaviour such that the need for a co-inducer to bind is lost. This effect has been demonstrated in well known LTTRs such as NodD (residues 95, 123 and 154) and AmpR (residues 102 and 135) among others (Jørgensen and Dandanell, 1999; Cebolla *et al.*, 1997; Burn *et al.*, 1989).

The actual sites at which LTTRs bind to DNA are typically in the intergenic region between the gene coding for the LTTR and the target gene itself. Typically these regions have been shown to be located from positions -40 to -20 or -35 to +20 with respect to the target genes transcriptional start site as well as at the -55 position when associated with DNA bending (Porrúa *et al.*, 2007; Lochowska *et al.*, 2001; Belitsky *et al.*, 1995). However

it may not be comprehensive enough to make such sweeping generalizations as LTTR DNA binding sites have also been shown to be located at positions ranging from -218 to +350 (such sites being referred to as internal binding sites) (Viswanathan *et al.*, 2007; Wilson *et al.*, 1995). There are also examples of palindromic sequences to which LTTRs bind such as the LTTR box seen upstream of the *nod* gene in *Rhizobium* spp which has been dubbed the 'Nod-box". Many LTTRs are functionally active as tetramers and are able to cover a region of 50-60 bp on DNA (Muraoka *et al.*, 2003b,a). Such a pattern is suggestive of the ability of these proteins to bind several locations on a DNA molecule with the particular region bound being dependent on whether the inducer molecule is bound to the protein. It is believed that in the presence of an inducer molecule the two dimeric forms of the LTTR which respectively bind the RBS and an alternative binding site are able to link together to form an active tetramer via the bending of the DNA (Tropel and van der Meer, 2004).This tetramer is then believed to form a complex with RNA polymerase which would lead to the initiation of transcription. The degree of DNA bending in the known examples of LTTRs is quite variable ranging from as little as 9 degrees to as much as 100 degrees (van Keulen *et al.*, 1998). One of the best studied examples of this has been shown for OccR from the Ti plasmid of *Agrobacterium tumefaciens*. It has been demonstrated that OccR bound to DNA occupies a region from -80 to -28 bp upstream of its target gene. In the presence of the inducer octopine however this bound region shifts to the -80 to -38 position as the angle of bending is reduced. It has been shown that in the case of OccR there are five distinct regions at which the protein binds to DNA (Akakura and Winans, 2002a,b).

1.16 Review of LysR Transcriptional Regulators in Rhizobia

1.16.1 *nodD* and *syr* genes

By far the most studied group of LysR transcriptional regulators in the Rhizobia are the *nodD* and *syr* genes associated with induction of root nodule formation on the host plant. In this case it is not the LysR proteins that act directly upon the plant but the ability of these proteins to induce the genes associated with synthesizing the subunits of the Nod factors that makes the LysR proteins so critical in establishing the bacterial-plant symbiosis. The Nod factors are responsible for morphological and physiological changes in the plant that prepare the plant for the entry of the bacteria. These changes include the development of curled root hairs that envelope the bacteria allowing them to be taken up by the plant as well as initiating the development of root nodules in the plant by inducing mitosis and gene expression in the root cortex and pericycle. In essence it is the LysR proteins that play a critical role in coordinating the bacterial response to signals from the plant under conditions in the soil environment favourable to the establishment of symbiosis.

The Nod factors whose production is induced by the LysR NodD proteins are molecules characterized by a β-1,4-linked *N*-acetyl-D-glucosamine backbone typically with 4 or 5 residues. At the C-2 position of a terminal residue there is an acyl chain that varies in length and structure in a manner characteristic for different species of rhizobia (Cullimore *et al.*, 2001; Mergaert *et al.*, 1997; Demont *et al.*, 1993; Schultze *et al.*, 1992; Roche *et al.*, 1991; Lerouge *et al.*, 1990). Many rhizobia are characterized by having the *nodABC* genes and these three genes are required for the synthesis of the Nod factor backbone structure

(though there are slight species differences in these genes that result in species variation in the core Nod factor) (Martinez *et al.*, 1990; Goethals *et al.*, 1989). There are also other *nod* genes found across different rhizobia such as *nodI* and *nodJ* which are found in the same operon as *nodABC* as well as *nodFE*, *nodL* and *nodM* (van Rhijn and Vanderleyden, 1995). In addition to the basic *nod* genes found in all rhizobia there are additional species specific *nod* genes. These include the *nodPQ* genes of *S. meliloti* (Schwedock and Long, 1989) and the *nodZ* gene of *Bradyrhizobium japonicum* (Stacey *et al.*, 1994). It is also these species specific genes that give rise to the unique side chains characteristic of the Nod factors of different rhizobia. It is these differences, along with variations in the quantity of Nod factor produced by different rhizobial species and different inducer molecules that play key roles in explaining differences in host plant specificity among rhizobial species (Pacios Bras *et al.*, 2000; Kurkdjian, 1995; Heidstra *et al.*, 1994; Relić *et al.*, 1993; Mergaert *et al.*, 1993; Schultze *et al.*, 1992; Price *et al.*, 1992; Spaink *et al.*, 1991; Lerouge *et al.*, 1990; Kondorosi *et al.*, 1984).

Critical to the induction of rhizobial *nod* genes are signals received from the potential host plant. These signals take the form of flavonoids also known as 2-phenyl-1,4-benzopyrone derivatives. These compounds are characterized by having a three ringed structure consisting of two aromatic rings (referred to as A and B) and a heterocyclic pyrone or pryone ring (the C ring) (Harborne and Williams, 2001, 2000). The use of *nod* gene *lacZ* fusion assays has demonstrated the induction of *nod* genes in rhizobia by flavonoids such as luteolin and genistein (Cullimore *et al.*, 2001; Barbour *et al.*, 1991; Kape *et al.*, 1991; Bassam *et al.*, 1988; Peters and Long, 1988; Zaat *et al.*, 1987; Peters *et al.*, 1986; Mulligan and Long, 1985; Okker *et al.*, 1984). Numerous classes of flavonoids exist with variations of the basic structure such as anthocyanidines, flavonols, chalcones among others (Harborne

and Williams, 2001, 2000). Those flavonoids that have the ability to induce *nod* genes are hydroxylated at the C-4 and C-7 positions of the A ring and the specific flavonoids secreted by the host plant is a determinate of which rhizobia will be able to form a symbiosis with a particular species of host plant (Cunningham *et al.*, 1991; Rolfe, 1988).

There are several theories as to why plants produce and release such a variety of flavonoids (Peters and Long, 1988; Peters *et al.*, 1986; Firmin *et al.*, 1986). While it could be speculated that plants release these compounds with the express purpose of attracting symbiotic rhizobia, several other ideas have been proposed that make the release of flavonoids a byproduct of plant physiology and the attraction of rhizobia an indirect consequence of these other purposes. Some of these alternative ideas are based on the finding that plants secrete more flavonoids when nitrogen is limited so the release of flavonoids is the byproduct of the plant deaminating nitrogen containing organic molecules and then removal of the excess carbon rings released in this process (Coronado *et al.*, 1995). It has also been suggested that these compounds play a role in plant defence as it has been shown that plants which produce higher levels of flavonoids are less susceptible to attack by fungal pathogens such as *Botrytis* (Harborne and Williams, 2001, 2000). One interesting observation however is that plants release the optimal concentrations of flavonoids in the region of new emerging root hairs which just happen to be the part of the plant most favoured by the rhizobia (via the root curling response of the plant to Nod factors) for entry into the plant (Zuanazzi *et al.*, 1998). In addition to flavonoids there are some additional classes of compounds that have been shown to induce rhizobial *nod* genes. In *S. meliloti* it has been shown that the compounds trigonelline and stachydrine produced an increase in β-galactosidase activity from a *nodC-lacZ* fusion. Since *nodC* is under the control of the NodD2 LysR protein it was suggested that these compounds converted NodD2 to an active

form which initiates transcription of *nodC* (Phillips *et al.*, 1992). Both of these compounds are betaines, quaternary ammonium compounds produced by plants under conditions of osmotic stress. These compounds are abundant on legume roots and seeds as well as in root exudate (Chen and Murata, 2002). In addition several other compounds have been shown to induce *nod* genes including aldonic acids and phenolics such as vanillin (Gagnon *et al.*, 1998; Kape *et al.*, 1991; Bassam *et al.*, 1988).

The flavonoids (or other compounds) released from the plant are thought to interact directly with the NodD (LysR) proteins synthesized by the rhizobia. The binding or interaction of the flavonoid with the LysR causes a conformational change in the LysR such that its altered conformation makes it able to bind to DNA and alter the expression of genes responsible for synthesizing the components of the bacterial Nod factors (Schell, 1993). The synthesized Nod factors then act upon the plants causing changes leading to the establishment of symbiosis with the rhizobia. The actual biochemical analysis of NodD proteins has been limited however. This is due to many of these LysR proteins being insoluble and/or membrane associated and thus resistant to purification attempts. It has been demonstrated that in at least one case a NodD protein requires the GroESL chaperone in order to fold properly and the GroES protein has been shown to co-purify with the NodD in this case (Yeh *et al.*, 2002; Ogawa and Long, 1995). A NodD of *R. leguminosarum* has been purified (Feng *et al.*, 2002).

Structurally LysR proteins contain two functional domains. On the C-terminal end there is a variable domain believed to interact with activating compounds such as flavonoids. The variable sequence in this domain being explained by the need of different LysRs, for example different NodDs, to interact with different flavonoid compounds depending upon

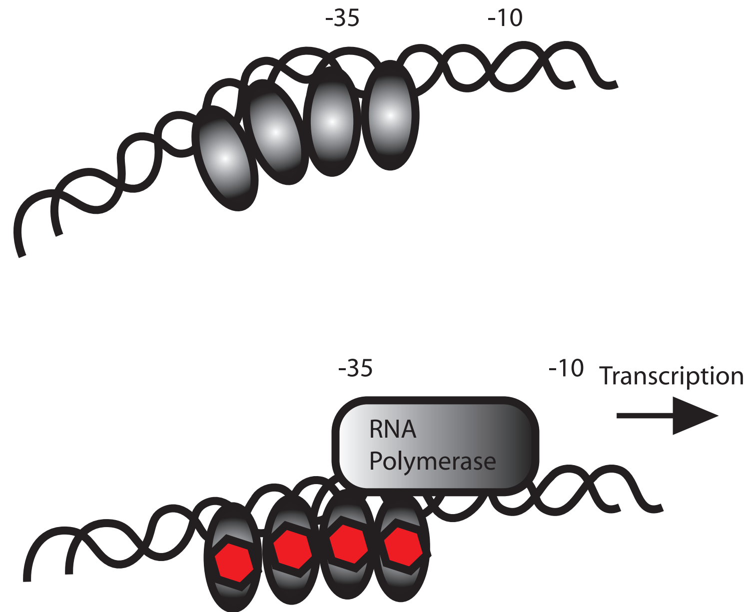

Figure 1.2: Model of LysR activation of gene transcription. In the absence of inducer molecules (top) the LysR protein binds to DNA as a tetramer causing a bend in the DNA. When inducer molecules (hexagons) such as flavonoids bind to the LysR protein a conformational change occurs which reduces the bending angle of the DNA and allows RNA polymerase to interact with DNA and gene transcription to proceed.

the particular species or legume and rhizobia forming the symbiotic pair. On the other end of the LysR is a highly conserved N-terminal DNA binding domain believed to interact with a specifically recognized sequence upstream of the *nod* genes whose synthesis is activated by the LysR (Schell, 1993). This specific DNA sequence that is targeted by the NodD is referred to as the *nod* box and helps to ensure that the particular NodD only activates its target *nod* gene. This model for NodD has been inferred from X-ray crystallography studies conducted with the limited number of other LysR proteins that have been purified such as CbnR or the purified C-terminal domains of DntR and OxyR (Smirnova *et al.*, 2004; Choi *et al.*, 2001; Tyrrell *et al.*, 1997). These studies demonstrate that these LysRs the C-terminal regions undergo a conformational change in the presence of a ligand. In each case, the two subdomains that make up the C-terminal region move closer together in the presence of a ligand but are further apart in its absence. The evidence for C-terminal involvement of NodD in binding to flavonoids is based on genetic studies in which the *nodD* genes of different rhizobia (*S. meliloti*, *R. trifolii* and *R. leguminosarum*) were cloned into a vector which was transferred into the non Sym plasmid containing strain LPR5045 along with another vector containing fusions to the upstream regions of *nodABCIJ*. It was demonstrated that different flavonoids caused different levels of expression in the presence of different *nodD* genes (Spaink *et al.*, 1989).

It has been demonstrated that *nodD* genes can be activated in the presence of flavonoids including luteolin (*S. meliloti nodD1*) as well as eriodictyol, 7-hydroxyflavone, luteolin and naringenin (*Rhizobium leguminosarum* bv *viciae*) and 7-hydroxyflavone, luteolin and naringenin (*Rhizobium leguminosarum* bv *trifolii*) (van Rhijn *et al.*, 1994; Dénarié *et al.*, 1992; Davis and Johnston, 1990).

The *nod* box to which the C-terminal domain of a NodD binds is located just upstream of the promoters of the *nod* operons targeted by the NodD. Typically the *nod* boxes are around 50 bp in length and it is believed that this length suggests that NodD binds to this region as a tetramer as has been demonstrated for other LysR-like proteins (Goethals *et al.*, 1992; Kondorosi *et al.*, 1989; Rostas *et al.*, 1986; Pabo and Sauer, 1984). Curiously it has been shown that NodD will bind to DNA containing a *nod* box sequence even in the absence of a flavonoid or flavonoid-like compounds (Fisher *et al.*, 1988; Hong *et al.*, 1987) and there is at least one example of this kind of binding resulting in a bending of DNA at the location of the *nod* box . While the LysR protein can bind to *nod* box DNA without an inducer molecule being present it seems that the inducer molecule is required for the LysR to be functional as has been suggested by DNA foot printing assays as well as *in vitro* NodD-*nod* box binding assays (Goethals *et al.*, 1992; Kondorosi *et al.*, 1989).

The sequence of the *nodD* gene among different rhizobia is not uniform. In addition, different rhizobia species can have more than a single copy of the gene or *nodD* genes of different sequences within the same species (Cullimore *et al.*, 2001; Downie, 1998). A homolog of NodD known as SyrM (symbiotic regulator) has also been identified. SyrM like NodD is an activator of *nod* genes. It is hypothesized that the slightly different NodD proteins coded for by the various *nodD* genes differ in their binding affinity for different *nod* boxes and may also differ in their flavonoid binding. This combination of different NodD and/or SyrM proteins coded for by different rhizobia can thus help to explain differences in host specificity among the rhizobia (Hanin *et al.*, 1998; Michiels *et al.*, 1995; Swart *et al.*, 1994; Michiels *et al.*, 1993; Mulligan and Long, 1989). It has been demonstrated that the transfer of *nodD* genes from one species of *Rhizobium* to another can alter the host specificity of the bacteria whereas point mutations in *nod* genes can alter the flavonoid

specificity of the proteins. In *S. meliloti* flavonoids are not required at all for SyrM and NodD3 to act as regulators (Swart *et al.*, 1994; McIver *et al.*, 1989; Horvath *et al.*, 1987; Spaink *et al.*, 1987; Burn *et al.*, 1987). In terms of regulating the expression of the *nodD* genes themselves, it has been demonstrated that NodD proteins act as their own repressors by binding to sequences between the divergently expressed *nodD* gene and the operon containing the regulated *nod* genes (Schlaman *et al.*, 1998; Spaink *et al.*, 1989). In addition, it has been demonstrated that NolR is also able to negatively regulate *nod* genes. NolR is a protein that contains a DNA binding motif similar to that of LysR regulators. NolR binds to the promoter region of *nod* genes as a dimer and blocks transcription of the *nod* regulon. Like many NodD proteins NolR negatively regulates itself (Cren *et al.*, 1995; Kondorosi *et al.*, 1991).

1.16.2 Other LysR proteins found in Rhizobia

Apart from the *nodD* and *syr* genes critical for the establishment of nodulation with the host plant very few LysR regulators have been well characterized in rhizobia. In *S. meliloti* there are 90 putative LysR genes that have been classified based on their DNA sequence similarity to LysR genes from non-rhizobia species. These genes are located on all three large genetic elements of the *S. meliloti* genome (pSymA, pSymB and the chromosome). Of these 90 genes, the vast majority have not been biochemically characterized in any detail, for example identification of the gene target or activating molecule. Luo (Luo *et al.*, 2005) characterized mutants of these predicted LysR genes but the characterization was limited to the impact of the mutation upon symbiosis. It was shown that for the vast majority of these genes the mutation does not impact symbiosis with the plant in an obvious way thus the true phenotypic impact and role of the vast majority of these putative regulators

remains unknown.

Of these 90 LysR genes found in *S. meliloti* a few have been phenotypically characterized. These include *bioS* which aids in the perception of signals from the plant which regulate biotin production (Heinz *et al.*, 1999). Of those *S. meliloti* LysR genes analyzed for their impact upon establishing nodulation with alfalfa plants, two were demonstrated to have a significant negative impact. These were *lsrA*(SMc00037) and *lsrB*(SMc01225) (Luo *et al.*, 2005). In the case of the SMc00037 mutant strain strain showed normal growth when free living and exhibited normal swarming behaviour but was unable to establish pink nitrogen fixing nodules when used to inoculate alfalfa. The SMc01225 mutant showed a phenotype quite distinct from the SMc00037 mutant. In this case the mutant strain displayed significant inhibition of free living growth and swarming behaviour. When inoculated on alfalfa, pink nitrogen-fixing nodules were established but the overall percentage of nodules that were pink and nitrogen fixing was only about half that of the wild-type strain. In both cases, mutants were recovered from white root nodules demonstrating that the mutants were able to reach the nodule thus demonstrating that the loss or reduction in nitrogen fixation observed was not the result (or at least not entirely the result) of the bacteria not reaching the nodule. In the case of the *lsrA* mutant this suggests that the LysR protein coded for by this gene is involved in regulating a step beyond establishing nodule invasion perhaps in regulating leghemoglobin biosynthesis though there is little evidence to support this hypothesis (Luo *et al.*, 2005). In the case of the *lsrB* mutant, while nitrogen fixation was established, the resulting plants were smaller and the leaves not as green as those plants inoculated with the wild-type Rm1021 strain. This suggested that these mutants may display a less efficient or lower level of nitrogenase activity. It was also shown in this case that the resulting pink nodules had a shape distinct from those occupied by Rm1021 which

may suggest that there are several points of abnormal gene expression in the establishment of symbiosis with this mutant strain. The *lsrA* gene shows closest homology to the *ampR* gene of *Enterobacter cloacae* (34% nucleotide indentity). This gene is located upstream of a glutathione S-transferase, *gst2*, and downstream of two overlapping *gst1* genes. The *lsrB* gene has highest similarity (29%) to the *crgA* gene of *Neisseria meningitidis* which is also a LysR-type regulator that is a negative regulator of genes responsible for pili and capsule production when the bacteria come in contact with animal cells. The *lsrB* gene is located downstream of the thioredoxin reductase gene *trxB* which also has a possible LysR recognition site in its upstream region. In *E. coli* the *trxB* gene is regulated by the LysR-type transcriptional regulator OxyR which is sensitive to redox potential within the bacterial cell. In *S. meliloti* redox potential within the nodule must be controlled to ensure efficient nitrogen fixation. This suggests a possible role for the *lsrB* in *S. meliloti* though this idea is not yet supported by direct experimental evidence (Luo *et al.*, 2005).

The same study that identified the *lsrA* and *lsrB* mutants also identified six additional putative LysR encoding genes that resulted in an altered phenotype. These included SMa1979, SMb21715, SMc00820, SMc04163, SMc03975 and SMc04315 (Luo *et al.*, 2005). In each case the resulting mutant strain was able to establish a symbiosis with alfalfa and carry out nitrogen fixation. However in each case a reduction in free-living motility was identified in complex growth media. Each of these genes were found to have sequences highly homologous to *lrhA* and *hexA*. In *E. coli* the *lrhA* gene is known to be involved in controlling the expression of genes related to chemotaxis and motility including genes involved in the production of flagella via the regulation of *flhDC* which acts as a "master regulator" of genes involved in motility and flagella production. The *E. coli lrhA* sequence also has a high degree of similarity to *hexA* (64%) of *Erwinia carotovora* and *pecT* of

Erwinia chrysanthemi. Both *hexA* and *pecT* are known to act as repressors of motility and as virulence factors related to cell lysis. Analysis of these six *S. meliloti* mutants revealed altered phenotypes related to motility and quorum sensing. These included large reductions in the production of *N*-acyl homoserine lactone (AHL) in SMa1979, SMb20715 and SMc04163 mutants compared with the wild-type Rm1021 strain suggesting that mutation in these LysR genes results in impaired quorum sensing (Luo *et al.*, 2005).

A recent study has identified the role of two LysR coding genes in *S. meliloti* that are involved in regulating the expression of the glycine cleavage operon *gcvTHP*. These LysR genes have been designated *gcvA1* and *gcvA2*. It was determined that both of these LysR genes were required for the full expression of *gcvT*. Mutation of these regulatory genes resulted in reduced expression of *gcvT* by 47% and 46% respectively. It was also demonstrated that the action of *gcvA1* is glycine inducible while that of *gcvA2* was not (MingSheng *et al.*, 2009).

One of the few well characterized LysR type transcriptional regulators apart from NodD in *S. meliloti* is coded for by the *pcaQ* gene. This LysR is involved in regulating the *pcaDCHGB* operon. The genes of this operon code for enzymes involved in the β-ketoadipate pathway which is involved in the first step of metabolizing aromatic acids from plants into catechol or protocatechuate. This metabolic pathway is found in many of the *Rhizobiaceae*. The PcaQ protein was able to be purified, and analyzed using gel mobility shift assays to determine that the protein binds to a region upstream of the transcriptional start site of *pcaD* between the -78 and -45 positions. It was also demonstrated in this case that the PcaQ DNA binding site showed a high degree of symmetry (5'-ATAACC-N_4-GGTTAA-3') and that PcaQ shows negative autoregulation. It still remains unclear, however, what if

any molecule acts in a manner analogous to flavonoids on PcaQ (MacLean *et al.*, 2008).

1.17 This Work

Previous to this work, the primary focus on understanding purine degradation in *S. meliloti* centered on the *bdhA-xdhA2-xdhB2* mixed function operon which demonstrated a link between the PHB cycle and the degradation of purines. In addition, a series of random Tn5 mutants were isolated that were found to lack the ability to utilize HX. The genetic loci of these mutants were determined and the HX utilization deficiencies were characterized.

A survey of the *S. meliloti* genome annotation database (http://iant.toulouse.inra.fr/bacteria/annotation.cgi) suggests that the region containing the *bdhA-xdhA2-xdhB2* operon is not the only genomic region associated with purine degradation. The mixed function operon and *guaD2* are the only genes in that region of pSymB that seem to be associated with this activity. In addition, this operon only contains the genes associated with two of the copies of the xanthine oxidase / xanthine dehydrogenase subunits. As mentioned previously, the genes coding for the other two XO/ XDH subunits (*xdhA1* and *xdhB1*) are located in a cluster of genes in a different location on pSymB which also seems to include many genes associated with purine degradation and related or similar functions. This region has not been previously investigated in any detail and the role of almost all the genes in this region in terms of purine degradation had not been experimentally determined. Thus purine catabolism in *S. meliloti* has, at best, only been partially characterized.

To begin to expand our understanding of purine degradation in *S. meliloti*, we formulated several hypotheses as a framework for our investigation. We wanted to test the assigned

annotations for the purine degradation genes. We assumed that genes assigned a given predicted role in the purine degradation pathway should display an obvious phenotypic defect when a strain containing a mutation in such a gene is grown on its predicted substrate. We also predicted that genes annotated as playing a role in this pathway should display elevated levels of expression when grown in the presence of their predicted substrates. In addition, we predicted that genes annotated as transcriptional regulators located in the same genomic region as the purine degradation genes would impact the expression of these genes. We would predict such impacts to be observable as growth defects or alterations in expression of purine degradation genes.

The initial phase of this project involved characterizing the symbiotic phenotype of the previously isolated and characterized Tn5 mutants (Capstick, 2004). These strains had not been previously used to inoculate alfalfa plants to determine if there were any possible symbiotic deficiencies that could be identified in these strains. This information could point to additional genes or regions of the *S. meliloti* genome that are involved in purine degradation that had not been previously analyzed and suggest additional steps involved in a purine degradation cycle. In the primary portion of this project we set out to analyze the role of several genes putatively believed to be involved in purine degradation in an effort to expand our understanding of purine catabolism in *S. meliloti*. Doing so involved the construction of mutants and reporter fusions to these genes and the subsequent investigation of molecules that act as their inducers, the impact of these mutations on gene expression and phenotype as well as the possible regulatory role of genetic elements found on pSymB near these putative purine catabolic genes.

Chapter 2

Materials and Methods

2.1 Bacterial Culture and Microbiological Techniques

2.1.1 Bacterial Strains, Plasmids, Transposons and Phage

Bacterial strains, plasmids, transposons and phage used in this project are outlined in Tables 2.1 and 2.2. For preservation and long term storage cultures were grown to a density of approximately 0.6-0.8 at OD_{600} and were then combined in a 1:1 ratio with liquid bacterial growth media (TY for *S. meliloti* and LB for *E. coli*) containing DMSO (16.5% in TY and 14% in LB [v/v]) and stored in glass or plastic vials at -70°C. From these frozen cultures working cultures were obtained by using a sterile wooden inoculating stick to scrape some of the frozen culture from the vial and streak onto the appropriate media containing agar.

2.1.2 Bacterial Growth Media

Media were sterilized by autoclaving at 121°C and 100 kPa for 30 min. Antibiotics and other solutions that could not be autoclaved, were filter-sterilized using either 0.2 μm or 0.45 μm syringe filters (PALL Corporation, Port Washington, NY, USA). TY media (Beringer, 1974) used to culture *S. meliloti* consisted of (g/L): tryptone, 5; yeast extract, 3; $CaCl_2$, 0.44. LB media used to culture *E. coli* consisted of (g/L): tryptone, 10; yeast extract, 5; NaCl, 10. *S. meliloti* (Miller, 1972) was also grown on defined M9 salts media (Miller, 1972) supplemented with a variety of nitrogen and carbon sources. M9 was prepared from a 20x stock solution composed of (g/L): Na_2HPO_4. 116;KH_2PO_4, 60;NaCl, 10;NH_4Cl, 20. The use of an alternate nitrogen source involved the exclusion of NH_4Cl from the M9 media and its replacement with the alternate nitrogen source at a concentration of 15mM. A 200 ml volume of M9 media consisted of 0.1 ml of 0.5 M $CaCl_2$, 0.2 ml of 1M $MgSO_4$, 0.2 ml biotin (0.3 mg/ml), 10 ml 20X M9 salts and a carbon source to a final concentration of 10 or 15 mM.

Autoinduction medium used for protein over-expression was made with 6 g/L Na_2HPO_4, 3 g/L, KH_2PO_4 , 20 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl. The pH of this solution was adjusted to 7.2 with NaOH and autoclaved. 0.2% lactose, 0.6% glycerol, and 0.05% glucose were added to the medium just prior to the inoculation using stock solutions of 8%, 60%, and 10% respectively.

Table 2.1: Bacterial strains used in this study.

Strain	Relevant characteristics	Reference or source
Sinorhizobium meliloti		
Rm1021	*str-21*, Sm^r, wild type	Meade *et al.* 1982
SmP110	Rm1021 $pstC^+$	Yuan *et al.* 2006
Rm5422	*ntrA75*::Tn5	Finan *et al.*:1988
Rm11421	Rm1021 *ccmC*::Tn5	Capstick: 2004
Rm11424	Rm1021 Smb21292::Tn5	Capstick: 2004
Rm11425	Rm1021 *ccmC*::Tn5	Capstick: 2004
Rm11428	Rm1021 *ccmC*::Tn5	Capstick: 2004
Rm11429	Rm1021 Smb20684::Tn5	Capstick: 2004
Rm11507	SmP110 SMb21291::ΩNm^r (pK19*mob*)	This Study
Rm11513	SmP110 SMb21285::pTH1703	This Study
Rm11517	SmP110 SMb21284::pTH1703	This Study
Rm11518	SmP110 *xdhA2*::pTH1703	This Study
Rm11519	SmP110 SMb21281::pTH1703	This Study
Rm11520	SmP110 *guaD2*::pTH1703	This Study
Rm11522	SmP110 *guaD1*::pTH1703	This Study
SmRL2522	Smp110 *xdhA1*::pTH1703	Finan *et al.* 2006

Continued on next page

Strain	Relevant characteristics	Reference or source
Rm20847	SmP110 SMb20847::ΩNm^r (pK19*mob*)	This Study
RmF117	ΔΩ5060-5033::Tn5-233	Charles and Finan: 1990
SmUW208	SmP110*xdhC*::ΩNm^r (pK19*mob*)	This Study
SmUW209	ϕSmRL2522→ Rm11507	This Study
SmUW210	ϕRm11517→ Rm11507	This Study
SmUW211	ϕRm11522 → Rm11507	This Study
SmUW212	ϕRm11513 → Rm11507	This Study
SmUW213	ϕRm11518→ Rm11507	This Study
SmUW214	ϕRm11519→ Rm11507	This Study
SmUW215	ϕSmRL2522→ Rm5422	This Study
SmUW216	ϕRm11517→ Rm5422	This Study
SmUW217	ϕRm11522→ Rm5422	This Study
SmUW218	ϕRm11513→ Rm5422	This Study
SmUW219	ϕRm11518→ Rm5422	This Study
SmUW220	ϕRm11519→ Rm5422	This Study
SmUW221	SmP110*xdhA1*::ΩNm^r(pK19*mob*)	This Study
Escherichia coli		
DH5α	F$^-$, ϕ80d*lac*ZΔM15, *end*A1, *rec*A1, Δ(*lac*ZYA-*arg*F)U169, *hsd*R17(r$_K^-$ m$_K^+$,)	Hanahan.: 1983)

Continued on next page

Strain	Relevant characteristics	Reference or source
	*deo*R, thi-1, *sup*E44, λ^-, *gyr*A96, *rel*A1	
MT616	MT607 pRK600, mobilizer	Finan :1986
BL21(DE3)pLysS	F^-, *omp*T, *hsd*S_B (r_B^-, m_B^-), *dcm*, *gal*, λ(DE3), pLysS, Cm^r	Studier:1986

Table 2.2: Plasmids used in this study.

Plasmid	Relevant characteristics	Reference or source
pRK600	Helper plasmid for conjugal transfer, Cm^r	Finan*et al.*.:1986
pGEM-T Easy®	Cloning of PCR products, Amp^r	Promega (USA)
pK19*mob*	Suicide vector. *mob*	Schafer *et al.*: 1994
pTH1703	*lacZ gusA* integrating reporter Gm^r	Cowie *et al.* 2006
pMG1	pLAFR1 clone from Rm1021 genomic library.	M. Georgiou
	Complements Rm11507, SmUW209, 210, 211, 212, 213, 214	
pDC007	EcoR1 fragment of pMG1 in pUC18	Capstick.: 2004
pDC084	SMb21284::pTH1703	This Study
pDC089	*guaD1*::pTH1703	This Study
pDC087	*guaD2*::pTH1703	This Study
pDC86	SMb21281::pTH1703	This Study
pDC077	SMb21284-85 intergenic region::pTH1703	This Study
pDC085	*xdhA2*::pTH1703	This Study
pDC060	SMb21291::pK19*mob*	This Study
pDC064	SMb20847::pK19*mob*	This Study
pDC006	169bp fragment of *xdhC* in pK19*mob*	This Study
pSP329	Broad-host-range vector	Zhan *et al.*.: 1990

Continued on next page

Plasmid	Relevant characteristics	Reference or source
pKW1	Fragment including SMb21284 in pSP329	This Study
pKW2	Fragment including *xdhA1* and *xdhB1* in pSP329	This Study
pKW3	Fragment including SMb21285 in pSP329	This Study
pKW4	Fragment including *xdhA1*,*xdhB1* and *xdhC*	This Study

Table 2.3: Primers used in this study.

Name	Sequence (5'-3')	Description
SMb21291F	gccgatctttcaagaatcc	SMb21291 mutant
SMb21291R	ccctcgtcgatatgctgttt	SMb21291 mutant
SMb20847F	gaagcagttcccgctgaat	SMb20847 mutant
SMb20847R	ggtcttgagggtcgatgatg	SMb20847mutant
SMb21284F	aggccttcctctcggaaat	SMb21284 pTH1703 fusion
SMb21284R	ggcgtagagaacgtcgaaac	SMb21284 pTH1703 fusion
SMb21293F	ccggctcctatcgtttcttc	SMb21293 pTH1703 fusion
SMb21293R	caggagttcgtccatgaagct	SMb21293 pTH1703 fusion
SMb20849F	caaggtatcgacgacagtgc	SMb20849 pTH1703 fusion
SMb20849R	cgggtctcgtcatagctca	SMb20849 pTH1703 fusion
SMb21281F	gtcttcatctccggcatcat	SMb21281 pTH1703
SMb21281R	catcgacagaaccgttacca	SMb21281 pTH1703
IR84TO85F	gatcgcggggatatctcata	SMb21284-85 intergenic region
IR84TO85R	tccgtcttgaacttgcattct	SMb21284-85 intergenic region
SMb20847X	ctcggattccaatcgaacat	To confirm SMb20847 fusion
SMb21284X	acttgcattctgtcgctcct	To confirm SMb21284 fusion
SMb21293X	gatgggaaggcataatcgtg	To confirm *guaD1* fusion
SMb20849X	cggcctatgaaggaattgtg	To confirm *guaD2* fusion
SMb21281X	cgaattcgctggaaaacttg	To confirm Smb21281 fusion
IR84TO85X	cgccctcctcgtaattgat	To confirm Smb21284-85 intergenic fusion
SMb21011X	tcaacgaggtgatgctcaag	To confirm *xdhA2*fusion

Continued on next page

Name	Sequence (5'-3')	Description
GFP Fwd	acagttttcgcgatccagac	To check orientation of insert in pTH1703
xdhA1F	agacgctgctggactttctg	*xdhA1* mutant
xdhA1R	cgaaccgttggcgatatt	*xdhA1* mutant
xdhCF	ctcgttcaaggggcgaac	*xdhC* mutant
xdhCR	gcgccgcagtcagtatct	*xdhC* mutant
KWxdhA1FS	cgtccatctctcgcatc	To confirm insert in *xdhA1* mutant
KWxdhA1FA	gaccgaaatccgcaat	To confrim insert in *xdhA1* mutant
KWM13F	gttttcccagtcacgacg	To confirm inserts in *xdhC*, *xdhA1* and SMb21291 mutants
KWM13R	gtgagcggataacaatttca	To confirm inserts in *xdhC*, *xdhA1* and SMb21291 mutants
KWxdhcFA	ggtggcggacagcttga	To confirm insert in *xdhC* mutant
KWxdhcFS	tcggcatgataggctcg	To confirm insert in *xdhC* mutant
KW21291FA	tctcgataatgtgcgtgtg	To confirm insert in SMb21291 mutant
KW21291FS	ggagtattactggaccttg	To confirm insert in SMb21291 mutant

Antibiotic stock solutions were stored at 4°C in ddH_2O or 95% ethanol. Antibiotic solutions were filter sterilized. Antibiotics in powder form were supplied by Sigma chemical, ICN Biomedical, BioShop, or Boehringer Mannheim.

Antibiotic concentrations used for *S. meliloti* were as follows (μg/ml): Gentamicin sulfate, 20 (40 for pTH1703 strains); neomycin sulfate, 100 to 200; streptomycin sulfate, 200; tetracycline hydrochloride, 10.

For *E. coli* antibiotic concentrations were as follows (μg/ml). ampicillin (sodium salt), 100; chloramphenicol, 10; kanamycin sulfate, 20; streptomycin sulfate, 200; tetracycline hydrochloride, 10.

X-Gal (5-bromo-4-chloro-3-indolyl-D-galactopyranoside) (BioShop) was dissolved in DMSO at a concentration of 40 mg/ml and used at a concentration of 80 μg/ml.

Bacterial cultures were grown in 5 ml broth in 16 x 150 mm test tubes (200 rpm) or on media containing 1.5% agar in 95 x 15 mm Petri plates (Fisher or VWR). Large culture volumes were grown in Erlenmeyer flasks at 200 rpm on an Innova 4430 shaker (New Brunswick Scientific). Cultures were grown at either 30°C (*S. meliloti*) or 37°C (*E. coli*).

2.2 Bacterial Genetic Techniques

2.2.1 Bacterial Conjugation

Cultures of donor, recipient and mobilizing strains were grown to mid- to late-log-phase in liquid media with appropriate antibiotics. Cultures grown with antibiotics were washed and resuspended in fresh antibiotic free media just prior to use. Equal volumes of donor, recipient and mobilizing strains mixed and a drop of this mixed culture (100 μL) was placed on LB agar and incubated overnight at 30°C. The mating spots were then resuspended in 0.85% NaCl and plated on selective media. To transfer a plasmid from a donor *E. coli* strain to a recipient *S. meliloti* strain the donor, mobilizer and recipient strains were mixed on an LB agar plate and incubated overnight at 30°C. The mating spot was then transferred using a sterile inoculating stick and streaked onto an agar plate with selective antibiotics.

2.2.2 Single-Crossover Mutagenesis

The pK19*mob* suicide vector ($Nm^r Km^r$) containing an internal fragment of the gene targeted for disruption was introduced into the wild type *S. meliloti* strain SmP110 by tri-parental mating. Transconjugants were plated on media supplemented with Sm-Nm which selected for single recombination events at the homologous position in the *S. meliloti* genome while counter-selecting against the *E. coli* donor strain. The resulting *S. meliloti* strain contains a pK19*mob* flanked by two truncated copies of the targeted gene.

2.2.3 Preparation of ϕM12 Transducing Lysate

To prepare transducing lysates, *S. meliloti* cultures were grown in 5 ml of TY broth to mid- to late-log phase. 50 μl of ϕM12 lysate from the wild type *S. meliloti* strain was added to the recipient culture and incubated at 30°C until complete cell lysis was observed. To kill any remaining viable cells, 30 μl of chloroform was added to the lysate and this mixture was allowed to stand at room temperature for 15 min. The lysate was then transferred to a polypropylene screw-cap tube and stored at 4°C.

2.2.4 Transductions using ϕM12 Phage

Transductions were carried out by mixing 0.5 ml of a 1:50 donor lysate diluted in TY with 0.5 ml of an *S. meliloti* culture grown to mid- to late-log phase. The mixture of cells and diluted phage was incubated at room temperature for 20 min to allow time for the phage to adsorb to the bacterial cell surface. Five ml of 0.85% NaCl was added and the cells were centrifuged at 4,000 rpm for 5 min (IEC 21000R centrifuge, IEC fixed place rotor). Cells were resuspended in 0.1 ml of 0.85% NaCl and spread on LB with selective antibiotics (Charles and Finan, 1990).

2.2.5 Creation of Fusion Strains using pTH1703

Fusion strains were constructed as described by Jacob (Jacob *et al.*, 2008). DNA fragments of genes known or suspected to be involved in the *S. meliloti* purine degradation cycle were PCR amplified from genomic DNA of the the SmP110 strain. PCR was carried out using Pfu DNA polymerase. The resulting PCR products were then cloned into the

intermediary vector pGEM-T Easy (Promega Corp.) in the T7 orientation. The inserts were cut from the T-easy vector via digestion with NotI and purified using a GFX PCR DNA and Gel Band Purification Kit (GE Healthcare). The purified fragments were then ligated into NotI-digested and dephosphorylated pTH1703 and used in the transformation of *E. coli* DH5α. Insert orientation for each clone was confirmed by PCR using gus and gfp primers along with the corresponding forward primer from the initial PCR amplification. The new constructs containing the PCR product in pTH1703 were subsequently mobilized from *E. coli* DH5α into *S. melilotii* SmP110 by triparental mating using an *E. coli* DH5α containing pRK600 as the helper. Transconjugants were isolated on TY containing gentamicin and streptomycin. Fusion strains containing a gene fragment internal to the gene resulted in a mutation in that strain while fusions to fragments that included regions from within a gene and external to the gene (for example the DNA fragment also included the region upstream of the gene) just resulted in a fusion to the genes promoter region.

2.2.6 Complementation of Mutant Strains

Complementation of mutant strains was carried out by ligating PCR amplified DNA fragments containing intact copies of mutated genes (including the promoter region) into the broad host range vector pSP329 and transferring the vector into the mutant strains via triparental mating. Complementation of SMb21291 mutant strains was carried out by transferring the cosmid clone pMG1 (containing the genomic region including SMb21291) into the mutant strains via triparental mating.

2.3 Molecular Biology Techniques

2.3.1 DNA Isolation and Purification

Isolation of Plasmid DNA

Plasmid DNA was isolated based on the alkaline lysis method described by (Birnboim and Doly, 1979). *E. coli* was grown overnight at 37°C and 1.5 ml of this culture was used to isolate plasmid DNA. The cells were centrifuged at 4,000 rpm for 5 min and resuspended in 100 μl of TEG (50 mM Tris-HCl, 20mM $Na_2EDTA \bullet 2H_2O$, 1% glucose [w/v], RNase A [200 μg/ml], pH=8.0). To the resuspended cells 200 μL of alkaline lysing solution (ALS) (0.2 M NaOH, 1% SDS) was added. The solution was gently mixed by inverting and 150 μl of ice-cold 7.5 M ammonium acetate was added and the solution was incubated on ice for 15 min followed by centrifugation at 15,000 rpm for 20 min. Following centrifugation the supernatant was transferred to a clean tube and 500 μl of 95% ethanol was added and the solution was incubated at -70°C for up to 30 min. Following this incubation, the solution was centrifuged at 15,000 rpm for 20 min, the supernatant discarded and the pellet washed with 70% ethanol and allowed to air dry. The pellet was resuspended in sterile nuclease-free ddH_2O.

2.3.2 Genomic DNA Isolation

Isolation of genomic DNA was carried out using an UltraClean DNA Isolation kit (MO BIO Laboratories Inc., USA) or by following the method outlined by Ausubel or Charles (Ausubel *et al.*, 1992; Charles and Nester, 1993). 250 ml *S. meliloti* cultures were grown

in TY followed by centrifugation to collect cells. The cells were then washed with 20 ml of 0.85% NaCl and 20 ml of TES (10mM tris-HCl, 25 mM $Na_2EDTA2H_2O$, 150 mM NaCl, pH=8.0) followed by resuspension in 20 ml of $T_{10}E_{25}$ (10 mM Tris-HCl, 25 mM $Na_2EDTA2H_2O$, pH=8.0). Cells were lysed by adding 1.25 ml of 20% SDS, 1.25 ml of pre-digested Pronase E (10 mg/ml), and 2.5 ml of 5 M NaCl. This mixture was incubated at 68°C for 30 minutes following by the addition of 7.5 M ammonium acetate (final concentration 2 M) and incubation for 1 h at -20°C. The thawed solution was then centrifuged at 4,500 rpm for 15 min (IEC 21000R centrifuge, IEC fixed-angle rotor). The supernatant was transferred to a clean centrifuge tube, extracted with 1:1 phenol-chloroform followed by extraction with chloroform. The top aqueous layer was then transferred to a clean centrifuge tube and precipitated with two equal volumes of 95% ethanol followed by centrifugation and washing with 70% ethanol. The pellet was resuspended in $T_{10}E_1$ (10 mM tris-HCl, 1 ml $Na_2EDTA2H_2O$, pH=8.0) or ddH_2O.

2.3.3 Preparation of Competent *E. coli* cells

Competent *E. coli* cells were prepared from a 100 ml culture grown in LB to an OD_{600} between 0.4 and 0.6. Cells were incubated for 10 min at 4°C and collected by centrifugation at 4,000 rpm for 10 min at 4°C (Sorvall GSA rotor). Cells were maintained at 4°C for the entire procedure. The pellet was resuspended in 50 ml of ice-cold 100 mM $CaCl_2$ and incubated on ice for 2 hours. Cells were collected by centrifugation at 4,000 rpm for 10 min (Sorvall GSA rotor) and suspended in 10 ml of ice-cold 100 mM $CaCl_2$ containing 15% glycerol and aliquoted into pre-chilled microcentrifuge tubes and stored at -70°C.

2.3.4 Transformation of *E. coli* Competent cells

Transformation of cells was performed by mixing plasmid DNA and 50 μl of 100 mM $CaCl_2$ in a 1.5 ml centrifuge tube followed by the addition of 100 to 200 μl of *E. coli* competent cells that were thawed on ice. The DNA-cell mix was incubated on ice for 30 min followed by heating at 42°C for 90 s and subsequent transfer to ice for 2 min. The cells were then incubated in LB broth for 1 h at 37°C. Following incubation the cells were resuspended in 100 μl of LB and plated on LB agar containing the appropriate antibiotic(s) (Sambrook *et al.*, 1989).

2.3.5 DNA Manipulation

Restriction Digestion of DNA

Restriction enzymes were obtained from Fermentas. Standard digestions were performed in a total volume of 20 μl with 0.1 to 1 U of enzyme. 10X reaction buffers used were the specific enzyme appropriate buffers provided by Fermentas. Unless otherwise specified reactions were carried out by incubation at 37°C for 1-3 h. Enzymatic inactivation was carried out at the manufacturer recommended temperature.

DNA Ligation reactions

Ligations were performed using 10 or 20 μl volumes using ligase and ligase buffer from Fermentas Inc.

Agarose Gel Electrophoresis

Gels were made using 1X TAE buffer and an agarose concentration of 0.8%. DNA fragments within agarose gels were visualized using either 3 μl of ethidium bromide (10 mg/ml) or 2 μl of Gel Red (Biotium Inc, Hayward CA, USA) added to 50 ml of molten agar just prior to casting the gel. DNA fragment sizes were estimated using Lambda HindIII or 1 kb ladder (Fermentas, Burlington, ON, Canada).

PCR Amplification of DNA

PCR reactions mixtures consisted of 200 ng of DNA template, 2 μl of 10X PCR buffer (Fermentas Inc, Burlington ON, Canada or Toyobo Inc, Osaka, Japan), 0.5 μl of dNTP, 1 μl each of forward and reverse primers (10 μM/μl), 1 μl of Taq, pfu (Fermentas inc) or KOD Hot Start DNA polymerase (Toyobo Inc) and ddH_2O to a final volume of 50 μl. Reactions were carried out using an Eppendorf Mastercycler personal (Eppendorf, Hamburg , Germany) using the following protocol; 94°C for 5 min, 30 denaturation cycles at 94°C for 1 min, annealing temperature suitable for each primer set, extension at 72°C or 68°C (for KOD) followed by a final extension at 72°C or 68°C for 10 min.

Cloning PCR Products

Cloning of PCR products was done using standard methods. For example with the the pGEM-T Easy vector system (Promega Corp, Madison,WI, USA). Ligations were carried out using 3 μl of PCR product (200 ng/μl), 1 μl of pGEM-T easy DNA (50ng/μl), 5 μl of 2X Rapid Ligation Buffer and 1 μl of T4 DNA Ligase (3 Weiss units/μl). The ligation mixture was incubated overnight at 4°C following the transformation of competent *E. coli*

with the ligation mixture. The cells transformed with the correct vector plus ligation product was isolated using *lac* alpha-complementation by blue/white screening on LB_{amp} agar plates supplemented with X-Gal.

DNA Sequencing

DNA sequencing was carried out at the Center for Applied Genomics at SickKids Hospital (Toronto, ON, Canada), Mobix (McMaster University, Hamilton, Ontario) or the Department of Biology, University of Waterloo.

2.4 Biochemical Methods

2.4.1 Analysis of Fusion Strains

β-Galactosidase Assay Fusion strains were analyzed following the method described by Miller (Miller, 1972). 5 ml cultures were grown in M9-glucose. Cells from these cultures were then centrifuged and washed three times to remove any residual carbon source. 1 ml of this culture was then used to inoculate a 250 ml flask or 50 μl to inoculate a 5 ml tube containing M9 plus a carbon and nitrogen source (for example M9 with glucose and hypoxanthine). Flasks were then incubated with shaking at 30°C. At mid-log phase 2 ml of cells were collected , pelleted and resuspended in the same volume (2 ml) and the OD_{600} was recorded. 0.5 ml of cells were combined with 0.5 ml Z-buffer followed by the addition of 100 μl of chloroform and 80 μl of 0.1% SDS. The mixture was vortexed for 10 seconds followed by incubation for 10 minutes. The colourmetric reaction was started by the addition of 200 μl of ONPG (4 mg/ml) and terminated upon the development of

yellow colour by the addition of 500 μl of Na_2CO_3. The time from the start to termination of the reaction was recorded and used to calculate activity as follows;

Activity is expressed in Miller units as 1000 x (Absorbance at OD_{420} - (1.75 x OD_{550}))/ (Time of reaction x volume of cells x OD_{600})

The composition of Z buffer was as follows : (per 1 L) 16.1 g $Na_2HPO_47H_2O$, 5.5 g NaH_2PO7H_2O, 0.75 g KCl, 0.246 g $MgSO_47H_2O$, 2.7 ml β-mercaptoethanol. The β-glucuronidase assay was carried out in a similar manner to the β-galactosidase assay except the initial step of dilution with buffer is omitted as the addition of GUS buffer was used to start the reaction (unlike Z-buffer, GUS buffer contains the substrate for the reaction). Upon the development of yellow colour the reaction is terminated by the addition of 1 M Na_2CO_3 and ODs at 420 and 550 nm were determined as before. Activity, in Miller units, was calculated as above. GUS buffer (per 10 ml) consisted of Phosphate-EDTA buffer (7.5 ml), 0.1% SDS (1.25 ml), 500 nM DTT (1.0 ml) and 4-nitrophenyl β-D-glucuronide (250 μl of 17.6 mg/ml solution).

2.4.2 Cell-Free Extracts

250 ml cultures were grown in TY broth on a shaking incubator at 30°C and 200 rpm. Samples were checked for contamination by streaking a sample of this culture on TY agar plates. Cells from the 250 ml culture were harvested by centrifugation (4,000 rpm for 5

min) and washed four times with buffer (20 mM Tris-HCl pH=7.8, 1 mM $MgCl_2$). Cells were resuspended in 4 ml of buffer (20 mM Tris-HCl pH=7.8, 1 mM $MgCl_2$, 10 % glycerol, 10mM β-mercaptoethanol) per g wet weight of pellet and kept on ice. The resuspended cells and 0.1 mm glass beads (Biospec) were combined in equal ratios in 2 ml screw-cap vials (Biospec). The cells were disrupted by bead beating (Biospec Mini Beadbeater-8) at 4°C for 3 min. The tubes were transferred to ice for 1 min following each minute of bead beating. Following bead beating tubes were centrifuged at 12,000 rpm for 20 minutes (IEC 21000R centrifuge, IEC 24-place fixed-angle rotor) to remove cell debris. Cell extract was transferred to clean tubes used immediately in the native gel assay or stored at -70°C.

2.4.3 Enzyme Assay and Activity

Xanthine Oxidase

The detection of xanthine oxidase (EC 1.1.3.22) activity was carried out using non-denaturing polyacrylamide gel electrophoresis (Selander *et al.*, 1986). Cell-extracts were electrophoresed in a 5 % native acrylamide gel (4 % stacking gel) in Tris-glycine buffer (124 mM Tris, 1M glycine, pH=8.3) at 90 V for 3 hours. Xanthine oxidase activity was detected using a method modified from Selander (Selander *et al.*, 1986). A staining solution (0.2 M Tris-HCl buffer pH=8.0, 25 mL, HX, 25 mg, MTT (4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide 12.5 mg/ml, 1 ml phenazine methosulphate (PMS), 10 mg/ml, 0.5 ml) was incubated with the gel for 30 min at 30°C in the dark. Following incubation the gel was washed with ddH_2O and fixed with a solution of methanol, acetic acid and ddH_2O prepared with a 5:1:5 ratio. An image of the gel was then taken with a digital camera (Canon Powershot A520, Canon Inc, Toyko, Japan).

2.5 Plant Assays

2.5.1 Inoculation of Plants with Bacterial Strains

Medicago sativa cv. *Iroquois* seeds were sterilized by exposure to 95% ethanol for two minutes followed by 2.5% sodium hypochlorite solution for 15 minutes. Sterilizations and washings were done in a sterile centrifuge tube by pouring off the liquid between washings. The seeds were then washed 10x with sterile double distilled H_2O. Seeds were then transferred to sterile water (0.1%) agar plates, placed in the dark at room temperature and allowed to germinate for 48h (Wang *et al.*, 2007).

Surface sterilized seeds were then aseptically transferred to Magenta jars (Sigma) containing 200 ml of sterile Jensen's nutrient solution and sterile vermiculite (Jacob *et al.*, 2008). Jensen's medium consisted of (g/L): $CaHPO_4$, 1;HPO_4, 0.2; $MgSO_47H_2O$, 0.2; NaCl, 0.2; $FeCl_3$, 0.1 and 1 ml of trace elements solution (pH=7.2). Trace elements solution (1000X) consisted of (g/L): H_3BO_3, 0.1; $ZnSO_47H_2O$, 0.1; $CuSO_45H_2O$, 0.05; $MnCl_24H_2O$, 0.05; $Na_2MoO_42H_2O$, 0.1; $Na_2EDTA2H_2O$, 1 and NaFeEDTA, 0.2 (Jensen, 1942). The magenta jars were then transferred to a growth chamber for 48 h. Concurrent with the preparation of the seeds and growth containers bacterial cultures of each strain to be assayed were inoculated in 5 ml of TY media and grown for 48-72 h at 30°C. 100 μl of each bacterial culture was diluted in 5 ml (1:50 dilution) of sterile saline (0.85% NaCl) and an entire 5 ml dilution was used to inoculate each of 3 magenta jars containing 5 seedlings per jar. The Magenta jars were then re-covered and returned to the growth chamber (Percival Scientific, DiaMed Lab Supplies Inc.) under the following conditions (22°C light/18°C dark) on an 18 h day / 6 h night cycle and allowed to grow for up to 6 weeks.

Obtaining Plant Shoot and Root Dry Weight

After the plants were grown for 6 weeks the Magenta Jars were opened and the entire plant removed. The plant shoots, roots and nodules were removed and the shoots and roots were placed in a drying oven at 80°C and periodically weighed until no further weight reductions were observed.

Bacterial competition assays

Alfalfa plants were co-inoculated with wild type *S. meliloti* Rm1021 and a mutant strain and grown in Magenta jars as described above. Cultures were mixed in proportions ranging from 1:1 and 1:9 for mutant and wild type mixtures or 1:1 for wild type and mutant strains mixes. After four weeks of growth the root nodules were aseptically removed from plants using sterile forceps and the nodules were surface sterilized with 1% sodium hypochlorite for 15 min and crushed in a few drops of TY containing 0.3 M sucrose onto TY agar plates. The crushed nodule suspension was then streaked onto TY agar. Portions of bacterial colonies were subsequently subcultured onto antibiotic selective TY agar to differentiate wild type from mutant strains for each nodule and the counts for each type (wild vs. mutant) were obtained. A total of 80-100 colonies were screened from each root nodule by randomly selecting colonies from the agar plate and screening isolate for the correct antibiotic-resistance marker or growth phenotype (e.g inability to grow on hypoxanthine).

Nodulation Kinetic Assays

Plants were inoculated individually with wild type or mutant *S. meliloti* strains as described above. At two day intervals plants, were removed from magenta jars and nodules

were screened for the presence of the bacterial strain by plating the crushed nodules on TY agar with selective antibiotics and the total counts for the number of colonies were obtained for every second day for up to 35 days.

Acetylene Reduction Assays on *Sinorhizobium* Mutants

Alfalfa plants were prepared and inoculated as described above. Nitrogen fixation activity was determined by the acetylene reduction method. 10 ml of acetylene gas was injected into 100 ml bottles containing the alfalfa plants that had been inoculated with various *S. melilot*i strains. The bottles were incubated for 3 h before being analyzed. 200 μl of gas was removed via a syringe from each sample bottle for analysis of acetylene and ethylene levels using a gas chomatograph (Shimadzu GC-17A. Shimadzu Corp. Kyoto Japan) fitted with an HP-AL/M column (30 m, Agilent Technologies, Santa Clara, CA, USA) and a flame-ionization detector. The carrier gas flow rate was set to 35 ml/min, air at 350 ml/min, and helium at 30 ml/min. The GC oven program was isocratic for 6 min. Following the acetylene reduction assay the nodules were removed from the roots, counted and dried for weighing as previously described.

2.5.2 Statistical Analysis

Student's t-test was used to analyze differences between means in different treatments. For example to compare the β-galatosidase activity in the SmP110 strain grown in M9 with glucose and ammonium chloride compared with the same strain grown in M9-hypoxanthine at a 95% confidence level.

Chapter 3

Phenotypic Analysis of Mutants Unable to utilize Hypoxanthine

3.1 Introduction

A regulatory link in *Sinorhizobium meliloti* has been demonstrated between intracellular carbon and nitrogen stores via the biotin-inducible *bdhA-xdhA2-xdhB2* mixed function operon (Hofmann *et al.*, 2000a; Aneja and Charles, 1999) . This operon synthesizes the enzymes D-3-hydroxybutyrate dehydrogenase and xanthine oxidase. D-3-hydroxybutyrate dehydrogenase catalyzes the second step in the degradation of the carbon storage compound poly-2-hydroxybutyrate (PHB) while xanthine oxidase is an enzyme involved in the conversion of the purine subunits hypoxanthine to xanthine and xanthine to uric acid in the purine salvage pathway. Via the PHB degradation pathway pathway, PHB granules within the bacterial cell can be degraded to provide a source of carbon and energy when the

bacteria encounter a carbon limited environment. In addition, the degradation of purines via the purine degradation pathway, of which xanthine dehydrogenase is a key enzyme, provides a source of nitrogen to the bacteria when environmental nitrogen sources are limited or when there are available purines in the soil environment due to degradation of DNA and RNA from other biological sources (Charles, 2002; Anderson and Dawes, 1990; Tombolini and Nuti, 1989).

The initial earlier phase of this project involved the generation and screening of Tn5 mutants that were defective in hypoxanthine utilization either as a sole source of carbon and / or nitrogen (Capstick, 2004). The initial goal of this early phase of the project was the isolation of hypoxanthine utilization mutants that derived their phenotype as the result of mutations in the one of the xanthine dehydrogenase operons (either the *xdhA1-xdhB1-xdhC* operon or the mixed function *bdhA-xdhA2-xdhB2* operon). Surprisingly, no such mutants were isolated. Instead, sequence analysis showed that the isolated hypoxanthine utilization mutants were found in genes other than those expected. These mutants could be divided into three distinct classes. Class one being an insertion into SMb21292, a conserved hypothetical membrane protein. Class two consisted of the *ccmC* gene, a heme exporter C transmembrane protein and class three consisted of an insertion into SMb20684, a conserved hypothetical protein.

The objective of the work presented here was to characterize the symbiotic phenotype of these mutants using a variety of approaches. These included growing alfalfa inoculated with wild-type *S. meliloti* as well as the hypoxanthine utilization mutants and comparing the relative ability of the mutants to establish a symbiotic association with the host plant. The effectiveness of symbiosis was determined through a variety of measures. These

included observing the presence or absence of pink nodules, nodule numbers, the rate of nodule appearance, plant growth characteristics including size and weight and the competitiveness of the mutant strains relative to wild-type *S. meliloti* in terms of establishing nodule occupancy. This variety of measurements and observations were used to provide a picture of impact of mutations in these genes upon the alfalfa-*S. meliloti* symbiosis.

3.2 Results

3.2.1 Nodulation Phenotype

A collection of *S. meliloti* mutants previously identified as being unable to utilize hypoxanthine or xanthine as a sole nitrogen and/or carbon source or were growth inhibited by hypoxanthine (Rm11429) were identified during an earlier phase of this project (Capstick, 2004). The next step in analyzing these mutants involved determining what impact these mutations had on the ability of *S. meliloti* to establish a symbiosis with its host alfalfa plant. Observation of plant growth after 30 days (Fig. 3.1) showed that strains with mutations in the gene *ccmC* coding for a putative heme exporter (Rm11425), in SMb21292, a conserved hypothetical membrane protein (strain Rm11424), and in SMb20684, a conserved hypothetical protein (strain Rm11429), all produced plants with healthy green leaves/shoots and examination of the roots showed the presence of pink nodules. Strains Rm11424 and Rm11429 produced a similar number of pink nodules and also a similar number as the plants inoculated with the wild-type Rm1021 control strain. Plants inoculated with Rm11425 while producing pink nodules showed a notable reduction in the average number of pink nodules compared with plants inoculated with Rm1021, Rm11424 or Rm11429.

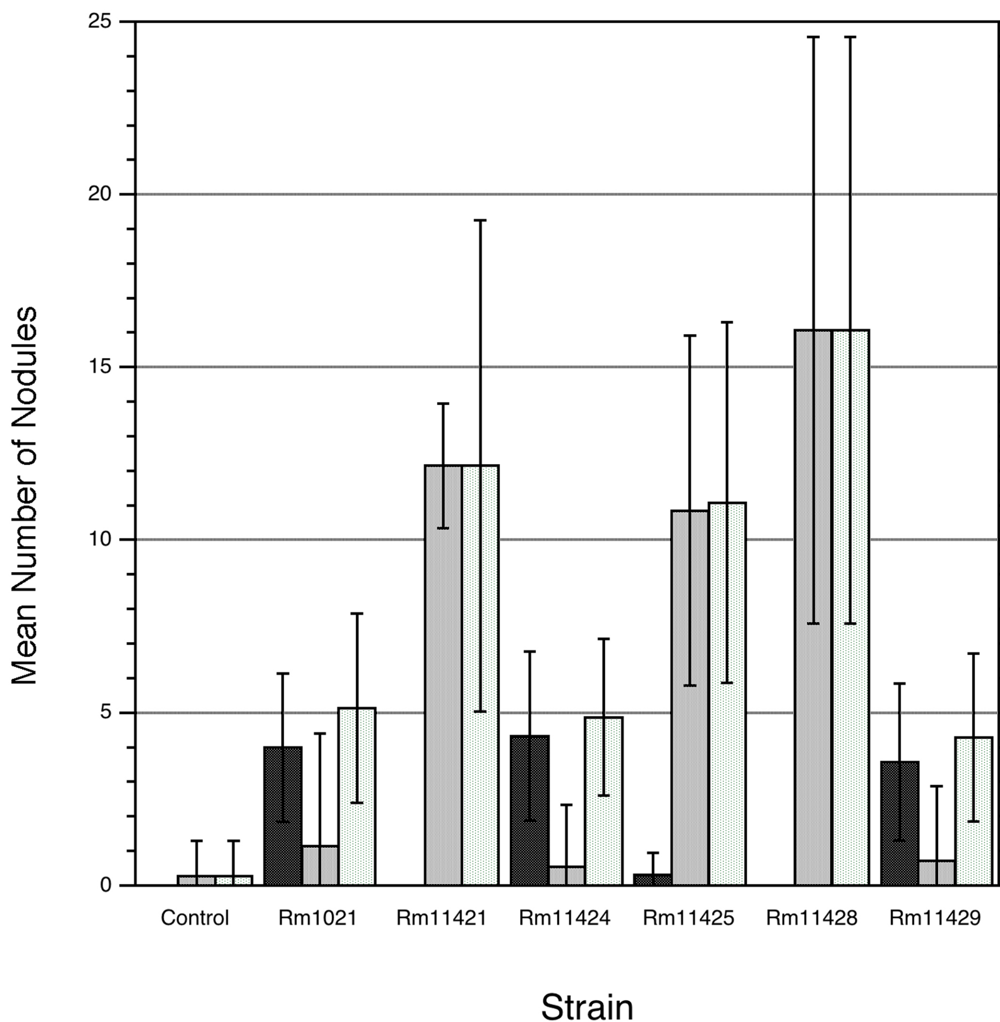

Figure 3.1: Phenotype of nodules from *Medicago sativa* cv *Iroquois* plants inoculated with *S. meliloti* Rm1021(wild type), Rm11421(*ccmC* mutant), Rm11424 (SMb21292 mutant), Rm11425 (*ccmC* mutant), Rm11428 (*ccmC* mutant) and Rm11429 (SMb20284 mutant). Control = uninoculated plants. N = 15 plants per bacterial strain.

Both strains Rm11421 and Rm11428 (both *ccmC* mutants) produced a plant phenotype that was quite distinct from the other strains. At 30 days, plants inoculated with these strains had pale green to yellow leaves and were notably shorter than plants inoculated with other strains. Examination of the roots showed a complete absence of pink nodules but a large number of white nodules (on average more than in the healthy plants).

3.2.2 Shoot Dry Weights

Weighing of dried shoots from plants (Fig. 3.2) inoculated with *S. meliloti* showed that on average plants inoculated with the wild-type strain Rm1021 and the nitrogen fixing strains Rm11424 and Rm11429 had the highest shoot dry weight among the plants inoculated with the different strains while plants inoculated with Rm11421, Rm11425 and Rm11428 had, on average, smaller shoot dry weights which were similar to those of the un-inoculated control plants. In all cases, however, there was quite a lot of variability in shoot weights among the plants in each treatment.

3.2.3 Nodulation Kinetics

The nodulation kinetics study (Fig. 3.3) showed that alfalfa plants inoculated with strains Rm11421 and Rm11429 displayed a notable lag in forming nodules on all plants compared with Rm1021. In the case of Rm11429, this lag was about 6-7 days while in the case of Rm11421 inoculated plants 100% nodulation was not observed for the duration of the study. Rm11424 and Rm11428 inoculated plants, reached 100% nodulation more quickly

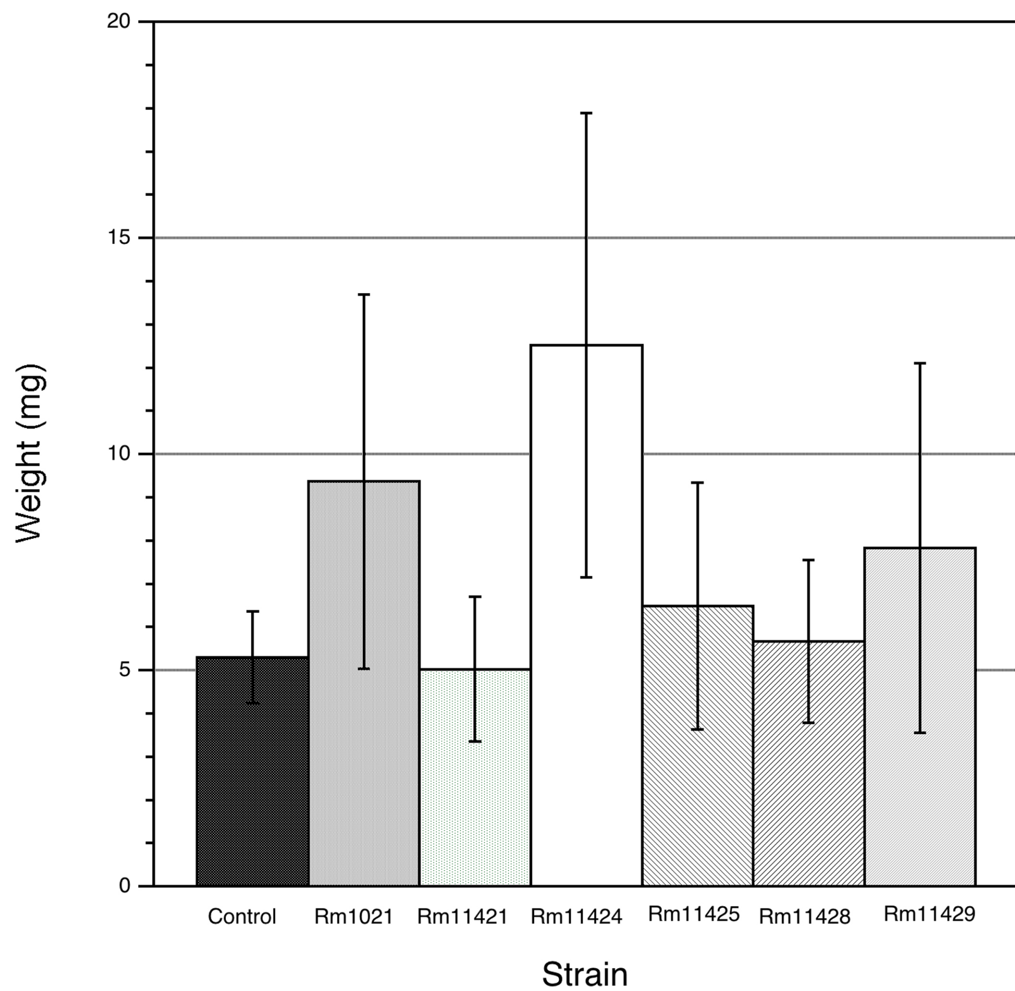

Figure 3.2: Shoot dry weight of *Medicago sativa* cv. *Iroquois* plants inoculated with *S. meliloti* Rm1021(wild type), Rm11421(*ccmC* mutant), Rm11424 (SMb21292 mutant), Rm11425 (*ccmC* mutant), Rm11428 (*ccmC* mutant) and Rm11429 (SMb20284 mutant). Control = uninoculated plants. N = 15 plants per bacterial strain.

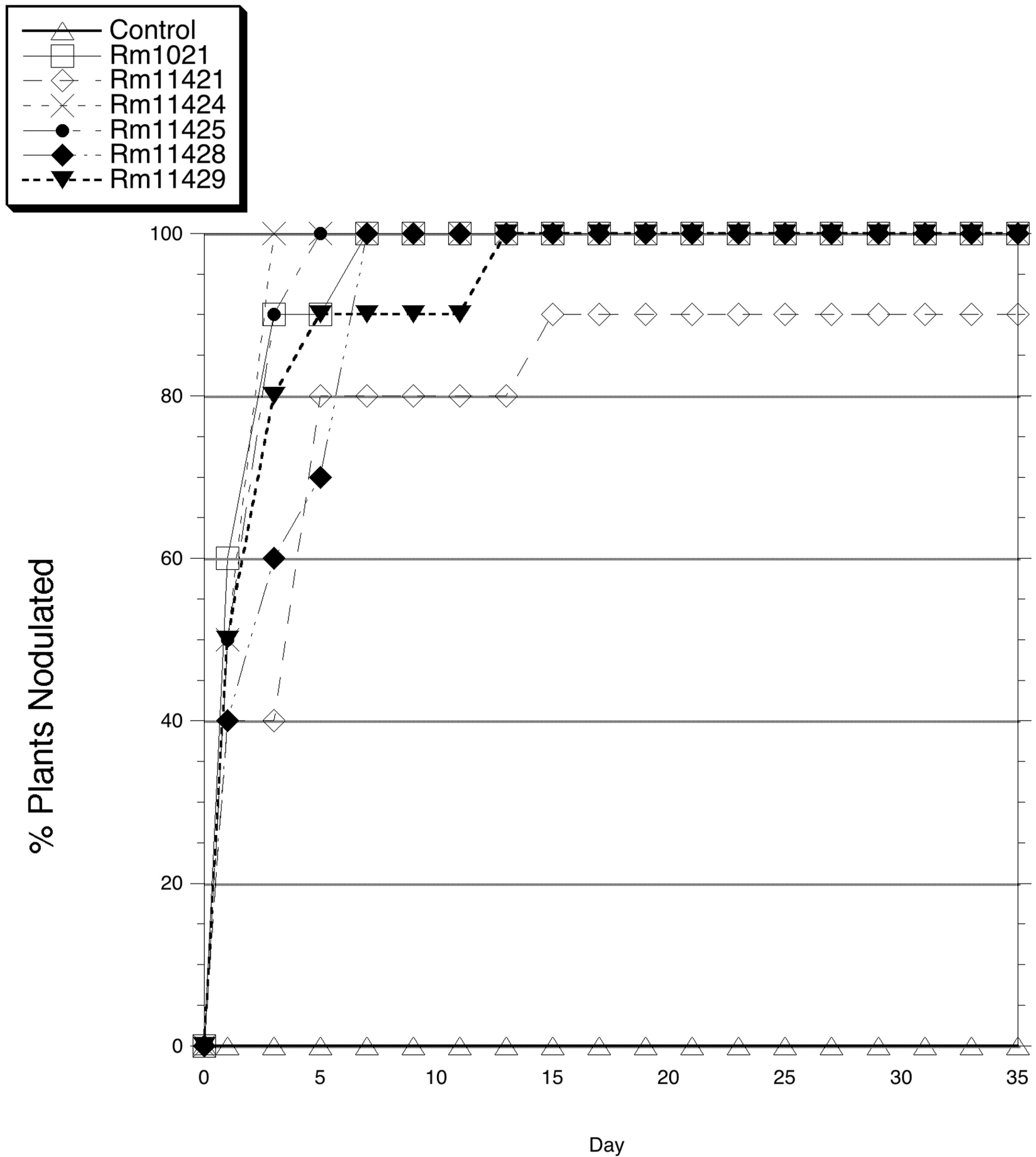

Figure 3.3: Nodulation kinetics for the percentage of *Medicago sativa* cv *Iroquois plants* nodulated by *S. meliloti* Rm1021 (control Fix +), Rm11421 (*ccmC* mutant), Rm11424 (SMb21292 mutant), Rm11425 (*ccmC* mutant), Rm11428 (*ccmC* mutant) and Rm11429 (SMb20284 mutant). Control = uninoculated plants. N= 15 plants per bacterial strain.

than Rm1021 inoculated plants (about 2 days vs. 4 days for Rm1021).

3.2.4 Nodulation Competitiveness

The nodulation competitiveness study (Fig. 3.1) showed that Rm1021 could be recovered from all inoculated plants. In the case of plants co-inoculated with both Rm1021 and a mutant strain, Rm1021 was recovered from a majority of nodules while mutant strains were recovered from as few as 8.5% of nodules in the case of the Rm1021/Rm11425 co-inoculated plants to a high of 21.2% in the case of Rm1021/Rm11429 co-inoculated plants. In all cases, a least a few nodules were found to contain mutant strains.

3.2.5 Acetylene Reduction Assays

Acetylene reduction assay results (Fig. 3.4) mirrored those seen from the physical examination of the plants. Plants with pink nodules on the roots and green leaves and stems showed levels of nitrogenase activity similar to those of plants inoculated with the wild-type Rm1021 strain. These plants were those inoculated with the strains Rm11424, Rm11425 and Rm11429. Plants inoculated with Rm11421 or Rm11428 produced plants with yellow leaves, no pink nodules and levels of enzymatic activity like the uninoculated control plants.

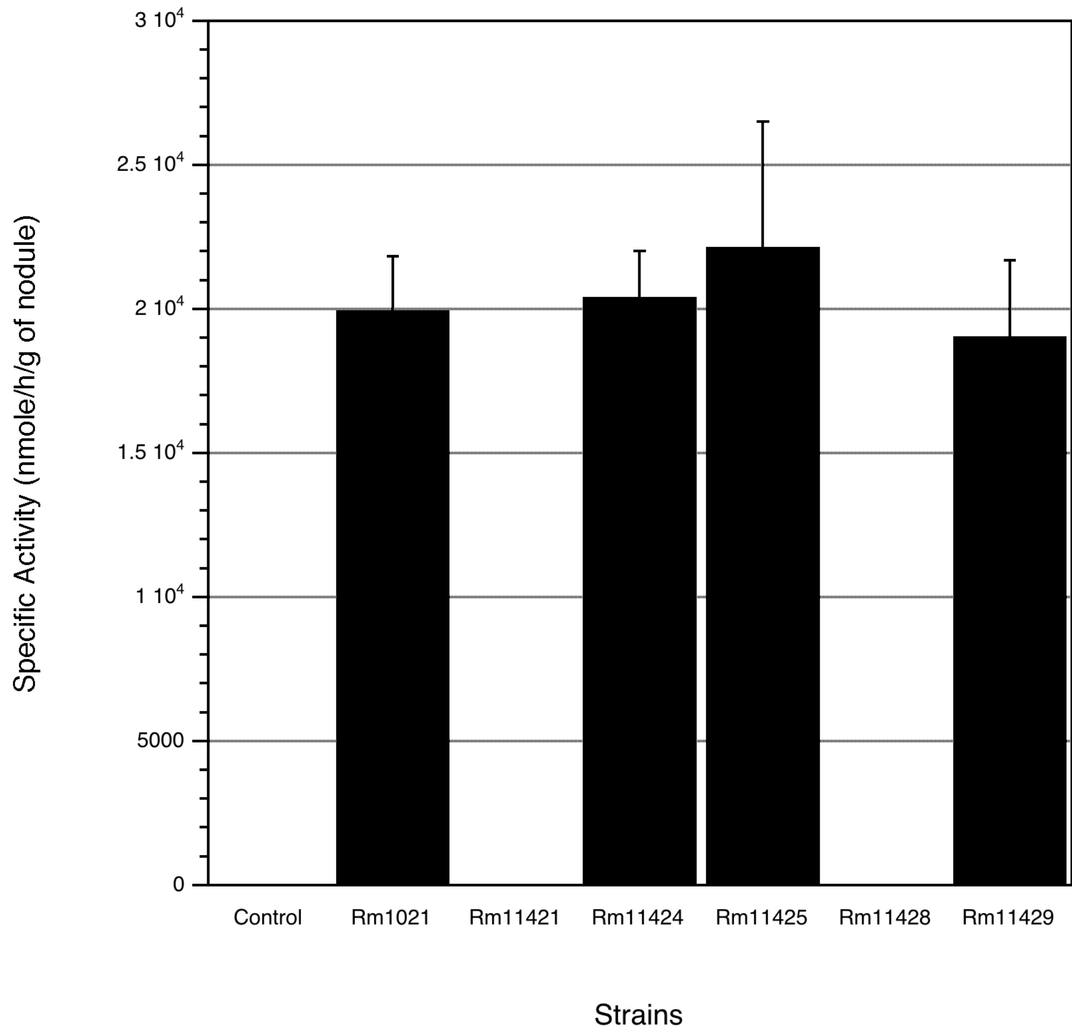

Figure 3.4: Acetylene reduction assay of *S. meliloti* hypoxanthine utilization mutant strains inoculated onto *Medicago sativa* cv *Iroquois* plants. *S. meliloti* Rm1021 (control Fix +), Rm11421 (*ccmC* mutant), Rm11424 (SMb21292 mutant), Rm11425 (*ccmC* mutant), Rm11428 (*ccmC* mutant) and Rm11429 (SMb20284 mutant). Control = uninoculated plants. N= 15 plants per bacterial strain.

Table 3.1: Nodulation competitiveness of *S. meliloti* mutants co-inoculated with Rm1201 on *Medicago sativa* cv *Iroquois* plants.

Strain	% Strain in inoculum	No. of Screened Nodules	% Wild Type	% Mutant	p-value.
Rm1021	100	76	100	N/A	N/A
Rm11421	52	46	82.6	17.4	<0.01
Rm11424	51	37	81.1	18.9	<0.01
Rm11425	54	47	91.5	8.5	<0.01
Rm11428	49	33	78.8	21.2	<0.01
Rm11429	51	41	82.9	17.1	<0.01

3.3 Discussion

Observation of plant appearance demonstrated that the only group of hypoxanthine mutants in this study that were unable to establish symbiosis and fix atmospheric nitrogen are those with mutations in the *ccmC* gene coding for a putative heme exporter protein associated with the electron transport chain (Rm11421 and Rm11428) (Fig. 3.1). The lack of green pigmentation in the leaves, the reduced size of the plants and the lack of pink nodules on the roots (suggesting that leghemoglobin associated with nitrogen fixation was lacking) are all characteristic of plants suffering from nitrogen deficiency. This observation was consistent with that reported by Pobigaylo and Yurgel (Pobigaylo *et al.*, 2008; Yurgel *et al.*, 2007). The most probable explanation for this phenotype would flow from the fact that fixing atmospheric nitrogen is a very energy demanding process. As there is a direct relationship between cytochrome C and respiration in bacterial cells, it is logical to conclude that a mutation in cytochrome C synthesis, in this case a failure to properly insert a heme group, would result in an impairment of function in the electron transport chain resulting in a loss or reduction in the cell's ability to use oxygen as a final electron acceptor. This would result in a reduction in ATP formation to the point that there would be insufficient ATP produced to supply the energy demands associated with nitrogen fixation (Delgado *et al.*, 1998). This was the case for strains Rm11421 and Rm11428. In the case of the third *ccmC* mutant (Rm11425) however there was a reduced number of pink nodules but nonetheless it was clear that nitrogen fixation was occurring with this strain. This suggests that while there may have been a mutation in the *ccmC* gene in Rm11425, the native of the mutation did not sufficiently distort the structure of the expressed protein to completely eliminate its function. Unlike the other mutations, the precise location of this mutation was not determined by sequencing but was shown to be in this region by transductional linkage. It is possible that this mutation is located in one of the nearby *ccm*

genes where it does not result in a non-fixing phenotype.

In strains Rm11424 (hypothetical membrane protein) and Rm11429 (conserved hypothetical protein), the lack of ability to utilize hypoxanthine as a sole carbon and/or nitrogen source did not result in a loss of ability to perform nitrogen fixation with root nodules or transfer fixed nitrogen to the host plant. In both cases the appearance of the host plants and the average number of nitrogen fixing nodules per plant was similar to plants inoculated with the wild-type Rm1021 strain of *S. meliloti* (Fig. 3.1). In the case of Rm11424 the mutation is located within a region of the pSymB megaplasmid populated with genes believed to play a role in purine metabolism so it seems likely that this gene would also be associated with these functions. Located next to this gene are SMb21291, a putative LysR transcriptional regulator, that plays a role in the purine degradation pathway (this will be discussed in greater detail later in this work) and *guaD1*, a gene that is part of the operon coding for a guanine deaminase enzyme that removes the NH_2 group from the purine guanine. The loss of ability to utilize hypoxanthine as a sole nitrogen and/or carbon source in this mutant strongly suggests that SMb21292 has a function in this pathway. The gene mutated in Rm11429 codes for a protein of unknown function. This gene is located on pSymB in a region populated by genes involved in metabolism of small molecules and macromolecules. The genes surrounding the mutated gene include those coding for DNA ligase, a LysR transcriptional regulator, a glyoxylate carboligase and a proprionate reductase. BLASTp analysis (expect threshold = 10, word size = 3, matrix = BLOSUM62, gap costs = existence:11 extension 1) (http://blast.ncbi.nlm.nih.gov/Blast.cgi) of the sequence shows that the closest match to the sequence in *Sinorhizobium* is to a protein of unknown function in *Sinorhizobium medicae*. The closest match to a gene with an identifiable function is to a gene believed to be involved in glyoxalate utilization in *Rhizobium*

sp. NGR234 which has an 84% identity with SMb20684. Interestingly, the next closest matching gene is identified as coding for putative allantoin catabolism protein in *Citreicella* sp. (a Rhodobacteraceae genus where the protein has a 97 % identity with SMb20684) which may suggest a possible role in purine catabolism.

Analysis of shoot dry weight (Fig. 3.2) suggested that those strains with a fix negative phenotype (Rm11421, Rm11428) as well as the Rm11425 strain (fix positive but with a reduced number of pink nodules) resulted in, on average, small plants as reflected in the smaller average shoot weight. For the fix negative strains this would not be surprising as the nitrogen deficiency in these plants would result in an impaired ability to synthesize proteins once the initial nitrogen supply in the seeds was exhausted. The result for Rm11425 is interesting as it suggests that while the plant is obtaining fixed nitrogen from the colonized root nodules there may be an impairment in the amount of nitrogen available to the host plant leading to reduced growth in an otherwise healthy looking plant. This more subtle phenotype follows logically from the observation of a reduced pink nodule count. Not surprisingly the other fix positive strains show shoot weights closer to that of plants inoculated with the Rm1021 wild-type strain which is consistent with the nodule count data described above. One significant caveat when considering this data is the high degree of variability in plant growth which is reflected in the error bars (Fig. 3.2). There was a wide range of plant sizes among the 15 plants in each inoculation trial with considerable overlap among plant sizes from the different inoculations. This high degree of variability in measurements of plant morphology was observed repeatedly and makes drawing definitive conclusions from this type of data quite difficult. One way to improve this analysis would be to increase the number of plants used in the experiment.

Examination of the nodulation competitiveness data (Fig.3.1) suggested that the impact of the mutations and changes in the hypoxanthine utilization phenotype had an impact on the various bacterial strains that was not always obvious just from observing the inoculated plants. In the cases of the *ccmC* mutants, where there was a loss of nitrogen fixing ability or a reduced ability to form fixing nodules (Rm11425), it might be expected that these bacteria would be impaired in forming nodules compared with the wild-type or were unable to enter nodules altogether. The data do support the idea that the *ccmC* mutants are severely impaired in their ability to colonize nodules relative to the Rm1021 strain. However, the data does not support the theory that these bacteria are unable to enter nodules as mutant strains were recoverable from some nodules in each case (this was also observed in earlier experiments to determine if the strains were able to fix atmospheric nitrogen where plants were inoculated with a single bacterial strain). This is similar to the results with the non-nitrogen fixing strains. In each case, the mutant strains were easily outcompeted by the wild type though the mutants were able to enter nodules suggesting that the impairment of colonization is both obvious (lack of pink nodules, pale yellow plants) and more subtle (reduced ability to enter nodules). Interestingly, all strains showed reduced nodulation competitiveness relative to the wild type strain even those that were able to establish nitrogen fixing nodules in numbers similar to plants inoculated with the wild-type Rm1021 strains (for example Rm11424). This suggests that even though these strains are not impacted in terms of their ability to fix atmospheric nitrogen they are impacted in terms of their ability to enter nodules in the presence of another non-mutated symbiont. This observation may suggest that the ability of the wild-type to utilize purines like hypoxanthine gives it an advantage over non-purine users. It seems possible that the wild-type strain may be scavenging purines from dead bacteria or from the plant and using these as a supplemental source of carbon and/or nitrogen to assist in powering its invasion of the plant. The relative lack of competitiveness of the nitrogen fixing but non-

hypoxanthine utilizing strains in all cases seems to support this theory.

The nodulation kinetics study (Fig. 3.3) recorded how long it took for all plants inoculated with a particular *S. meliloti* strain to form nodules. The expectation from this study would be that mutant strains with a deficiency in their symbiotic phenotype may show a delay in nodulating all of the plants or experience a delay on achieving 100% nodulation compared with those inoculated with the wild-type strain. In the case of plants inoculated with Rm11424 (the SMb21292 mutant), it was observed that this strain actually achieved 100% of plants becoming nodulated faster than any other strain. Clearly, mutating this gene did not have any negative impact upon the symbiotic phenotype in terms of the bacteria quickly invading the plant and quickly establishing nitrogen fixing nodules on the plant root though as mentioned above it did display reduced competitiveness in the presence of the wild-type Rm1021 strain. This observation is curious. One may expect there to be a relative delay in nodulation compared with Rm1021 if the inability to utilize purines was the only factor at work here. In this case, the rapid formation of nodules on Rm11424 infected plants is only observed for white, non-nitrogen fixing nodules which were observed to form on all plants more rapidly than with Rm1021. This formation of white nodules on non-nitrogen fixing mutants has been previously described by Paau and colleagues (Paau *et al.*, 1985). The appearance of pink nitrogen fixing nodules did in fact show a small lag (about 1 day) when compared with Rm1021. This observation is consistent with the purine as supplement theory though it leaves open the question of why non-fixing nodules are formed so quickly with this strain. In the case of plants inoculated with Rm11429 (conserved hypothetical membrane protein), it was observed that the appearance of nodules on all plants was delayed and did not occur until around day 8. In this case, the appearance of pink nodules on plants was delayed compared to Rm1021 plants which would be consistent

with purine degrading strains having an advantage over non-purine degrading strains in terms of the rate at which nodulation is achieved. In the case of the fix positive *ccmC* mutant strain Rm11425 the situation was similar to Rm11424 with an early appearance of white nodules but a slight delay in the appearance of pink nodules relative to the wild-type inoculated strains. In the case of the remaining two *ccmC* mutants nodulation was delayed (Rm11429) or did not occur on all plants (Rm11421) and the lack of pink nodules again was consistent with the non nitrogen fixing phenotype of these strains. These results suggest that these two *ccmC* mutants were more severely impacted in terms of being able to colonize plants as they were slower in forming nodules than all the other strains suggesting that these *ccmC* mutations resulted in the most severe impairment of symbiotic phenotype due to the double impact of not being able to utilize hypoxanthine nor being able to fix atmospheric nitrogen. The loss of ability to fix atmospheric nitrogen in *S. meliloti* strains in which key genes encoding electron transport chain components have been mutated is not surprising (Delgado *et al.*, 1998). The rhizobia have a branched electron transport chain whereby there are terminal oxidases with different oxygen affinities. Reducing equivalents from different energy sources enter a quinone pool and are subject to an initial oxidation by the bc_1 cytochrome complex. The resulting electrons derived from the initial oxidation are transferred either to cytochrome *aa* through the membrane spanning cytochrome *c* or to cytochrome cbb_3. An alternate route involves the bypassing of the bc_1 complex by the reduction of quinol oxidase. While the alternative electron transport chain may allow the bacteria to survive under unusual varying environmental conditions the cytochrome cbb_3 route seems to be essential for the generation of sufficient ATP to drive the energy demanding process of nitrogen fixation. Previous work has also shown that mutating genes (SMc03849 for example) involved in the formation of the cytochrome c resulted in a loss of nitrogen fixing ability (Pobigaylo *et al.*, 2008). This is presumably due to the inability of electrons to be transferred to cytochrome aa_3.

Analysis of the hypoxanthine utilization strains by acetylene reduction assay (measuring the conversion of acetylene to ethylene by nitrogenase) mirrored, and confirmed, the results seen with the nodule count assays. The fix positive strains Rm11424, Rm11425 and Rm11429 showed levels of enzyme activity similar to the Rm1021 wild type strain. These results demonstrated that the plants with a normal healthy appearance (green leaves, pink root nodules) contained nodules containing active nitrogenase. Those plants inoculated with Rm11421 and Rm11428 had the appearance of plants suffering from nitrogen deficiency such as yellow leaves. The lack of nitrogenase activity in root nodules on these plants confirmed the inability of these strains to fix atmospheric nitrogen.

Chapter 4

Purine Mutants

4.1 Introduction

A wide range of bacteria, including *S. meliloti*, are able to utilize purines as a source of carbon and/or nitrogen. Purines can be utilized by bacteria during periods when supplies of carbon or nitrogen containing nutrients are limited. Purines, such as adenine and guanine, can be made available to bacteria either from inside the cell as a result of the degradation of unstable RNA-DNA molecules or from the external environment when released or excreted from decaying biological tissue or cellular material (Burton, 1994; Nygaard, 1983).

The objective of this study was to analyze the expression of genes predicted to be involved in the degradation of purines in *S. meliloti* when grown in minimal media containing a purine as a source of carbon and/or nitrogen. In addition the goal was to determine the phenotype associated with mutations in some of these genes. This would allow for the deter-

mination of the validity of the annotation of these genes in addition to allowing the prediction of their position and role in the purine degradation pathway of *S. meliloti*. To do this a series of *lacZ* fusions to putative purine degradation genes located in a 20 kb region of the pSymB megaplasmid of the *S. meliloti* genome were created. A search using the Toulouse annotation website (http://iant.toulouse.inra.fr/bacteria/annotation/cgi/rhime.cgi) previously mentioned reveals that this region contains numerous genes whose sequences suggest a role in purine catabolism as well as genes whose function cannot be theorized just from sequence similarity to other non-*S. meliloti* genes. In this region of the genome there is the operon containing the *xdhA1*, *xdhB1* and *xdhC* genes. In addition there is a LysR transcriptional regulator, *guaD1*, as well as a uricase like enzyme and a putative membrane transport protein. The construction of such fusion strains allowed us to measure expression of the gene of interest under different conditions including growth on different purines, purine derivatives, carbon sources or inducer molecules. In several cases creation of the fusion simultaneously created a mutation of the gene whose promoter region was fused to the *lacZ* reporter gene allowing us to study the phenotypic impact of a particular mutation upon purine catabolism. In addition to reporter fusion derived mutations we were able to mutate additional genes of interest by inserting a plasmid containing a portion of a gene of interest resulting in a truncation and mutation of the gene.

4.2 Materials and Methods

Materials and methods used in the chapter are described in Chapter 2.

4.3 Results

4.3.1 Sequence Analysis

Sequence analysis of these genes using BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) (expect threshold = 10, word size = 28, match/mismatch scores = 1,-2, gap costs = linear) reveals varying levels of similarity with genes from other bacteria. The *xdhA1* gene for example has a high degree of similarity to annotated genes in other rhizobia such as *S. medicae* and *Rhizobia* spp. NGR234 (identities of 90 and 83 % respectively). It also shows high levels of similarity to genes annotated as xanthine dehydrogenase in closely related species such as *Brucella* with identities in the 73-74 % range and *Burkholderia multivorans* spp at 77 %. In all cases however, the classification of the sequences as *xdhA1* is based on sequence similarity not direct experimental analysis.In each case the protein sequences are predicted to a contain [2Fe-2S], a FAD binding and a CO dehydrogenase flavoprotein C-terminal domain. The *xdhC* gene on the other hand only shows a high degree of similarity (89% identity) to a few other sequences. These include a xanthine dehydrogenase accessory protein in *Sinorhizobium medicae* also annotated as *xdhC*. The next most similar genes are again found in the closely related *Rhizobium* species NGR234 which has an identity score of 78% and is annotated as a putative xanthine dehydrogenase cofactor. The same sequence has a similar annotation in *Ochrobactrum anthropi* where its closest match is annotated as a xanthine dehydrogenase accessory protein *xdhC* though the sequence similarity begins to fall off at this point (69%). Interestingly, in *Brucella melitensis* the closest match is annotated as an XdhC protein assisting in molybdopterin insertion into xanthine dehydrogenase. Each of these sequences contains a predicted XdhC/CoxI domain. Again in all these cases the annotation is based on sequence similarity not direct experimental evidence. In the case of *guaD1* the most similar sequences are again found in closely related species

including *S. medicae* WSM419 (also annotated as a guanine demaminase with an e-value of 0 and an identity of 86%). This is followed by a match in Rhizobium species NGR234 which is again annotated as a guanine deaminase (with an e-score of 0 and an 84% identity) and a guanine deaminase in *Rhizobium leguminosarum* bv. *trifolii* WSM2304 (e-value equals $4x10^{-54}$ with an identity of 71%). All these sequences are characterized by a predicted amidohydrolase domain. Again all these annotations are based on sequence similarity but direct experimental evidence for these particular genes functioning as a guanine deaminase is lacking. In the case of SMb21284, annotated as a uricase-like protein in the *S. meliloti* genome sequence again the closest matching sequences come from species closely related to *S. meliloti* Rm1021. These include *S. medicae* WSM419 where the closest matching sequence is annotated as a chitin deacetylase (with an e-score of 0 and an identity of 89%). This is followed by a match in *Rhizobium* species NGR234 annotated as a uricase (with an e-score of 0 and an identity of 84%). This is followed by an unannotated sequence from the pNGR234b plasmid of *Mesorhizobium loti* (e-score $1x10^{63}$, identity equals 75%). In each case sequences are predicted to include two distinct conserved domains, a polysaccharide deacetylase, a OHCU decarboxylase. This gene is annotated as a hypothetical protein or as a putative urate catabolism protein. Again direct experimental results with these closest matching genes, as with *S. meliloti*, are not available. *S. meliloti* SMb21281 is annotated as an inner membrane permease. Its most closely related sequences as determined by BLAST are from *S. medicae* WSM419 where it is classified as a xanthine/uracil/vitamin C permease (e-score = 0, identity = 89%) and from *Rhizobium* species NGR234 where it has the same annotation as the sequence in *S. medicae* WSM419 (e-score = 0, identity = 88%). These sequences are characterized by having a conserved permease family domain. In each of these cases this annotation was again not based on direct experimental evidence with the species in question. The only experimental evidence that exists to support this comes from *Bacillus subtilus* (Schultz *et al.*, 2001) and *E.coli* K-12 (Loh *et al.*, 2006). In each

case, none of the sequences in question from these species are among the closest matches as determined from megaBLAST analysis nor are they among those classified as being somewhat similar sequences based on BLASTn results.

4.3.2 Organization of *xdh* Genes in Species Closely Related to *S. meliloti*

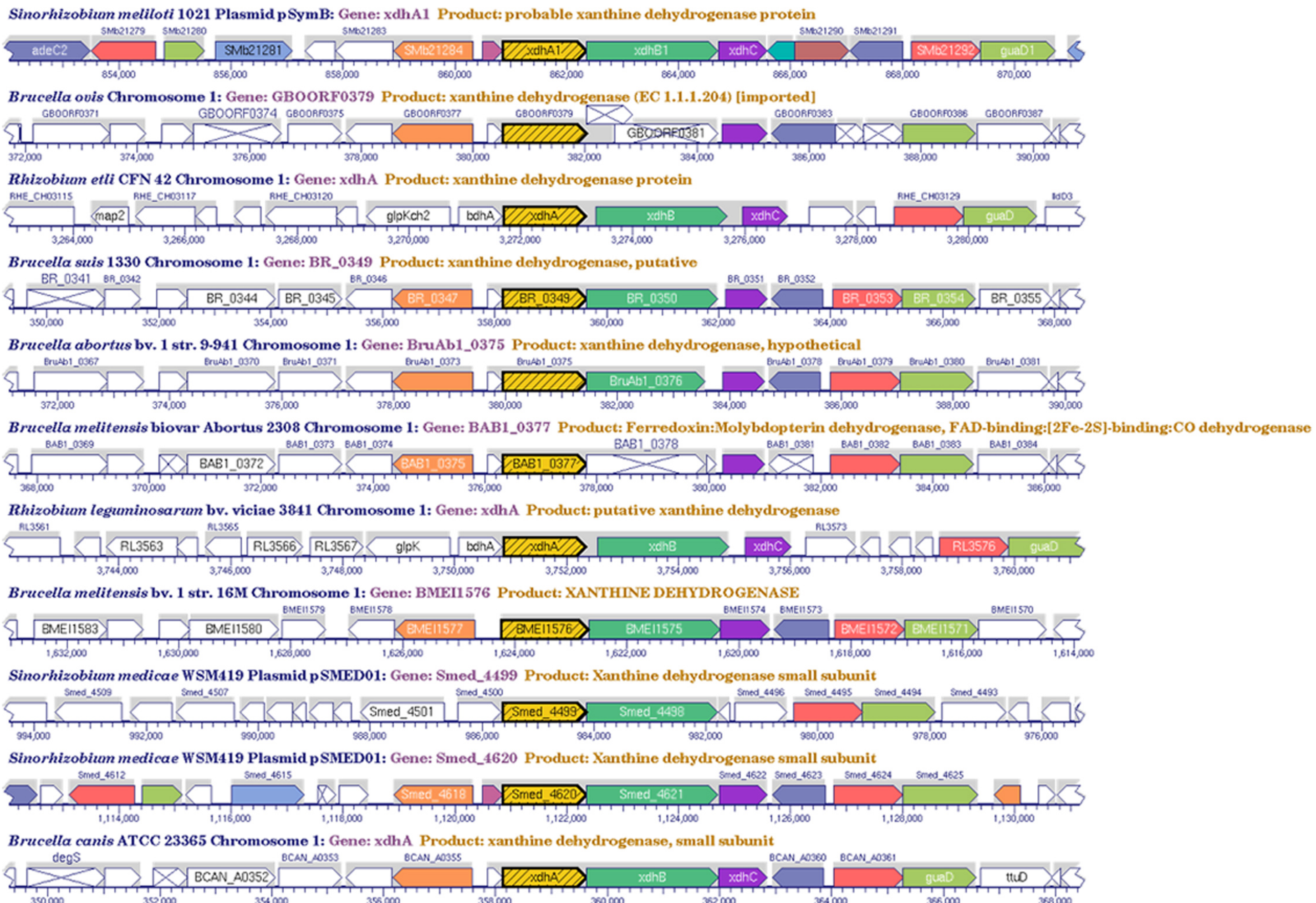

Figure 4.1: Organization of *xdh* genes in bacterial species closely related to *S. meliloti* obtained using biocyc.org

An examination of the arrangement of *xdh* genes in species closely related to *S. meliloti* using biocyc.org shows some interesting results (Fig. 4.1). In *S. medicae* the organization of the *xdh* genes is similar to that seen in *S. meliloti* with Smed4620,4621 and 4622 having the same order and annotation as found for *xdhA1xdhB1xdhC* in *S. meliloti.* The gene found upstream of this operon in *S. medicae* is however, a chitin deacetylase not a uricase. *S. medicae*, unlike the other species shown in Fig. 4.1, contains two copies of *xdhA* and *xdhB* (Smed4499 and Smed4620 being equivalent to *xdhA* and Smed4498 and Smed4421 being equivalent to *xdhB*. Like *S. meliloti*, *S. medicae* only contains a single *xdhC* gene, Smed4622. In some species such as *Rhizobium etli* and *R. leguminosarum* the gene organization appears to be a hybrid of the arrangement found in *S. meliloti.* In *S. meliloti* the *bdhA* and *xdhC* genes are found in distinct operons that include *xdhA2xdhB2* and *xdhA1xdhB1* respectively. In *R. etli* and *R. leguminosarum* the *bdhA* and *xdhC* genes are located in the same operon. In *R. etli xdhC* is the next gene following the operon containing *bdhAxdhAxdhB*. In both of these species the gene upstream of the *bdhAxdhAxdhB* operon is annotated as a glycerol kinase, not a uricase as in *S. meliloti.* In *Brucella ovis* the gene arrangement is similar to *S. meliloti* with *xdhA* and *xdhC* being found consecutively. However in this case *xdhB* seems to be missing as the space between *xdhA* and *xdhC* contains a region annotated as a pseudogene. The gene upstream of *xdhA* in the opposite orientation is annotated as a polysaccharide deacetylase. In *Brucella suis* the arrangement of *xdhAxdhBxdhC* (BR0349,0350 and 0351) is similar to that seen in *S. meliloti* except the xdhC (BR0351) is not predicted to be in the same operon. The upstream gene running in the opposite orientation has no predicted function and is annotated as an open reading frame. The same arrangement can be seen in *Brucella abortus.* In *Brucella melitensis* are arrangement is similar except the upstream, opposite gene is annotated as a chitoolligosaccharide deacetylase.

4.3.3 Strain RmF117

To gain an understanding of the importance of the region of pSymB containing several putative purine degradation genes in allowing *S. meliloti* to degrade purines the phenotype of the RmF117 strain was investigated. This strain, which has part of the pSymB megaplasmid containing the genes such as *xdhA1* deleted (Charles and Finan, 1990), was unable to grow on media containing purines such as hypoxanthine or xanthine as a sole source of carbon and or nitrogen. When grown on M9 containing either purine there was no visible growth or clearing of hypoxanthine or xanthine observed. This differed from the wild-type SmP110 strain which when grown on the same media showed both growth and clearing of both purines. This result demonstrated the requirement of genes from this part of the pSymB megaplasmid for purine degradation in *S. meliloti*.

4.3.4 Expression of *lacZ* Fusion Strains

Having established the requirement of a region of the pSymB megaplasmid for purine degradation more detailed experiments were undertaken. The goal of these studies was to characterize expression in response to purines by several genes predicted to be involved in purine degradation in *S. meliloti*. To accomplish this, strains that were constructed with a *lacZ* fusion to the promoter region of a given purine degradation gene were utilized.

4.3.5 *xdhA1* Fusion

The *xdhA1 lacZ* fusion strain (Fig. 4.3) showed a high level of *lacZ* expression relative to the background, wild-type, SmP110 strain when grown on M9 minimal media containing

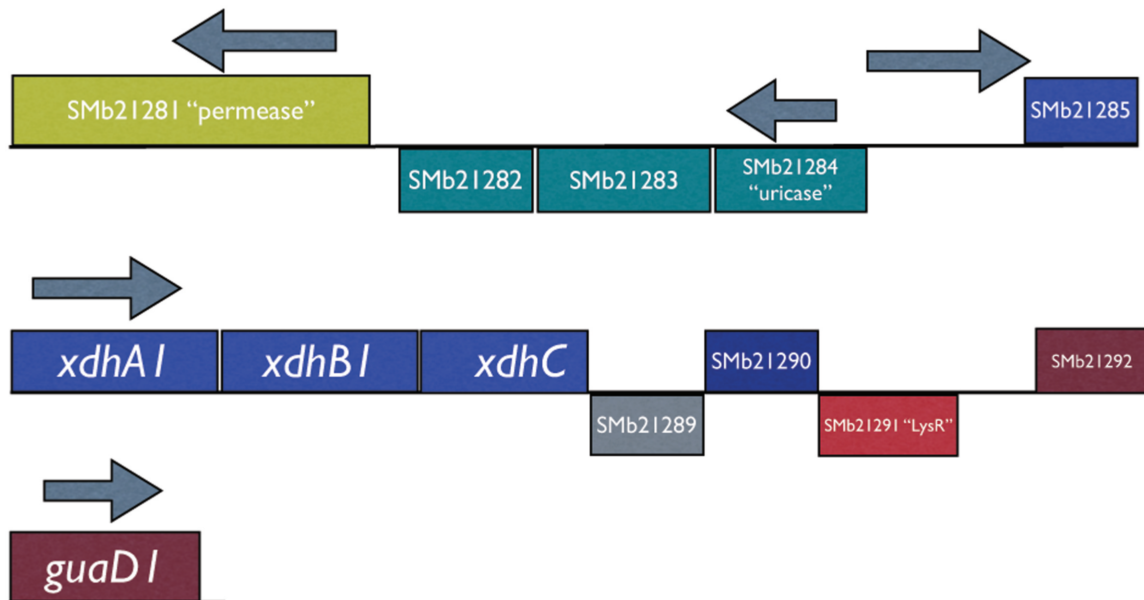

Figure 4.2: Relative location of putative purine degradation genes deleted in *S. meliloti* RmF117 strain. Genes with the same colours are predicted to be in the same operon. Arrows represent the location and direction of *lacZ* fusions in the respective strains discussed in the text. SMb21291 = amino acid accession number CAC49195.

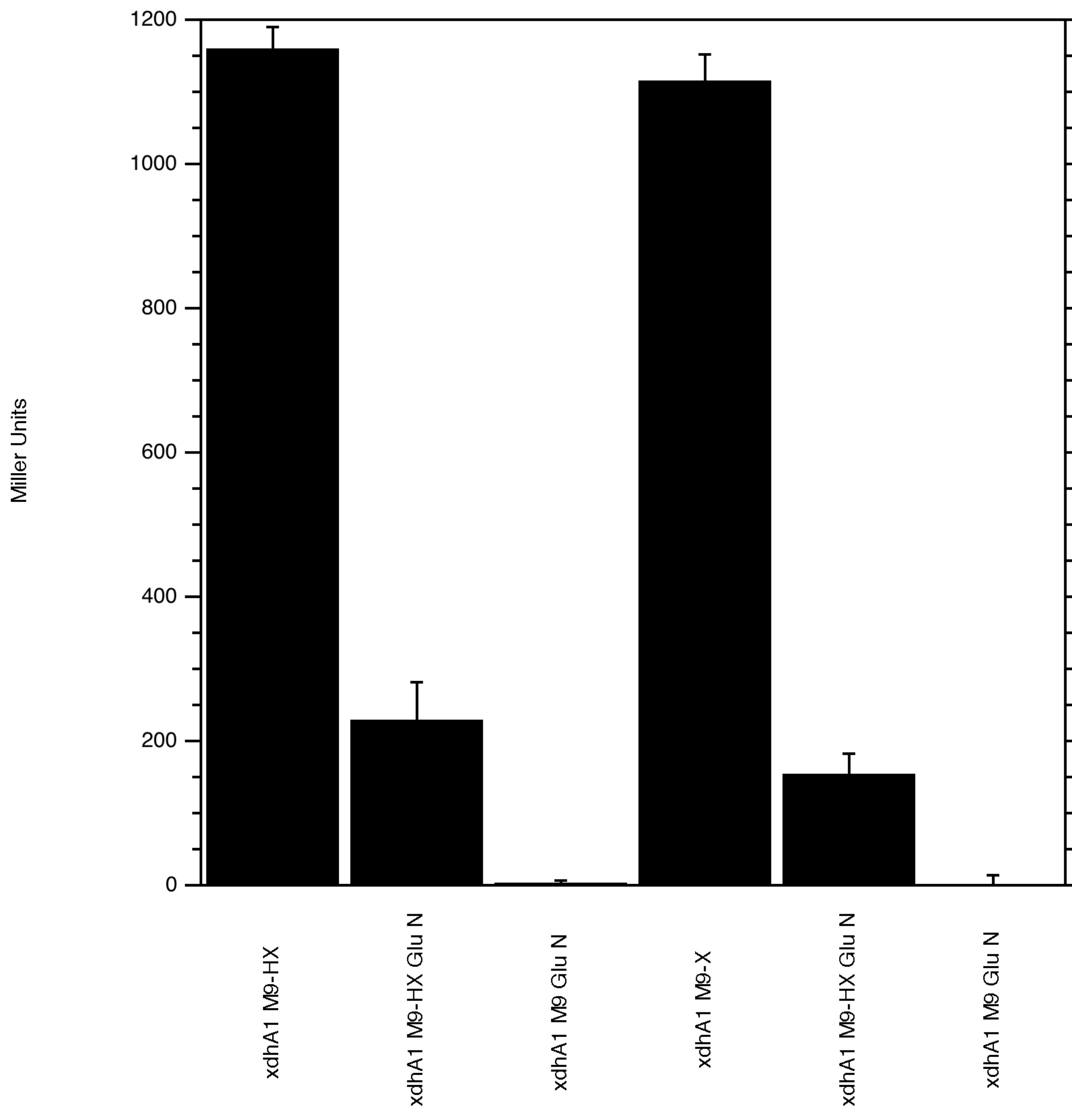

Figure 4.3: *lacZ* expression for SmRL2522 (*xdhA1lacZ* fusion strain) grown in various M9 minimal media. Expression of control strain SmP110, grown in each medium, is set as 0. M9 = M9 minimal media, HX = hypoxanthine, X = xanthine, Glu = glucose, N = ammonium chloride. xdhA1 = *xdhA1* fusion strain. N = 22.

hypoxanthine or xanthine as the sole source of available carbon and nitrogen. Both fusion strains showed close to 1200 Miller units of activity when grown in this media with the reactions turning bright yellow rapidly (within approximately 5 minutes after starting the reaction). When grown in the same M9 minimal media containing glucose and ammonium chloride as the respective carbon and nitrogen sources activity was the same as the non-fusion wild type strain SmP110. Expression in M9 media containing hypoxanthine or xanthine but also containing glucose and ammonium chloride showed expression levels lower than M9 with hypoxanthine or xanthine alone but still significantly higher then the fusion or wild-type strains in M9 lacking either of these purines.

4.3.6 SMb21285 Fusion

The strain containing a *lacZ* fusion to SMb21285 (first gene in the operon containing *xdhA1B1C*) and including the intergenic region between SMb21284 (a uricase like protein) and SMb21285 (a hypothetical protein of unknown function) showed a high and significant level of expression when grown in M9 minimal media containing either hypoxanthine or xanthine as the only carbon and nitrogen sources (Fig. 4.4). Expression levels were lower though still high and significant when the fusion strain was grown in M9 minimal media with glucose and ammonium chloride as carbon and nitrogen sources in addition to either hypoxanthine or xanthine. In each case, expression was far higher than when the fusion strain was grown in M9 containing glucose and ammonium chloride but lacking either of the purines. In this case the levels of *lacZ* expression were similar to the levels observed for the SmP110 wild-type strain when grown in M9 with or without hypoxanthine or xanthine.

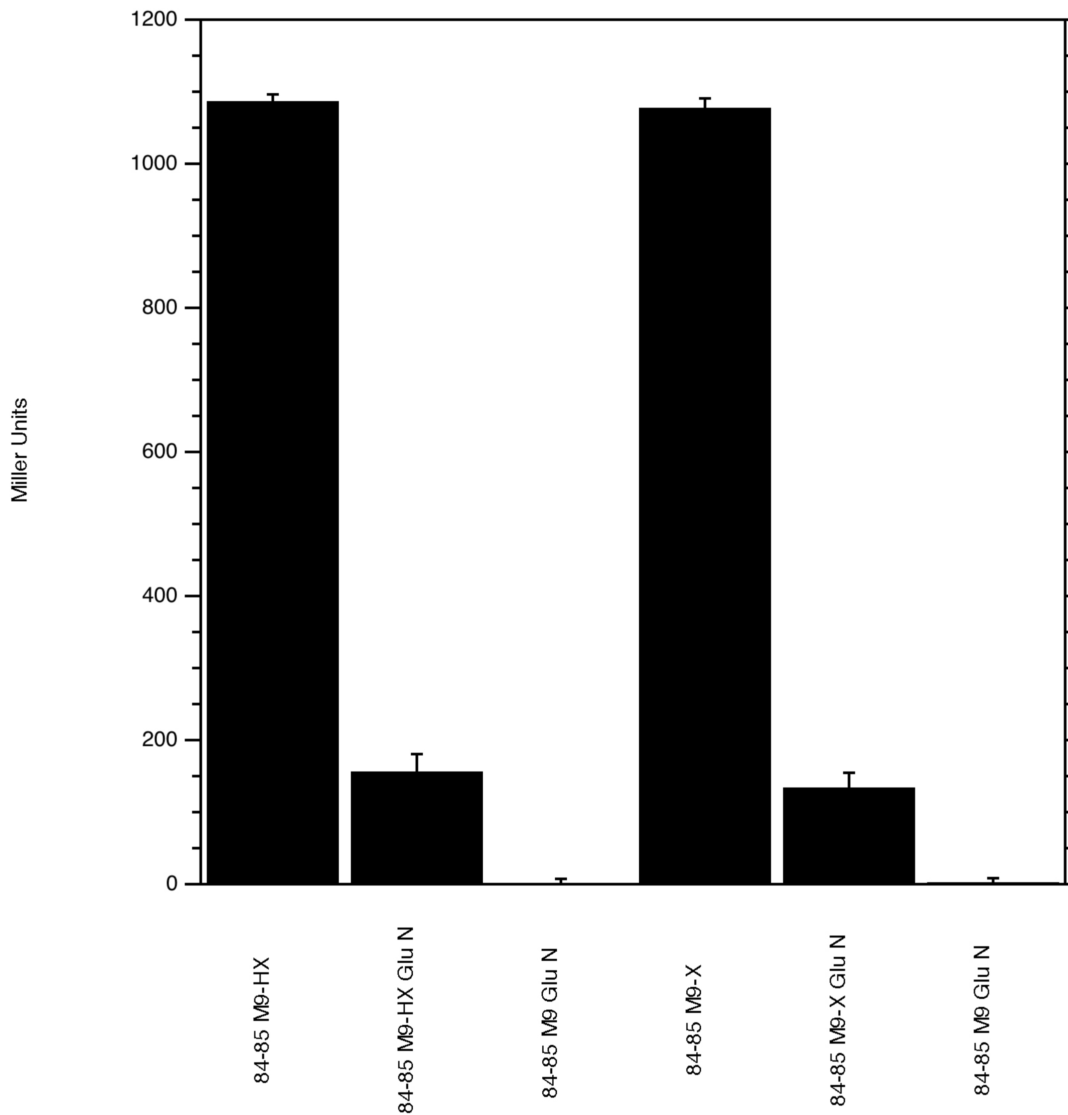

Figure 4.4: *lacZ* expression for Rm11513 (SMb21285*lacZ* fusion strain) grown in various M9 minimal media. Expression of control strain SmP110, grown in each medium, is set as 0. M9 = M9 minimal media, HX = hypoxanthine, X = xanthine, Glu = glucose, N = ammonium chloride. 84-85 = SMb21285*lacZ* fusion strain. N = 22.

4.3.7 *guaD1* Fusion

Strain Rm11522 containing a *lacZ* fusion to the *guaD1* gene when grown in M9 minimal media (Fig. 4.5) containing guanine showed high levels of *lacZ* expression compared with the same strain grown in M9 without either purine or the SmP110 wild-type background strain grown in media with or without added purines (guanine, hypoxanthine or xanthine). When the fusion strain is grown in M9 containing hypoxanthine or xanthine there is still a significant level of expression compared with the control although this level is much lower than when the strain was grown in the same media containing guanine instead.

4.3.8 SMb21284 Fusion

Strain Rm11517, containing a *lacZ* fusion to a predicted uricase like gene (SMb21284), when grown in M9 minimal media containing uric acid shows a significantly elevated level of expression (Fig. 4.6) compared with the same strain grown in M9 with glucose and ammonium chloride as the only carbon and nitrogen sources. Expression is also similarly higher when compared with the background expression levels seen in SmP110 when this wild-type strain is grown in M9 with uric acid or purines (hypoxanthine or xanthine). As with the Rm11522 (*guaD1* fusion) when the strain is grown in M9 which also contains the purines hypoxanthine or xanthine there is significantly elevated *lacZ* expression though again much lower than when the media contains uric acid. The Rm11517 strain grows very poorly when grown in M9 containing uric acid as the only carbon or nitrogen source.

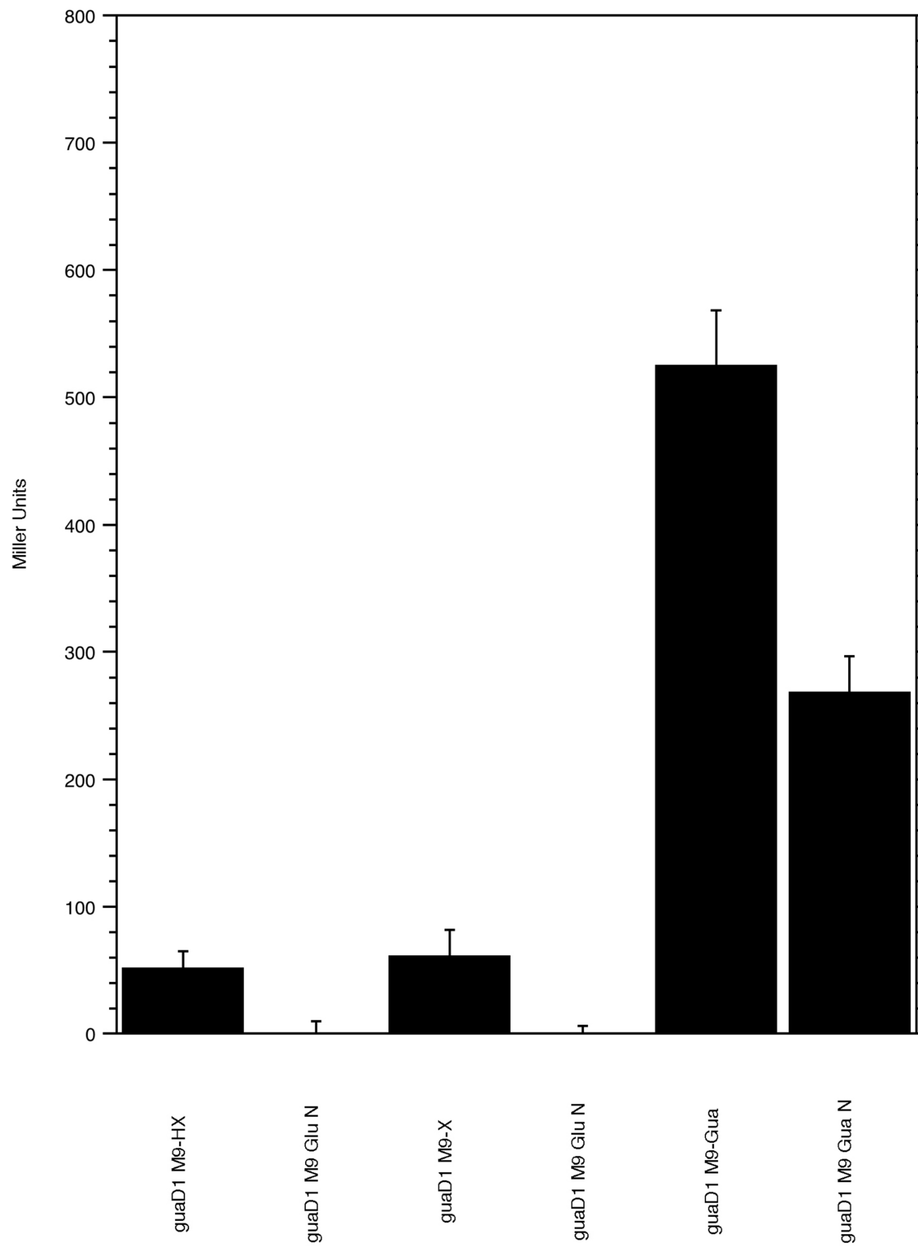

Figure 4.5: *lacZ* expression for Rm11522 (*guaD1lacZ* fusion strain) grown in various M9 minimal media. Expression of control strain SmP110, grown in each medium, is set as 0. M9 = M9 minimal media, HX = hypoxanthine, X = xanthine, Glu = glucose, N = ammonium chloride, gua = guanine. guaD1 = *guaD1lacZ* fusion strain. N = 22.

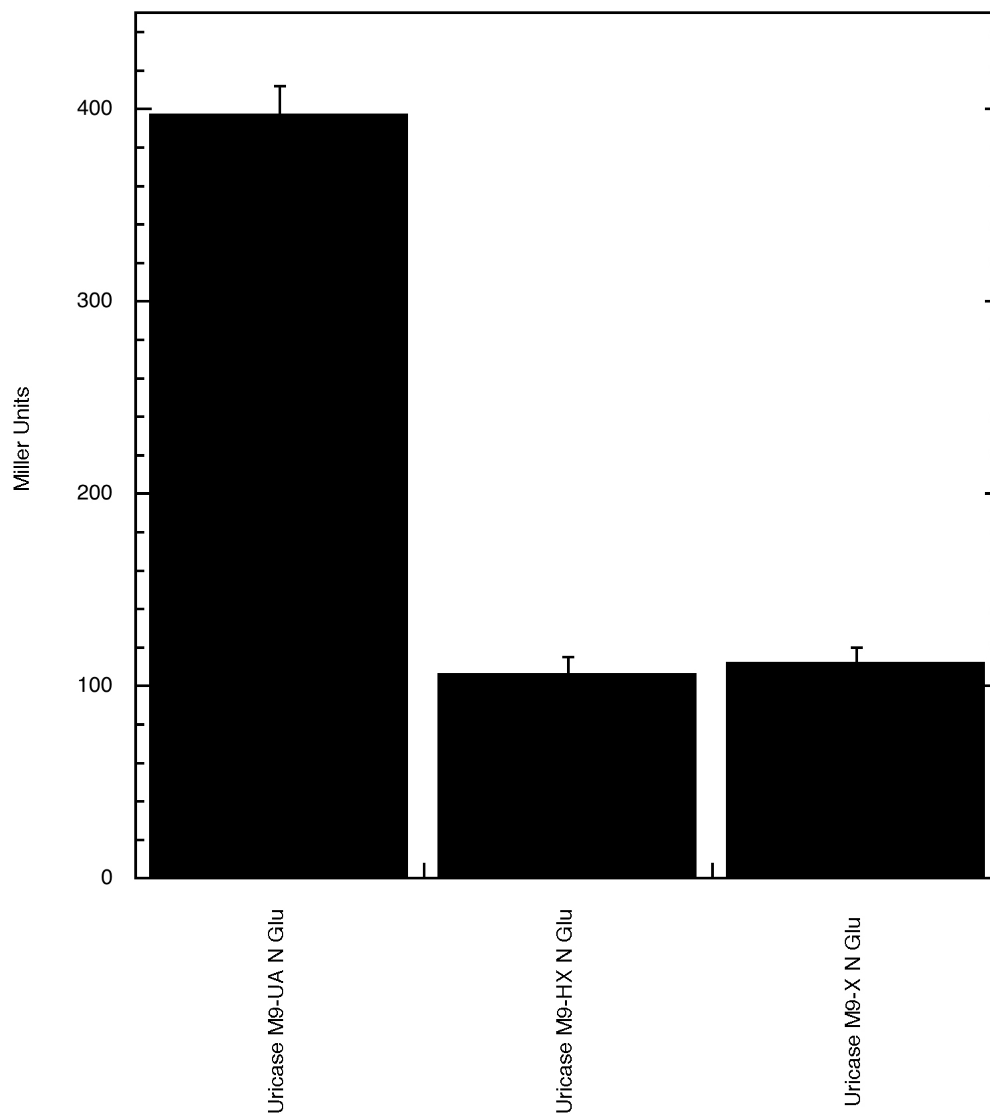

Figure 4.6: *lacZ* expression for Rm11517 (SMb21284("uricase like protein")*lacZ* fusion strain grown in various M9 minimal media. Expression of control strain SmP110, grown in each medium, is set as 0. M9 = M9 minimal media, HX = hypoxanthine, X = xanthine, Glu = glucose, N = ammonium chloride, UA = uric acid. SmP110 = *S. meliloti* strain SmP110, uricase = SMb21284*lacZ* fusion. N = 22.

4.3.9 SMb21281 Fusion

Strain Rm11519, containing a *lacZ* fusion to a predicted membrane transporter permease, does not show an elevated level of expression when grown in M9 containing hypoxanthine or xanthine (Fig. 4.7). Expression levels were similar to those of the SmP110 wild-type strain grown in the same media or M9 with glucose and ammonium chloride or the Rm11519 strain grown in M9 with glucose and ammonium chloride.

4.4 Analysis of *lacZ* Fusions to Purine Degradation Genes Found Elsewhere in the *S. meliloti* Genome

In addition to examining fusions to genes found in the genomic region encoding XdhA1/XdhB1, two additional fusions were constructed in the region of the *S. meliloti* genome encoding XdhA2/XdhB2. These were fusions to *xdhA2* and *guaD2*.

4.4.1 *xdhA2* Fusion

The *xdhA2* gene is the second gene in an operon that begins with *bdhA*. Growth of the *xdhA2lacZ* fusion in media containing hypoxanthine or xanthine did not result in a measurable increase in *lacZ* expression compared with the the fusion strain grown in M9 media lacking either of these purines (Fig. 4.8). Levels of expression were higher in all three cases than when SmP110 was grown in the same types of M9. This result suggested that

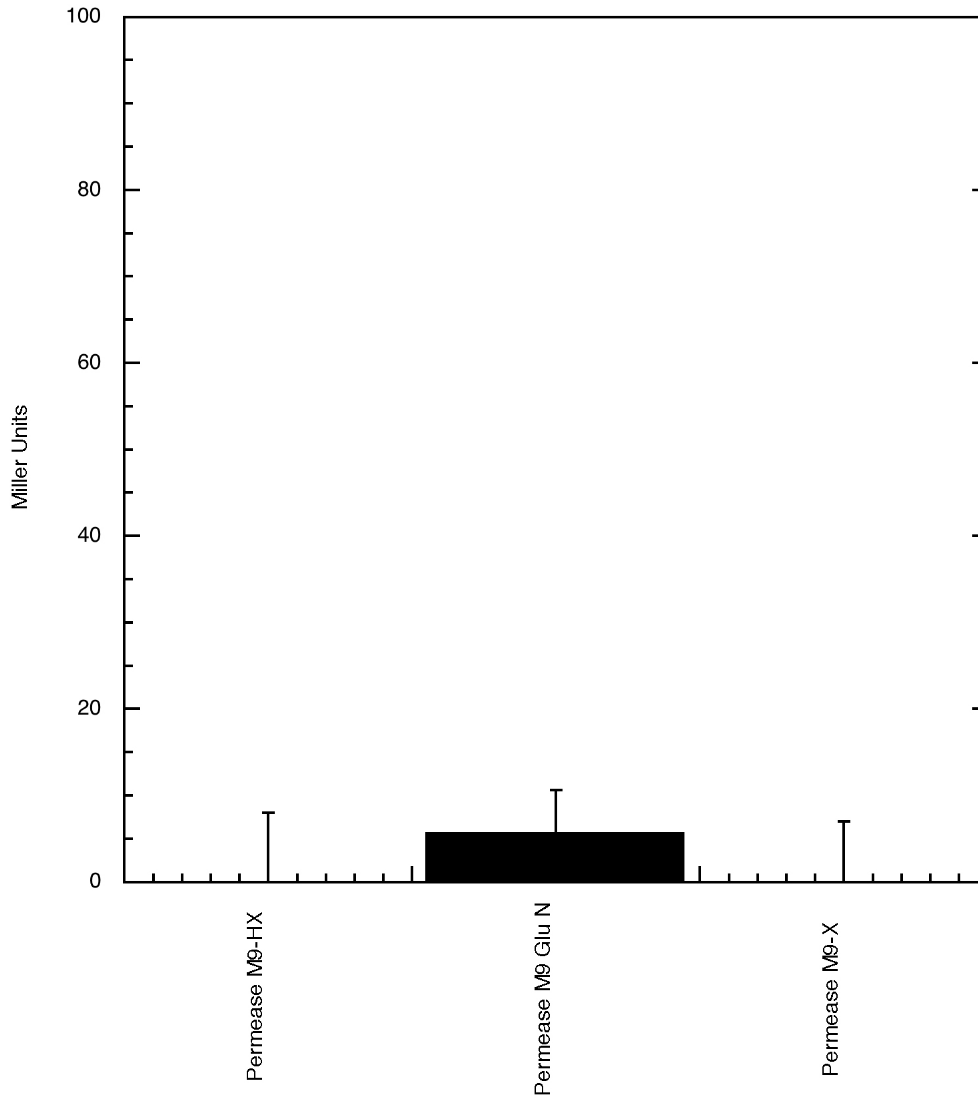

Figure 4.7: *lacZ* expression for Rm11519 (SMb21281*lacZ* fusion strain) grown in various M9 minimal media. Expression of control strain SmP110, grown in each medium, is set as 0. M9 = M9 minimal media, HX = hypoxanthine, X = xanthine, Glu = glucose, N = ammonium chloride. Permease = SMb21281*lacZ* fusion. N = 22.

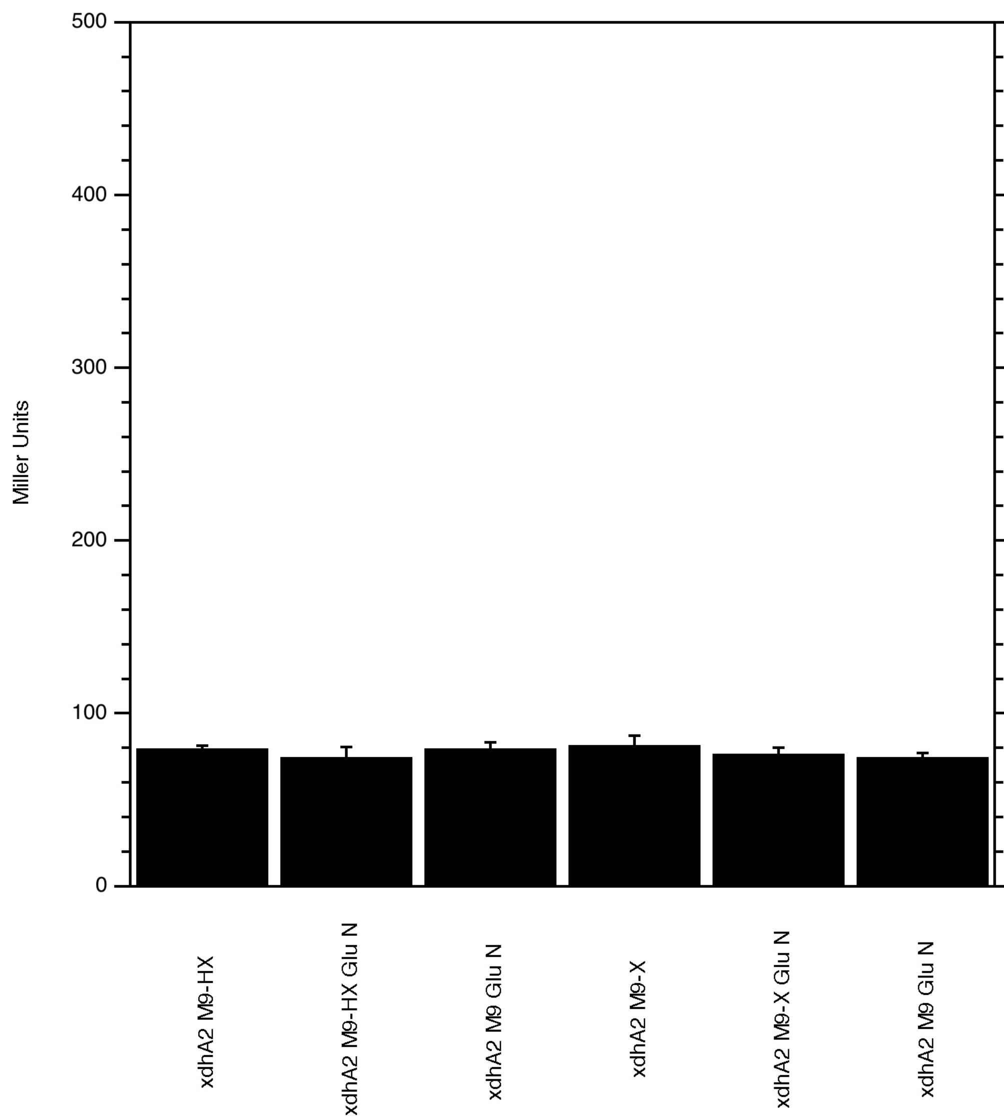

Figure 4.8: *lacZ* expression for Rm11518 (*xdhA2lacZ* fusion strain) grown in various M9 minimal media. Expression of control strain SmP110, grown in each medium, is set as 0. M9 = M9 minimal media, HX = hypoxanthine, X = xanthine, Glu = glucose, N = ammonium chloride. xdhA2 = *xdhA2lacZ* fusion strain. N = 22.

something other than purines may act as an inducer of this operon or that this gene has a higher basel level of expression.

4.4.2 *guaD2* Fusion

Not surprisingly the results from the analysis of the the *guaD2lacZ* fusion strain (Fig. 4.9) were quite similar to those of the fusion to *guaD1*. Again with this second copy of *guaD* the highest levels of expression were observed when the strain was grown in M9 minimal media containing guanine as the single carbon and nitrogen source. The addition of ammonium chloride as an alternate nitrogen source resulted in a significantly reduced level of expression though still well above that of the controls using the fusion strain grown in M9 without any added purine or the SmP110 wild-type strain grown with guanine in the media. As with the *guaD1* the presence of hypoxanthine or xanthine in the media resulted in an elevated level of *lacZ* expression though again this level of expression was far below that observed when guanine was in the media.

4.4.3 Phenotypic Analysis of Mutations in *xdhA1*, *xdhC* and SMb21291

In addition to *lacZ* fusion strains created to study changes in expression additional mutants were created in three specific genes, *xdhA1*, *xdhC* and SMb21291using single crossover mutagenesis and confirmed by DNA sequencing. This allowed for the analysis of the phe-

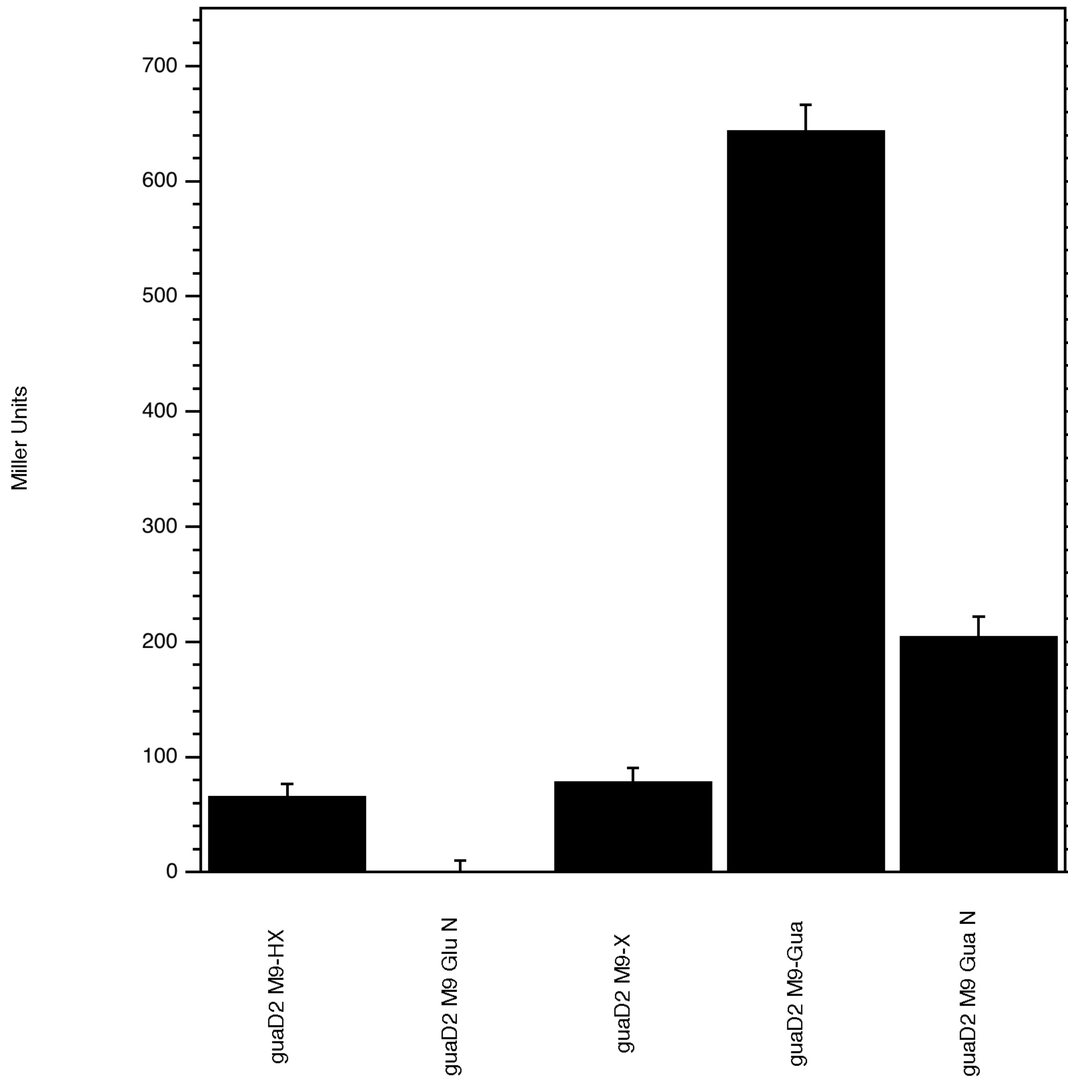

Figure 4.9: *lacZ* expression for Rm11518 (*guaD2lacZ* fusion strain) grown in various M9 minimal media. Expression of control strain SmP110, grown in each medium, is set as 0. M9 = M9 minimal media, HX = hypoxanthine, X = xanthine, Glu = glucose, N = ammonium chloride. SmP110 = *S. meliloti* strain SmP110, guaD2 = *guaD2lacZ* fusion strain. N = 22.

notypic impact of these mutations upon purine degradation in *S. meliloti*.

xdhA1 **Mutant**

The *xdhA1* gene is one of four genes found in the operon which also contains *xdhB1* and *xdhC*. The *xdhA1* mutant strain grows normally on rich media (TY broth) with no obvious growth differences from the SmP110 wild-type strain (Fig. 4.10). When inoculated in M9 minimal media with glucose and ammonium chloride as carbon and nitrogen sources the strain displayed growth indistinguishable from SmP110. When grown on M9 minimal media containing either hypoxanthine or xanthine however the *xdhA1* mutant shows a distinct phenotype. Unlike the SmP110 wild-type strain the *xdhA1* mutant is not able to grow on either purine as a sole source of carbon and or nitrogen. In liquid media there is no visible loss of either purine in the media or a detectable increase in culture density both of which are observable with SmP110. A similar result can be seen on M9 agar plates where purine crystals disappear as the SmP110 cells grow while there is no observable growth on plates inoculated with the mutant strain and purine crystals remain visible in the media.

Staining of native gels (Fig. 4.13) loaded with cell extract from the *xdhA1* mutant strain for xanthine oxidase/xanthine dehydrogenase activity revealed the absence of any detectable activity in the mutant strain with all three stained bands that are found in the SmP110 cell extract sample being absent from the lane containing the mutant strain cell extract.

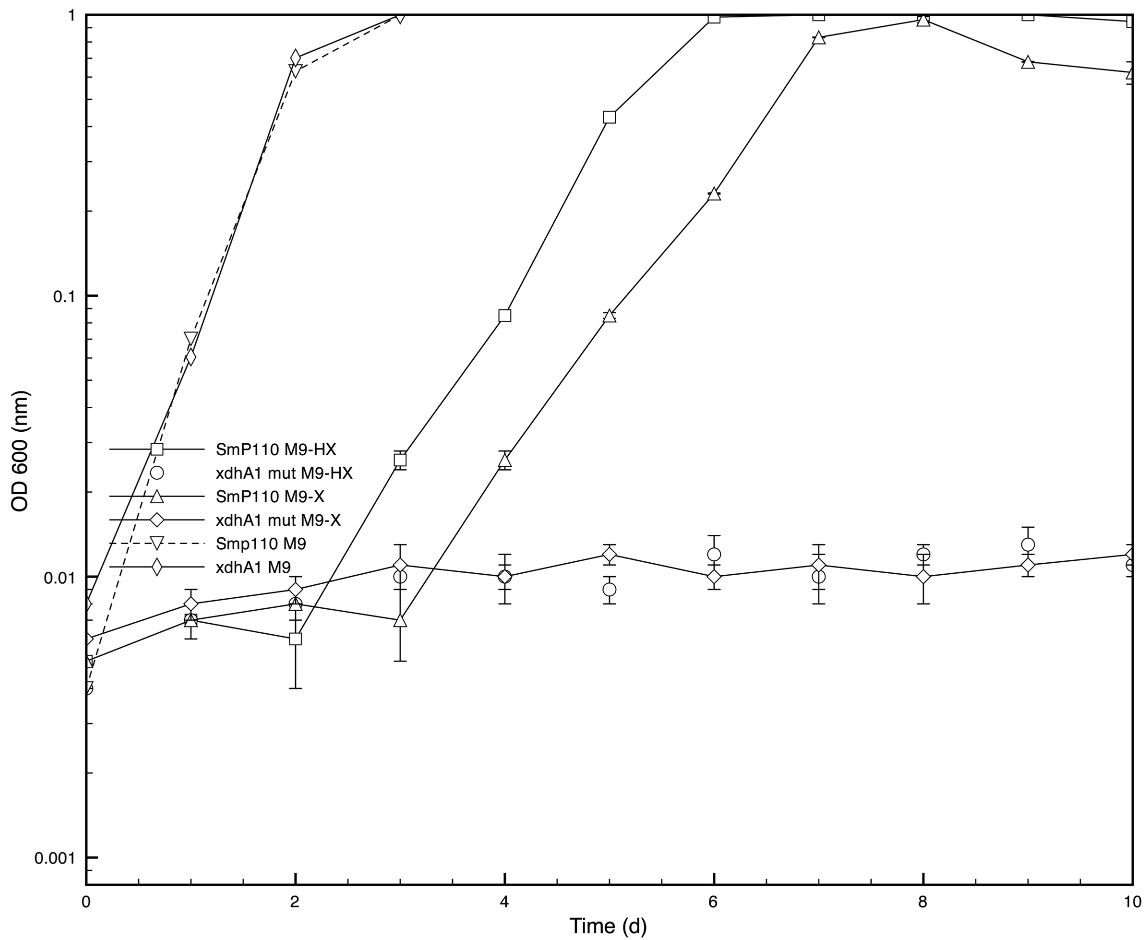

Figure 4.10: Growth of *S. meliloti xdhA1* mutant and SmP110 (wild-type) strains on M9 minimal media. M9 = M9 minimal media, HX = hypoxanthine, X = xanthine. SmP110 = *S. meliloti* strain SmP110, xdhA1 = *xdhA1* mutant strain. SmP110 M9 and xdhA1 M9 = the respective strains grown on M9 with glucose and ammonium chloride. Bars= +/- SE. N=3.

4.4.4 *xdhC* Mutant

The *xdhC* gene is the final gene in the predicted operon containing *xdhA1* and *xdhB1* and is believed to code for a protein that plays a role in helping to form the final functional protein but is not part of the Xdh protein itself. As was observed for the *xdhA1* mutant strain the *xdhC* mutant shows growth indistinguishable from the SmP110 wild-type strain when growing on M9 with glucose and ammonium chloride (Fig. 4.11). Growth in rich media (TY broth) is similarly indistinguishable. In each case, in minimal media hypoxanthine and xanthine crystals disappear as the SmP110 culture grows while the mutant strain does not grow and and the purine crystals remain.

When cell extract was collected from both the *xdhC* mutant and SmP110, electrophoresed through a non-denaturing acrylamide gel (Fig. 4.13) and stained for xanthine oxidase/xanthine dehydrogenase activity once again as with the *xdhA1* mutant there was a striking difference. The *xdhC* lacked any of the three stained bands that are present in the lane containing the SmP110 cell extract.

SMb21291 Mutant

The SMb21291 protein codes for a predicted LysR transcriptional regulator and is located among the genes predicted to play a role in purine degradation. Its role, if any, in purine degradation had never been studied and thus a strain with a mutation in this gene was made to study the possible role of SMb21291 in purine degradation. When grown on M9 minimal media with glucose and ammonium chloride as the carbon and nitrogen sources

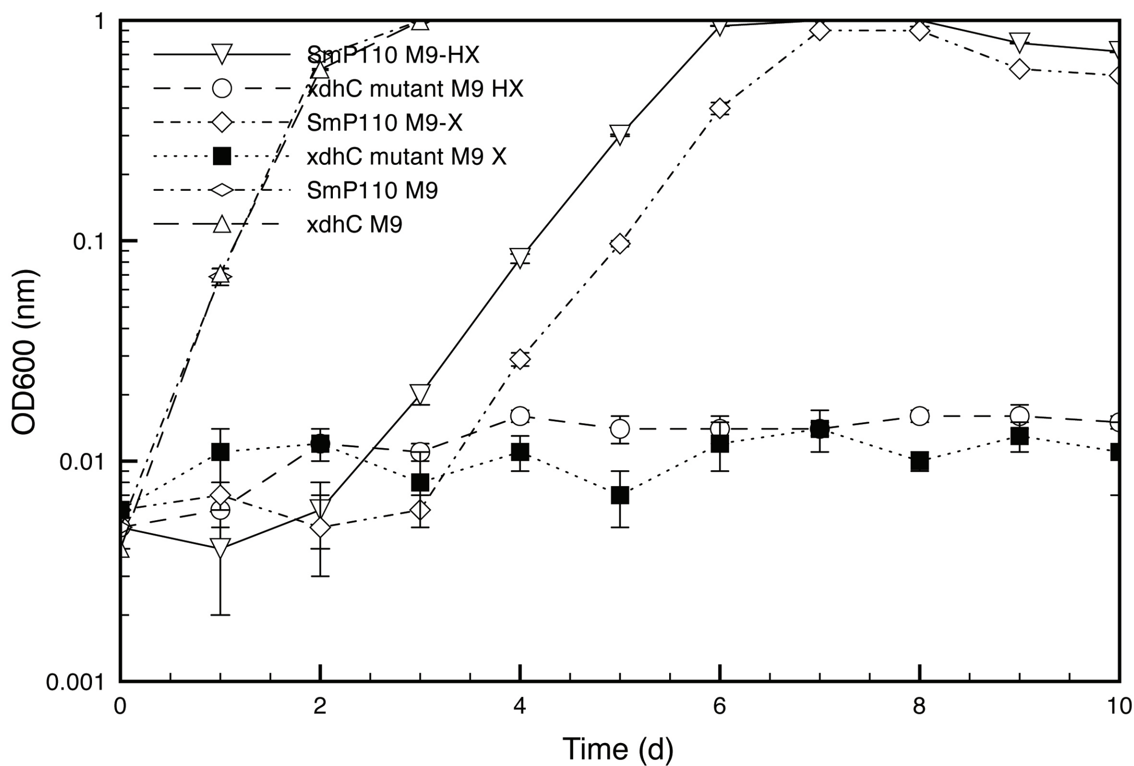

Figure 4.11: Growth of *S. meliloti xdhC* mutant and SmP110 (wild-type) strains on M9 minimal media. M9 = M9 minimal media, HX = hypoxanthine, X = xanthine. SmP110 = *S. meliloti* strain SmP110, xdhC = *xdhC* mutant strain. SmP110 M9 and xdhA1 M9 = the respective strains grown on M9 with glucose and ammonium chloride. Bars= +/- SE. N=3.

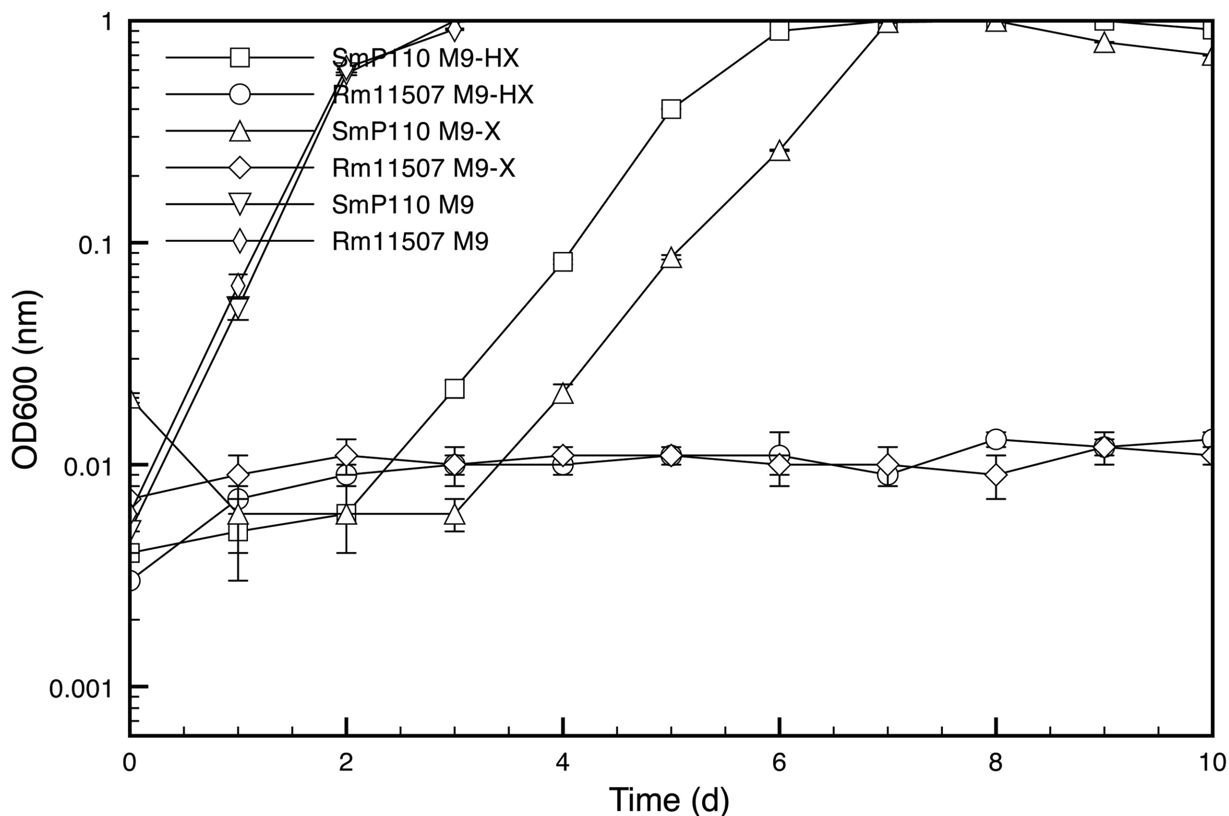

Figure 4.12: Growth of *S. meliloti* SMb21291 mutant (Rm11507) and SmP110 (wild-type) strains on M9 minimal media. M9 = M9 minimal media, HX = hypoxanthine, X = xanthine. SmP110 = *S. meliloti* strain SmP110. SmP110 M9 and xdhA1 M9 = the respective strains grown on M9 with glucose and ammonium chloride. Bars= +/- SE. N=3.

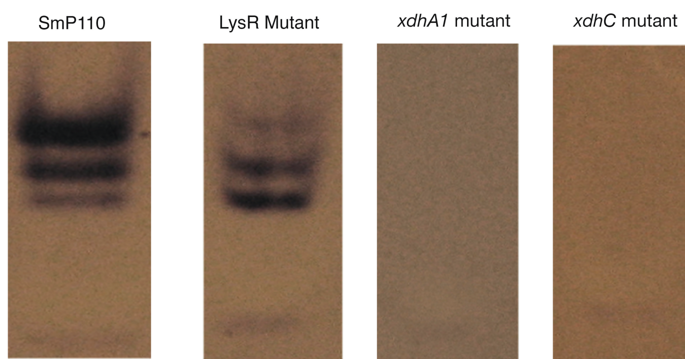

Figure 4.13: Staining of *S. meliloti* cell-free extract on a non-denaturing polyacrylamide gel for xanthine oxidase/xanthine dehydrogenase activity. From left to right SmP110, SMb21291 mutant, *xdhA1* mutant and *xdhC* mutant.

the SMb21291 mutant showed growth similar to the wild type SmP110 strain. This is also the case when grown in TY broth. When the SMb21291 mutant was inoculated in M9 minimal media with either hypoxanthine or xanthine as the only source of carbon and nitrogen the strain does not grow (Fig. 4.12) and purine crystals remain in the media. SmP110 grows on both purines with visible loss of purine crystals in the solution. On agar plates the situation is similar with the mutant strain being unable to grow with no apparent loss of purine crystals from the media unlike the wild-type strain.

Cell extract from the SMb21291 mutant run on a non-denaturing acrylamide gel (Fig. 4.13) and stained for xanthine oxidase/xanthine dehydrogenase activity reveals a unique pattern. While all three polypeptides are visible on the gel the high molecular weight polypeptide is very faint (though still visible) compared with the same band from SmP110 extract. The intermediate band also shows reduced staining while the low molecular weight band shows increased staining relative to the wild-type SmP110 strain. The overall pattern of activity of the three bands seems to be the inverse of what is normally seen in the wild-type strain.

4.4.5 SMb20847 Mutant

SMb20847, like SMb21291, is annotated in the Toulouse annotation database as a possible LysR transcriptional regulator but its actual role had never been investigated. It is located near the *bdhAxdhA2xdhB2* operon suggesting that a role in regulating some aspect of purine degradation was a possible. The SMb20847 strain Rm11510 shows growth similar to the SmP110 strain in M9 with glucose and ammonium chloride or on LB. In M9 with hypox-

anthine or xanthine Rm11510 shows growth like SmP110 with growth accompanied by a visible loss of purine crystals from the media. The situation is the same on solid agar media.

Staining of Rm11510 extract run on a non-denaturing polyacrylamide gel shows a band and staining pattern identical to that of the SmP110 strain. All three bands are present on the gel with a staining intensity identical to that of the wild-type.

4.4.6 Analysis of the Impact of the SMb21291 Mutation on *lacZ* Expression in Fusion Strains

In addition to creating a strain with a mutation in the SMb21291 putative LysR encoding gene for phenotypic analysis, the *lacZ* fusions were transduced into the SMb21291 mutant strain. This allowed for the study of changes in expression levels of possible purine degradation genes when the SMb21291 gene was mutated. Mutating SMb21291enabled the analysis of the possible role of the product of this gene in regulating purine degradation in *S. meliloti*.

xdhA1 Fusion with SMb21291 Mutation

The *xdhA1lacZ* fusion transduced from SmRL2522 into the SMb21291 mutant background resulted in a notable expression change in the *xdhA1* fusion (Fig. 4.14). When grown in M9 with hypoxanthine or xanthine there was a reduction in *lacZ* expression to about 1/3 the level in the non-LysR mutant background. The expression level of the transduced

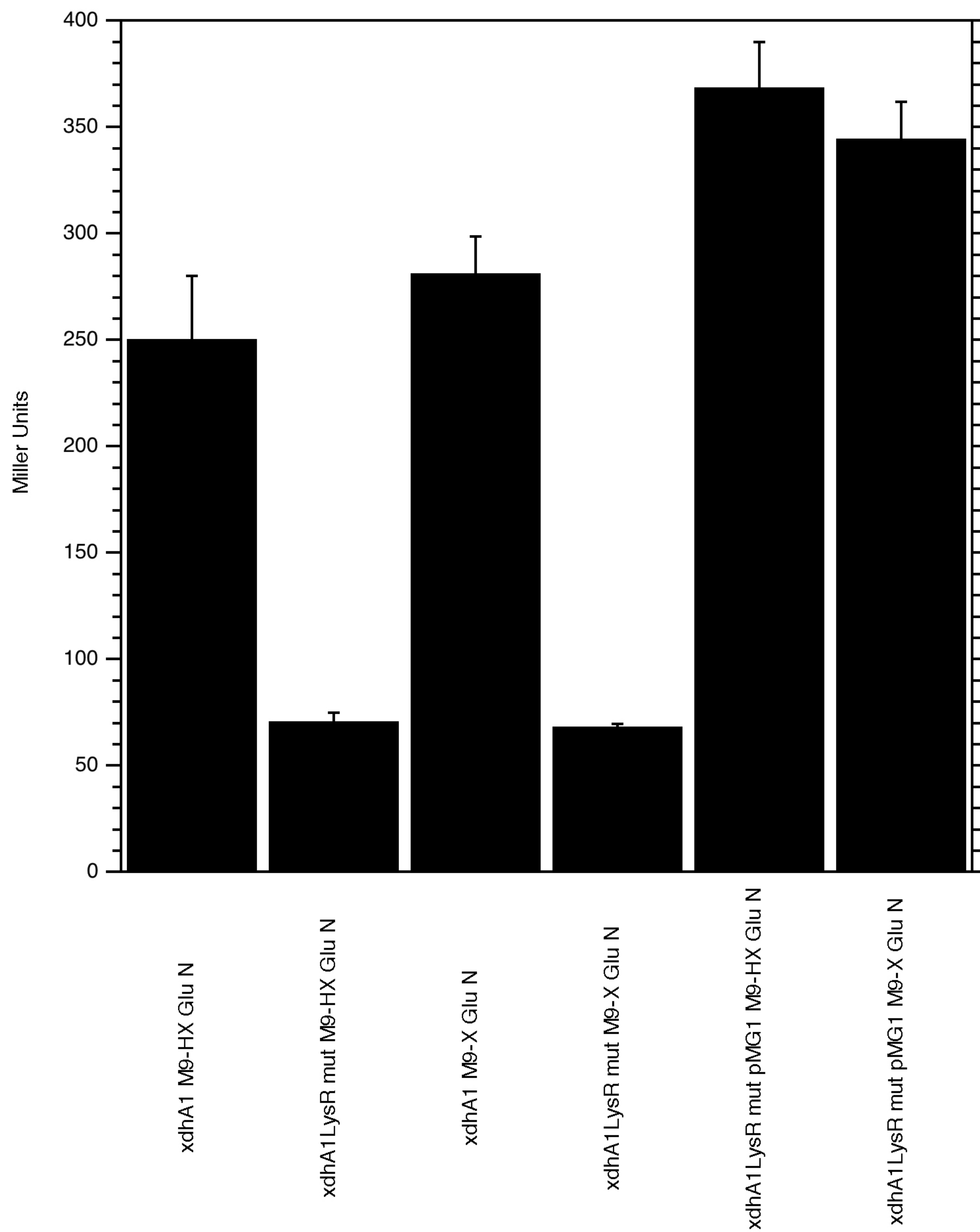

Figure 4.14: *lacZ* expression for *S. meliloti xdhA1lacZ* fusion strain transduced into the SMb21291 mutant grown in various M9 minimal media. xdhA1LysR = *xdhA1lacZ* in SMb21291 mutant strain. N=22.

strain remained significantly higher than of the SmP110 strain or the LysR mutant strain when grown in the same media. The transduced strains showed the same phenotype as the SMb21291 mutant with the new strain being unable to grow on hypoxanthine or xanthine as the only source of carbon and/or nitrogen. When grown in the same media supplemented with glucose and ammonium chloride growth of the transductant strain did not cause the disappearance of purine crystals from the media.

SMb21285 Fusion with SMb21291 Mutation

Transduction of the SMb21284-85 intergenic *lacZ* fusion (a fusion of the SMb21284-85 intergenic region plus the first 20 bp of SMb21285) into the SMb21291 mutant background gave results similar to those of the *xdhA1* fusion in the same genetic background (Fig. 4.15). There was a similar reduction in *lacZ* expression when the fusion was transduced into the LysR mutant strain to about 1/4 of normal activity found in the fusion background alone when grown in M9 containing hypoxanthine or xanthine. The resulting activity in the fusion/LysR mutant strain, though greatly reduced, remained significantly above that of the SmP110 and Rm11507 background strains in the same media.

guaD1 Fusion with SMb21291 Mutation

Upon transduction of the *guaD1lacZ* fusion into the SMb21921 mutant strain the levels of expression in M9 media containing guanine did not show a significant difference (Fig. 4.16) from those of the *guaD1lacZ* fusion strain without the SMb21291 mutation. The

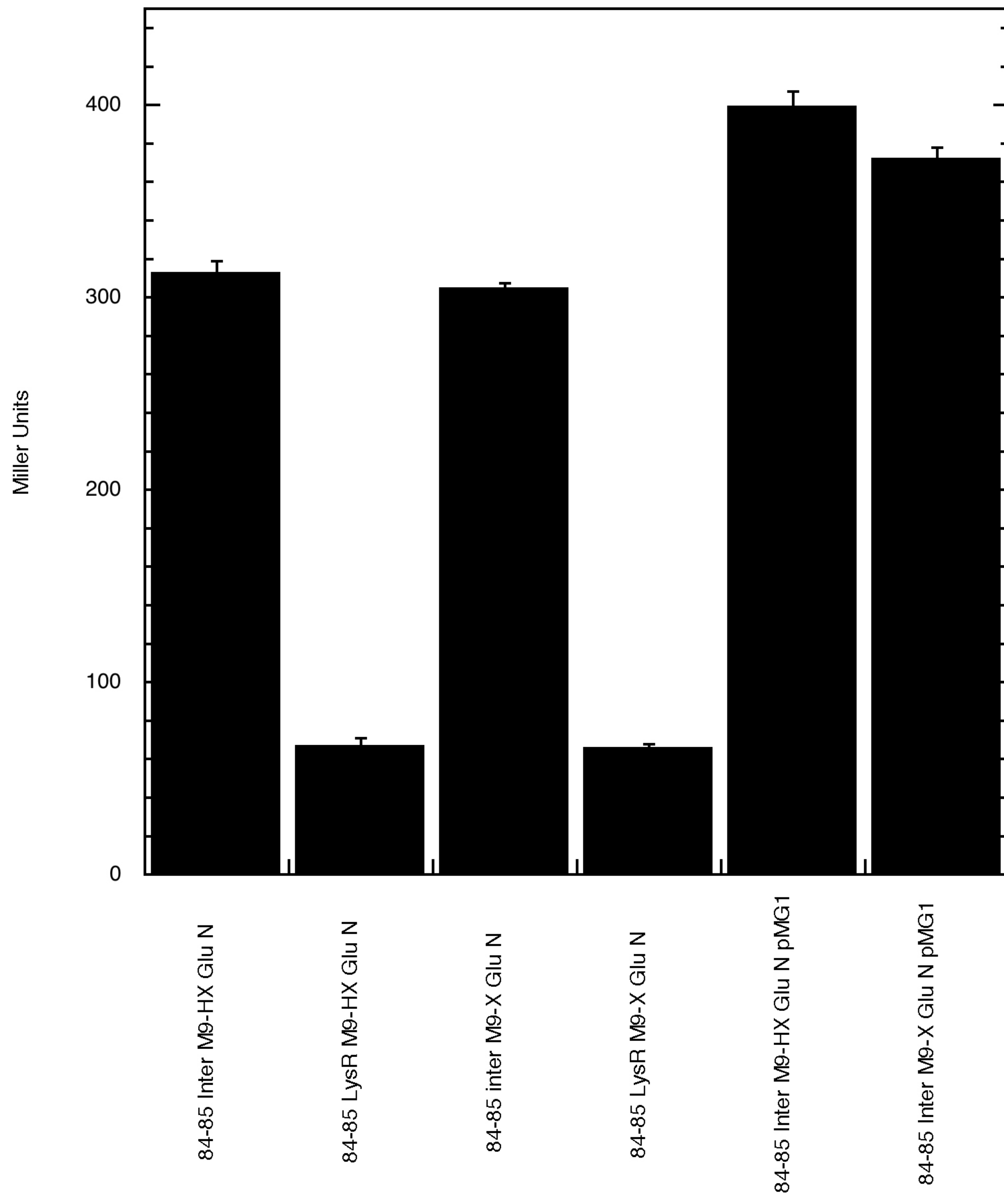

Figure 4.15: *lacZ* expression for *S. meliloti* SMb21285*lacZ* fusion strain transduced into the SMb21291 mutant grown in various M9 minimal media. 84-85 =SMb21285*lacZ* fusion strain in SMb21291 mutant. N = 22.

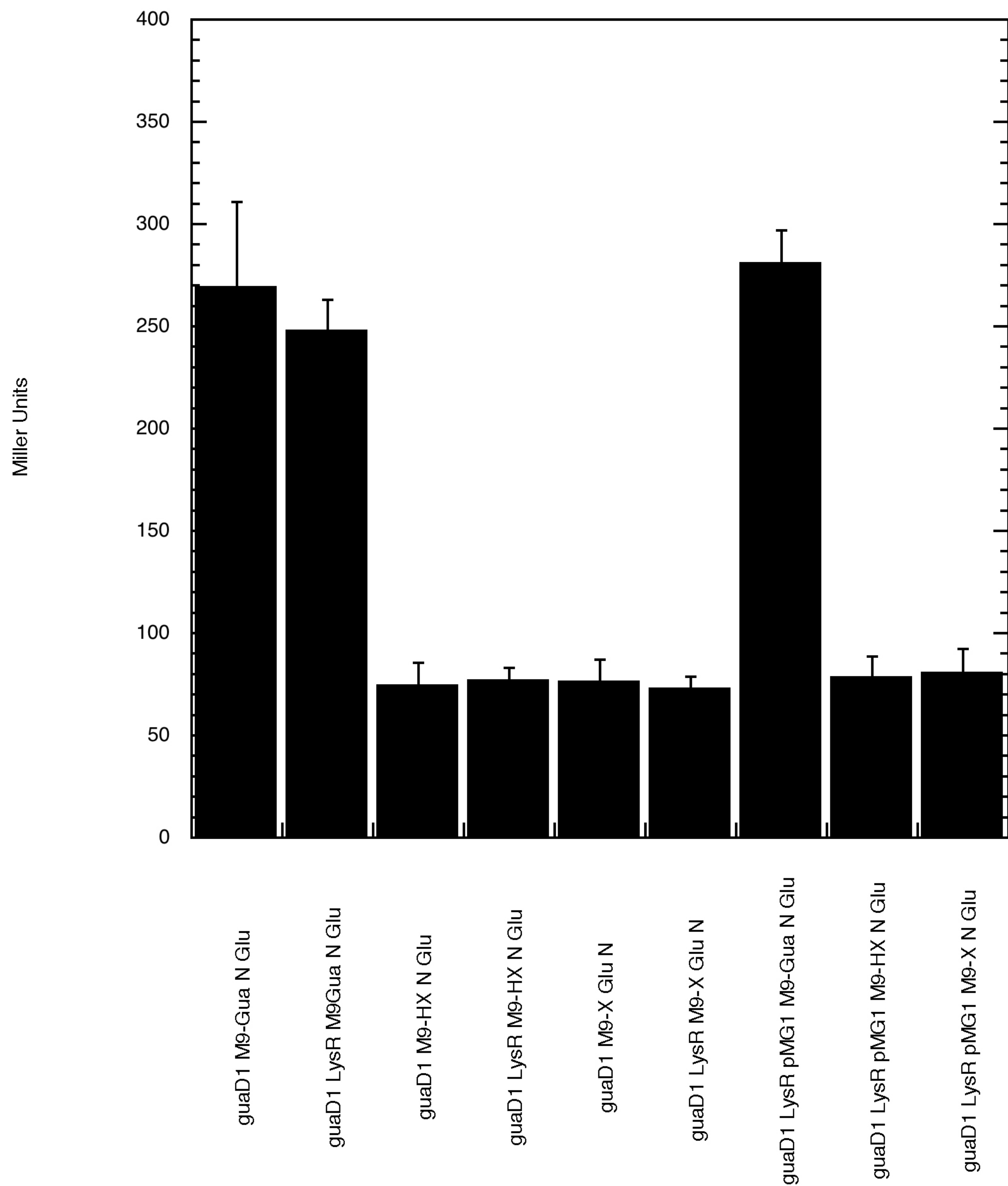

Figure 4.16: *lacZ* expression for *S. meliloti guaD1lacZ* fusion strain transduced into the SMb21291 mutant grown in M9 minimal media. Gua = guanine. guaD1 = *guaD1lacZ* fusion strain, guaD1LysR = *guaD1* fusion in SMb21291 mutant N = 22.

expression level of the transduced strain remained much higher than either SmP110 or Rm11507 background strains grown in the same media or with these strains grown in M9 media containing either hypoxanthine or xanthine. Transduction of the fusion into the SMb21291 mutant strain did not result in any change in the phenotype of LysR mutant strain and it remained unable to grow on hypoxanthine or xanthine as a sole source of nitrogen and/or carbon.

SMb21284 Fusion with SMb21291 Mutation

Transduction of the SMb21284 (a gene predicted to code for a uricase like protein) into the SMb21291 mutant background resulted in a reduction in *lacZ* expression when grown in M9 minimal media containing uric acid (Fig. 4.17). The levels of *lacZ* expression were reduced essentially to the background level found in the SmP110 or Rm11507 strains in the same media.

SMb21281 Fusion with SMb21291 Mutation

There was no detectable change in *lacZ* expression when the SMb21281 fusion was transduced into the LysR mutant strain. Expression levels remained similar to those of the background strains SmP110 and Rm11507 whether the strain was grown in the presence of hypoxanthine or xanthine or just with glucose and ammonium chloride.

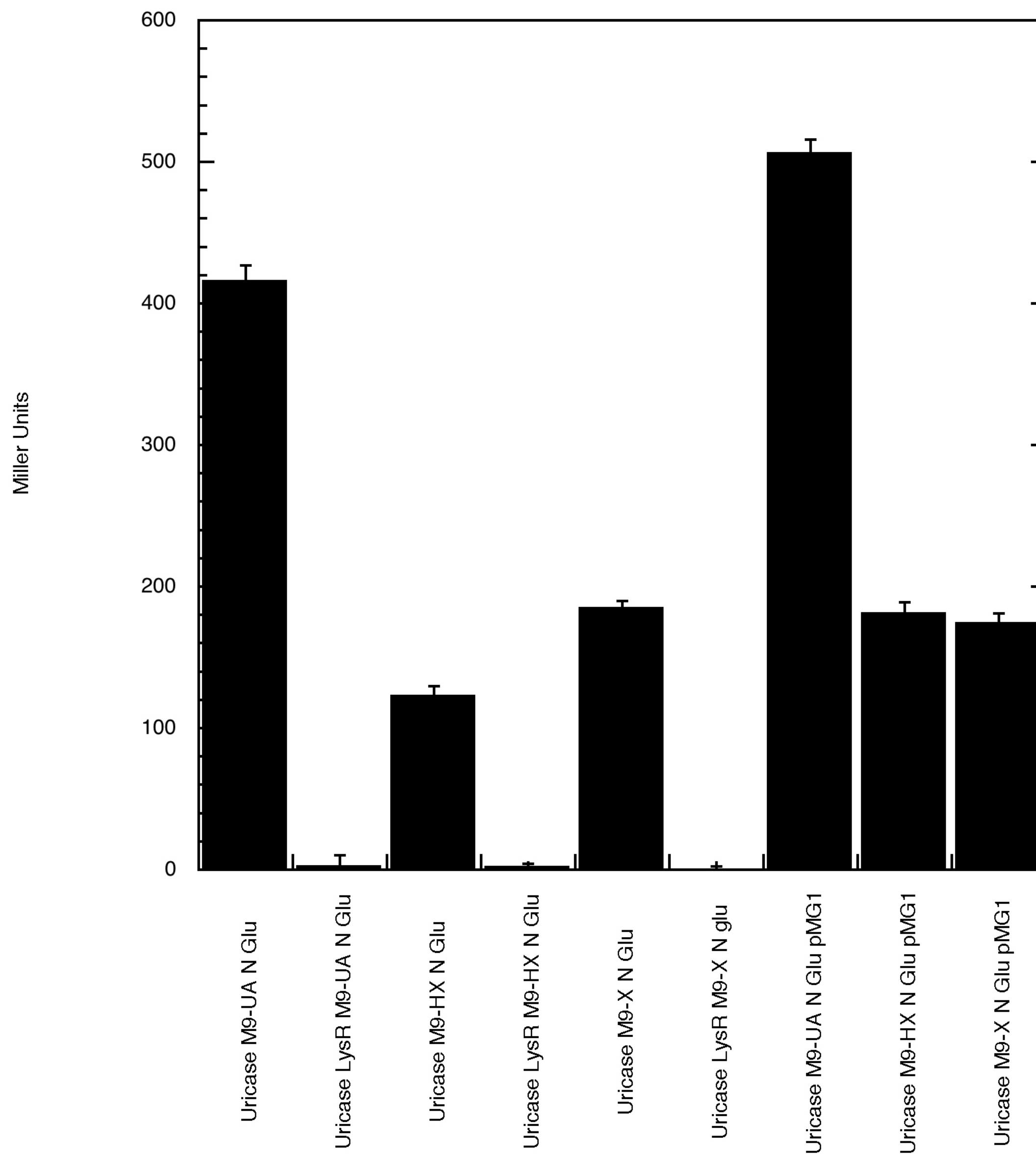

Figure 4.17: *lacZ* expression for SMb21284 *S. meliloti lacZ* fusion strain transduced into the SMb21291 mutant grown in various M9 minimal media. M9 = M9 minimal media, HX = hypoxanthine, X = xanthine, Glu = glucose, N = ammonium chloride. Uricase = SMb21284 *S. meliloti lacZ* fusion strain. N = 22.

guaD2 Fusion with SMb21291 Mutation

The transduction of the *guaD2* fusion into the Smb21291 mutant background did not result in any changes in *lacZ* expression levels (Fig. 4.18) compared with the *guaD2* fusion strain with the intact SMb21291. Expression levels and the patterns of expression were similar in both cases with the M9 media with guanine having the highest levels of expression. Once again, there was a reduced level of expression seen when hypoxanthine or xanthine was present in the media but the levels of expression were still significantly higher than was observed with the same strains grown in media lacking purines. In all cases the levels of expression were not significantly different between the *guaD2* fusion strain and the strain containing both the fusion and the mutation in the SMb21291 gene.

xdhA2 Fusion with SMb21291 Mutation

Transduction of the *xdhA2lacZ* fusion into the SMb21291 mutant strain resulted in noticeable changes in *lacZ* expression (Fig. 4.19) compared with the *xdhA2lacZ* fusion strain. Expression levels were higher in the transductant strain. The levels of expression for the strains containing the SMb21291 mutation was higher then the strain only containing the fusion. Expression in the transduced strains was similar with or without the presence of hypoxanthine or xanthine in the media. In each case expression was higher than with the SmP110 strain.

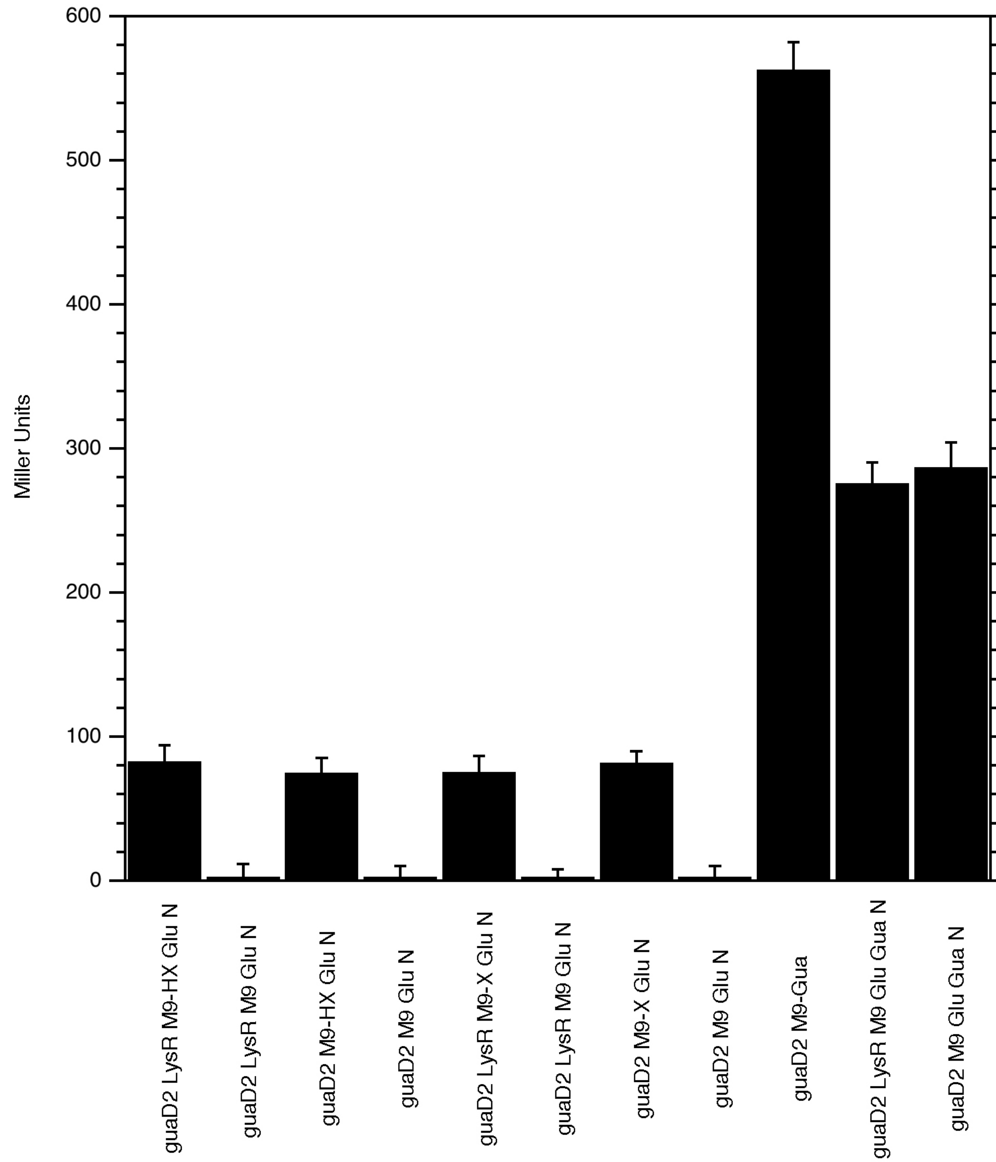

Figure 4.18: *lacZ* expression for *S. meliloti guaD2lacZ* fusion strain transduced into the SMb21291 mutant grown in various M9 minimal media. M9 = M9 minimal media, HX = hypoxanthine, X = xanthine, Gua = guanine,Glu = glucose, N = ammonium chloride. guaD2 = *guaD2lacZ* fusion. N = 22.

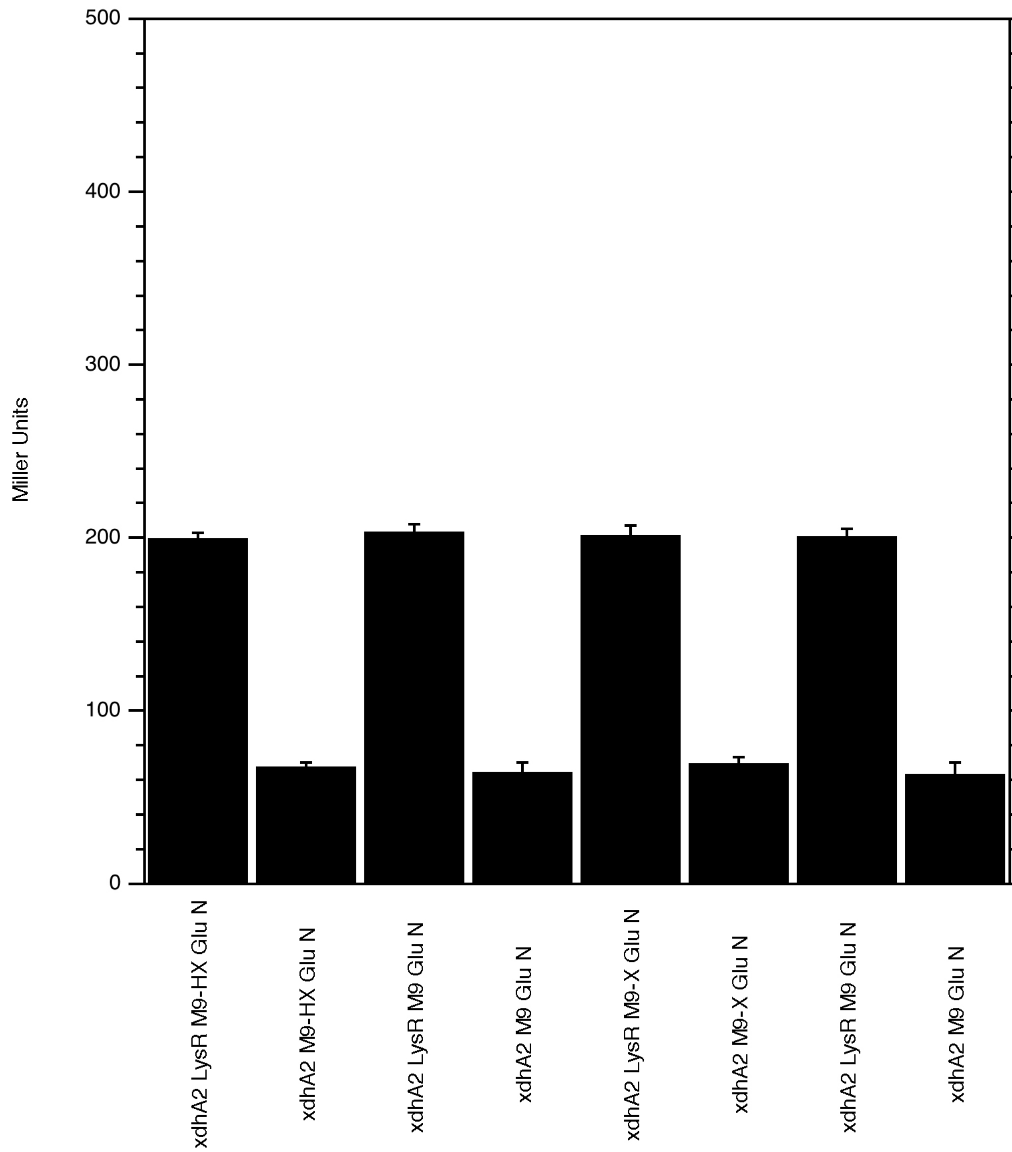

Figure 4.19: *lacZ* expression for *S. meliloti xdhA2lacZ* fusion strain transduced into the SMb21291 mutant grown in various M9 minimal media. M9 = M9 minimal media, HX = hypoxanthine, X = xanthine, Glu = glucose, N = ammonium chloride. xdhA2 = *xdhA2lacZ* fusion in SmP110. N = 22.

4.5 Growth and Expression of *lacZ* Fusion Strains on Different Carbon Sources

To determine if the *lacZ* fusion strains only showed elevated expression in response to purines the fusion strains were grown and tested for expression on a variety of different carbon sources (Table 4.1 and Table 4.2). These included different classes of compounds including carboxylic acids, pyrimidines, amino acids and hexoses among others. Of all the compounds tested only growth on purines (or at least the specific purine expected to be the substrate for the enzyme coded for by a specific gene) resulted in elevated levels of *lacZ* expression.

Table 4.1: Growth of *lacZ* fusion strains on M9 minimal media with various purines as sole carbon sources or with purines supplemented with glucose and ammonium chloride. [+++] represents growth equivalent to wild-type and [-] represents no growth. [Exp] represents fold increase in *lacZ* expression compared with wild-type.

Carbon and/or	SmRL2522		Rm11517		Rm11522		Rm11519		Rm11513	
nitrogen source	Growth	Exp.	Growth	Exp.	Growth	Exp.	Growth	Exp.	Growth	Exp.
Purines										
Gua	+++	5	-	0	+++	12	+++	1	+++	2.2
Gua, NH_4, Glu	+++	2	+++	2	+++	4.5	+++	1	+++	1.7
X	+++	19	-	0	+++	2	+++	1	+++	18
X, NH_4, Glu	+++	4	+++	3.8	+++	2.2	+++	1	+++	4
HX	+++	21	-	0	+++	2	+++	1	+++	18
HX, NH_4, Glu	+++	5	+++	4	+++	1.6	+++	1	+++	4
UA	+++	1	-	0	+++	1	+++	1	+++	1.1
UA, NH_4, Glu	+++	1	+++	8	+++	1	+++	1	+++	1
Ade	+++	3	-	0	+++	1.5	+++	1	+++	2.8
Ade, NH_4, Glu	+++	2	+++	2	+++	1.3	+++	1	+++	1.9

Table 4.2: Carbon sources on which *S. meliloti lacZ* fusions strains were grown and assayed for *lacZ* expression which showed no change in expression relative to SmP110.

Carboxylic acids	**Polyols**
lactate	D-mannitol
succinate	D-maltitol
mono-methyl succinate	glycerol
DL-malate	dulcitol
Amino acids	**Dipeptides**
L-alanine	Gly-Asp
L-asparagine	Ala-Gly
γ-aminobutyrate	Gly-Glu
aminovalerate	Gly-Pro
L-lysine	**Oligosaccharides**
L-glutamate	D-(+)-lactose
L-serine	D-melibiose
L-tyrosine	D-trehalose
L-ornithine	maltose
L-proline	D-raffinose
trans-4-hydroxy-L-proline	sucrose
Hexoses	**Pyrimidines**
D-galactose	cytosine
methyl-α-D-galactose	thymidine
methyl-β-D-galactose	uridine
N-acetyl D-galactosamine	**Pentoses**
N-acetyl D-glucosamine	D-ribose
D-fructose	D-xylose
D-glucose	L-arabinose
D-mannose	D-arabinose
D-tagatose	
6-deoxy-L-mannose	
L-fucose	

4.6 Discussion

The purine degradation pathway has, to date, not been extensively studied or defined in *S. meliloti*. The majority of the putative pathway for purine degradation has been based upon sequence analysis and comparison of sequences from the *S. meliloti* genome with sequences from other bacteria using standard sequence alignment tools such as BLAST. Even this approach has only been possible following the completion of the *S. meliloti* sequencing project (Barnett *et al.*, 2001; Capela *et al.*, 2001; Finan *et al.*, 2001; Galibert *et al.*, 2001). Thus virtually all genes believed to be associated with this pathway have only been given that designation based on non-experimental evidence and thus can only be described as "putative".

Previous work in our lab has begun to add some experimental weight to the analysis of *S. meliloti* purine degradation. This work has focused primarily on two areas. One being an investigation of the *bdhAxdhA2xdhB2* operon due to it containing the *bdhA* gene involved in the degradation of PHB which has been a key area of investigation in our research group (Aneja *et al.*, 2005; Aneja, 1999; Aneja and Charles, 1999; Charles *et al.*, 1997). This operon also contains one copy each of the *xdhA2* and *xdhB2* genes which code for two subunits of xanthine dehydrogenase, a critical enzyme in purine degradation. A second area of investigation previously undertaken in our lab has been to search for mutations that impair purine degradation by carrying out random Tn5 mutagenesis and testing for the inability of mutants to grow on purines such as hypoxanthine and xanthine as carbon and/or nitrogen sources (Capstick, 2004). This work identified three genes in three genomic regions that when mutated lead to a loss of ability to grow on these purines. These included a *ccmC* gene, a hypothetical membrane protein, SMb21292, and a hypothetical

protein SMb20684. Unexpectedly, only one of these genes is located in a genomic region containing any putative purine degradation genes (SMb21292) and none were found near the *bdhAxdhA2xdhB2* operon. It was also curious that none of the isolated mutants were found in genes believed to encode subunits of the xanthine oxidase/xanthine dehydrogenase protein nor were any isolated that were located in the genomic region containing the *bdhAxdhA2xdhB2* mixed function operon even though this area also contains genes such as *guaD2* which encodes a guanine deaminase and a predicted LysR regulator close to these genes which one might expect to play some role in purine degradation (SMb20487) (Capstick, 2004).

This work still left much of the purine degradation pathway in *S. meliloti* experimentally undefined or unconfirmed. In particular the region of the pSymB megaplasmid including the *xdhA1* and *xdhB1* genes coding for the additional and presumably higher molecular weight xanthine oxidase/xanthine dehydrogenase subunits. As annotated in the Toulouse database this region would seem to be critical to the purine degradation cycle in *S. meliloti* as it contains several genes that could be predicted to code for key enzymes in such a pathway (Wang *et al.*, 1999; Christiansen *et al.*, 1997; Burton, 1994; Nygaard, 1983; Vogels and van der Drift, 1976). These include genes that code for a second copy of guanine deaminase (*guaD1*), a predicted uricase like protein (SMb21284) which one would expect to play a role in the conversion of uric acid to allantoin and *xdhC*, predicted to be a chaperone which may have a crucial role in forming a functional xanthine oxidase/xanthine dehydrogenase protein. In addition, there is a gene encoding a predicted LysR type regulatory protein (SMb21291) located near the operon containing *xdhA1xdhB1* as well as the intriguing intergenic space between that operon and the operon containing the predicted uricase. These operons code for enzymes that one would expect to be utilized in sequential steps in a

purine degradation pathway (the conversion of hypoxanthine to xanthine followed by the conversion of xanthine to uric acid and uric acid to allantoin). Each of these operons is predicted to run in opposite orientation to one another and thus share a common intergenic region. There is also a gene predicted to encode a permease predicted to be involved in xanthine transport that has never been experimentally tested.

4.6.1 The *S. meliloti* RmF117 Strain

To begin examining the genomic region containing the genes described above including *xdhA1xdhB1xdhC* we began by examining the phenotypic impact on purine degradation when this entire genomic region has been deleted. To do this, we utilized the RmF117 strain which had previously been constructed and from which this region of the genome is missing (Charles and Finan, 1990). As seen in the results section, the loss of this region eliminates the ability of *S. meliloti* to grow on hypoxanthine or xanthine as a sole carbon and/or nitrogen source. This inability to grow on purines is observed even though the RmF117 strain still contains functional copies of genes coding for two subunits of xanthine oxidase / xanthine dehydrogenase, XdhA2 and XdhB2, as well as a functional guanine deaminase gene (*guaD2*). With functional copies of both genes required to produce the two subunits of the xanthine oxidase / xanthine dehydrogenase protein being present one might expect that a functional enzyme could be produced utilizing these two genes yet this does not appear to be the case. On the native protein gel, there is a complete absence of any detectable activity associated with this enzyme suggesting that there must be additional genes involved in the synthesis or activity of the functional protein. On the native gel not only are the bands associated with the enzymatic subunits coded for by *xdhA1* and

xdhB1 missing but those coded for by *xdhA2* and *xdhB2* are absent as well. This, despite these genes being unaffected by the genomic deletion. This result suggests that the required gene(s) is/are located on the portion of the pSymB megaplasmid missing from the RmF117 strain. It would also seem possible that other genes located in this region could be essential for purine degradation / utilization because they code for additional enzymes or regulators in this pathway without which the degradation of purines like hypoxanthine is blocked.

4.6.2 Analysis of *lacZ* Fusions to Putative Purine Degradation Genes

The *xdhA1lacZ* fusion strain clearly demonstrated the ability of hypoxanthine and xanthine to induce the gene coding for the XdhA1 subunit of the xanthine oxidase / xanthine dehydrogenase enzyme (Fig. 4.3). This result supports the sequence based classification of this gene. The fusion strain, in this case, leaves an intact copy of the *xdhA1* gene in place and is thus able to grow on hypoxanthine or xanthine just like the wild-type strain SmP110 while allowing for the measurement of activity of the promoter region via the measurement of β-galactosidase activity. It was also apparent that the level of *lacZ* expression was significantly higher in media that contains the purine than in media without either hypoxanthine or xanthine which strongly supports the idea of this gene being induced by these purines. The levels of expression indicate that hypoxanthine and xanthine induce the *xdhA1* gene to a similar extent. In addition the levels of *lacZ* expression in the SmP110 strain (the genetic background of the fusion strain) were very similar to those of the fusion strain grown in media lacking either hypoxanthine or xanthine. This suggests that the level of activity seen in the fusion strain grown in M9 with glucose and ammonium chloride was

the result of the background level of β-galactosidase found in *S. meliloti* and not due to activity from the fusion to *xdhA1*. It also seemed apparent that when provided with an alternative source of nitrogen and carbon (in this case glucose and ammonium chloride) the level of expression of the fusion was reduced significantly even if hypoxanthine or xanthine was also present in the media though it should be noted that the expression levels remained well above that of the strain grown in the absence of either purine. This situation was similar to that seen with *Bacillis subtilus* though in that case the repression of expression appears to be stronger as expression was almost eliminated (Christiansen *et al.*, 1997).

As *xdhA1* is predicted to share an operon with the genes annotated as *xdhB1* and *xdhC*, the induction of *xdhA1* will also result in the expression of these genes in the presence of hypoxanthine and xanthine. This idea was supported by the phenotypic impact of the *xdhA1* gene being mutated. In this case, the mutant lost its ability to grow on hypoxanthine or xanthine as a sole carbon or nitrogen source (Fig. 4.10). Examination of cell extract from the mutant reveals the source of this phenotype as all three bands that normally stain for xanthine oxidase / xanthine dehydrogenase activity were absent in the cell extract from the mutant indicating that there was a total (or virtually total) elimination of activity due to this enzyme. This result supports the prediction that *xdhA1* shares the same transcript with *xdhB1* and *xdhC* as the *xdhA1* mutation would appear to have a polar effect on these downstream genes. We will discuss *xdhC* in more detail shortly.

A *lacZ* fusion strain was also constructed to the SMb21285 gene. This gene is annotated as coding for a small hypothetical protein and is predicted to belong to the same operon as *xdhA1*, *xdhB1* and *xdhC* and could be expected to utilize the same promoter region for this operon (and hence for the genes coding for the XdhA1 and XdhB1 subunits of the en-

zyme). Examination of the *lacZ* expression data (Fig. 4.4) does in fact reveal very similar results to those seen with the *xdhA1* fusion. Once again, there was a relatively high level of induction when the strain was grown in minimal media with only hypoxanthine or xanthine as the sole source of carbon and nitrogen. As was found with the *xdhA1* fusion analysis, the levels of expression were far higher than those seen in the SmP110 genetic background strain when it was grown on the same purine containing media or when the fusion strain was grown in M9 minimal media lacking a purine. Again suggesting that the induction of the *lacZ* fusion was a response to the presence of hypoxanthine or xanthine in the media in which the fusion strain was grown. Likewise, the similar, relatively low level of expression seen in both the SmP110 strain and the fusion strain grown in M9 containing glucose and ammonium chloride but lacking a purine suggested that the level of induction observed was due to the presence of β-galactosidase genes known to be present in *S. meliloti* and not due to significant additional expression from the *lacZ* associated with the fusion itself. As was seen in the case of the *xdhA1* fusion, the levels of expression in M9 media containing either hypoxanthine or xanthine were similar. Also of note is that, as was the case with the *xdhA1* fusion, the level of expression was significantly reduced when the either purine containing media was supplemented with ammonium chloride and glucose suggesting that expression of the xanthine oxidase / xanthine dehydrogenase operon was repressed when alternative nitrogen or carbon sources were present in the media. It was also of note that the expression levels seen with the fusion to SMb21285 were very similar to those of the *xdhA1* fusion strain. It has been observed that fusions to genes in the same predicted operon seem to have very similar levels of expression while the expression levels of genes in different operons are quite variable.

Sharing the same intergenic region, but in the opposite orientation, with the operon

containing *xdhA1*, *xdhB1* and *xdhC* is the gene SMb21284. This gene is annotated as coding for a uricase like protein. Uricase is the enzyme that catalyzes the conversion of uric acid (itself the product of the conversion of xanthine catalyzed by Xdh) to allantoin (Nygaard, 1983; Vogels and van der Drift, 1976; Arima and Nose, 1968; Baum *et al.*, 1956). Thus this enzyme would be expected to be required for the purine degradation pathway which takes guanine or hypoxanthine (for example) and converts it eventually to ammonia to be used as a cellular nitrogen source. The phenotypic results would seem to support this hypothesis as the mutation caused by the insertion of pTH1703 resulted in the inability to grow on hypoxanthine, xanthine or uric acid. This somewhat complicates trying to study the *lacZ* expression in these strains as they must be grown using supplemental nitrogen and carbon sources such as ammonium chloride and glucose in addition to the expected substrate, uric acid. Under these conditions the level of *lacZ* expression was determined for the strain grown with the purines hypoxanthine or xanthine as well as uric acid supplemented with an additional carbon and nitrogen source. Under these conditions there was still a level of expression significantly higher than seen in M9 media containing only glucose and ammonium chloride (Fig. 4.6). In the case of the media containing hypoxanthine or xanthine, the level of expression was higher than the level seen in the SmP110 wild-type strain used as the genetic background for these strains though it was still a relatively low level compared with, for example, the *xdhA1* strain grown in the same type of media. This may suggest that xanthine or hypoxanthine is not a direct substrate for inducing the expression of SMb21284 but the induction may be the result of the conversion of hypoxanthine and xanthine to uric acid via the still functional Xdh enzyme in the strain. Looking at the level of expression observed when uric acid was added to the media we see a much higher expression level than under any of the other conditions suggesting that this compound was in fact causing the expression of this gene via a promoter located in the intergenic region. In all cases there did not appear to be any additional observed

induction when either hypoxanthine, xanthine or uric acid were in the media in which the wild-type background strain SmP110 was grown nor when the mutant strain was grown in M9 minimal media with only ammonium chloride and glucose. Again this suggests that the levels of *lacZ* expression only rise above the background levels when these compounds were available in the media.

The configuration of the operons that include the *xdhA1B1C* genes and the predicted uricase coding gene is rather interesting. These operons have an orientation opposite to each other (the *xdh* operon being coded on the plus strand of DNA while the operon containing the SMb21284, the "uricase like protein", is coded for on the negative DNA strand). Thus they share a common intergenic region. It is also interesting to note that these operons would be expected to code for successive steps in a predicted purine degradation pathway (Capstick, 2004; Nygaard, 1983; Vogels and van der Drift, 1976). This organization can also be viewed on Biocyc.org (http://biocyc.org). The operon containing *xdhA1B1C* would be expected to encode the enzyme (Xdh) that catalyzes the conversion of hypoxanthine to xanthine and xanthine to uric acid while the operon containing SMb21284 encodes enzymes involved in the following two steps in the pathway, the conversion of uric acid to allantoin and allantoin to urea. This arrangement of operons and the intergenic region offers the intriguing possibility that much of the putative purine degradation pathway could be under the control or influence of some common regulatory mechanism or master switch. Perhaps such a switch could sense the levels of some target molecule and thus act to regulate the expression of all four steps in the pathway via a single mechanism. This would provide a simple efficient system whereby a single or very few types of regulatory molecules could act to regulate the expression of two operons containing five genes and comprising four sequential steps of a biochemical pathway. It would seem log-

ical that such a regulatory molecule could interact with these operons through a shared region of DNA such as the common intergenic region. This shared region would make a good candidate as this small stretch of DNA (167 bp) would be expected to contain the regulatory / promoter regions of both operons thus suggesting the possibility that a common regulator could impact the expression of these two divergent operons concurrently.

The results observed when the *lacZ* fusions to the common promoter region of SMb21284 (the uricase-like protein) and SMb21285 (the first gene of the operon containing *xdhA1B1C*) were transduced into the SMb21291 (putative LysR transcriptional regulator) mutant background suggests a common mechanism of regulation for these operons (Fig. 4.17 and Fig. 4.15). In each case, the expression from the promoter region for the operons in both orientations was severely reduced compared with expression of the same fusions in the genetic background containing a functional version of SMb21291. Complementation with a plasmid containing a complete copy of SMb21291 in both cases restored the original patterns of expression in both strains containing a mutated SMb21291 again strongly supporting the idea that a single switch is involved in regulating both operons. The same situation was seen for the *lacZ* fusion to *xdhA1* where the initial high levels of expression seen when the fusion strain was grown in minimal media containing hypoxanthine or xanthine were sharply reduced when the fusion was transduced into the SMb21291 mutant strain. This naturally raises the question of how the single regulatory molecule (such as one coded for by SMb21291) could sense different inducer molecules that could cause it to bind to the intergenic region and promote expression of each operon at the same time. The answer may lie in the chemical structure of the presumed chemical inducers. Hypoxanthine, xanthine and uric acid are all heterocyclic aromatic organic compounds, that share virtually the same chemical structure (Voet and Voet, 1995a). This consists of a pyrimidine ring fused

to an imidazole ring. The only difference is in the number of oxygen molecules attached to the rings. The number of attached oxygen molecules increases from one, in the case of hypoxanthine to two with xanthine to three with uric acid. Otherwise, the structures of the molecules are identical. It seems plausible that a single transcriptional regulator could therefore sense all three of these near identical molecules. There is some evidence for this in the case of Xdh in *Rhodobacter* where *lacZ* expression of fusions to Xdh encoding genes was demonstrated to occur for both hypoxanthine and xanthine even though these molecules differ by one oxygen in their chemical composition (Leimkuhler *et al.*, 1998). In that case the levels of expression were found to be higher for xanthine suggesting that while the molecule that senses these purines is able to bind both purines it may have a stronger affinity for xanthine than hypoxanthine or that the conformational change induced by xanthine is more conducive to increasing levels of expression from the promoter region of the operon. Again however the key point being that expression is sensitive to both of these very similar molecules. The authors of this paper do not, however, suggest by what mechanism the levels of *xdh* gene expression may be controlled. The results obtained in the case of *S. meliloti* suggest a similar situation with the levels of expression of the operon encoding *xdhA1B1C* being elevated in response to both hypoxanthine and xanthine. In addition, we see that the levels of expression to the SMb21284, uricase encoding gene, also became elevated in the media containing uric acid. Expression of fusions to both operons were greatly reduced when the SMb21291 gene was mutated suggesting that it is this putative transcriptional regulator that is acting to stimulate the expression of both operons perhaps assuming an active conformation upon interacting with the hypoxanthine, xanthine or uric acid.

The *guaD1 lacZ* fusion strain was able to grow normally on M9 minimal media with

either hypoxanthine or xanthine as a sole source of carbon and nitrogen despite the *guaD1* gene being mutated. This is possible because there is a second guanine deaminase, *guaD2*, located in the *S. meliloti* genome near the *bdhAxdhA2xdhB2* operon (http://iant.toulouse.inra .fr/bacteria/annotation.cgi). When grown on guanine as a sole carbon and nitrogen source growth of the *guaD1* fusion was slow, however this is the case for the wild-type SmP110 strain as well indicating that guanine is not an optimal carbon and nitrogen source for *S. meliloti*. It seems clear that guanine is the preferred inducing substrate for this gene as high levels of induction were only seen when there was guanine in the M9 media. The level of induction was much higher (Fig. 4.5) when guanine was the single source of carbon and nitrogen as was seen with hypoxanthine and xanthine in the case of *xdhA1*. Induction dropped noticeably when an alternative source of carbon and nitrogen was provided in addition to guanine. In this case, the level of induction was reduced despite the better cell growth under these conditions. This was similar to the situation observed with *xdhA1*, for example, where the presence of another nitrogen source, presumably one preferable or at least easier for the cell to utilize, resulted in a reduced level of expression of the purine degradation cycle gene. Although the level of induction was indeed reduced, it still remained higher than was seen when the *guaD1* fusion strain was grown in the control media lacking guanine or when the wild-type SmP110 strain was grown in the same media. When grown in M9 media containing either hypoxanthine or xanthine there was still an elevated level of induction observed compared with the control strains grown in media without either of these purines. This seems to suggest that some of the xanthine (or hypoxanthine converted to xanthine by xanthine dehydrogenase) was being converted to guanine and thus the cells were using at least some of the provided purine substrate to produce this purine base. This does not seem to be too surprising as the reaction catalyzed by guanine deaminase is believed to be reversible (see http://BioCyc.org).

In the *S. meliloti* genome there are 198 genes annotated as possible permeases according to the Toulouse genome annotation database. Few of these have, however, been experimentally tested and the annotations are based on sequence similarity comparisons with other bacterial species such as *Pseudomonas*. The strain Rm11519 which contains a *lacZ* fusion to SMb21281 annotated as coding for a putative membrane transport permease involved in the transport of xanthine was tested for *lacZ* activity. When grown in M9 minimal media containing hypoxanthine, xanthine, glucose and ammonium chloride as well as hypoxanthine and xanthine with added glucose and hypoxanthine however there was no obvious increase in the levels of *lacZ* activity above that observed for the controls (Fig. 4.7). The levels of expression were not significantly different in M9 with or without added hypoxanthine or xanthine. The expression levels for the fusion strain grown in media containing a purine remained about the same as the SmP110 background strain and the fusion strain grown in M9 media containing either purine. Despite the annotation of this gene there was no indication that it is expressed (or showed elevated expression) in the presence of xanthine or hypoxanthine. As an additional check, the *gusA* fusion to this gene was also tested for induction under the same growth conditions with similar results. This suggests that caution is necessary in assigning functionality to genes based only on sequence similarity. Ultimately, there is no substitute for experimental evidence. The Rm11519 strain was able to grow on minimal media even with the permease gene mutated thus the gene did not seem to be required for growth on hypoxanthine or xanthine as a single source of carbon and nitrogen. The function of the gene (or at least the substrate it acts upon) remains unclear at this point though it does not seem to be associated with the purine degradation cycle based on these results. Previous work by Mauchline *et al.* did however identify elevated levels of expression in transporters in *S. meliloti* in response to the purines xanthine (SMb20127) and allantoin (SMc02415) (Mauchline *et al.*, 2006). Neither of these genes are located in close proximity to the genes investigated here.

4.6.3 Analysis of *xdhC*

The *xdhC* gene is the last gene found in the operon that also contains the *xdhA1* and *xdhB1* genes and shares an intergenic region with the operon containing SMb21284 which codes for a uricase-like protein. The *xdhC* gene is unique to this operon as there is no equivalent gene found in the operon containing *xdhA2* and *xdhB2*, which is found elsewhere in the genome and codes for the two lower molecular weight xanthine oxidase/xanthine dehydrogenase subunits (Toulouse annotation database). In the *S. meliloti* Toulouse database, *xdhC* is described as a 288 amino acid protein of approximately 31 kDa and annotated as a "probable xanthine dehydrogenase". In the phototrophic purple bacterium *Rhodobacter capsulatus*, there is a similar three gene configuration of xanthine dehydrogenase into two subunits coded for by *xdhA1* and *xdhB1* genes with a third downstream gene co-transcribed with *xdhA1* and *xdhB1* also called *xdhC* (Leimkuhler *et al.*, 1998). In *R. capsulatus* the XdhA subunit contains two iron-sulfer binding domains, and a FAD cofactor binding domain. In addition, there is a molybdenum cofactor located within the XdhB subunit. In *R. capsulatus* the XdhC is described as a protein that is not a subunit of Xdh but is required for the insertion of the molybdenum cofactor into XdhB or as a chaperone required for the correct folding of the Xdh protein during the process of molybdenum cofactor insertion. In this case, it was demonstrated that an *xdhC* mutant strain of *R. capsulatus* lacked the molybdenum cofactor. Analysis of the crystal structures of several molybdoenzymes demonstrated that the molybdenum cofactor is found deep within the enzyme at the base of a channel whose size and shape match that of the enzymatic substrate (Kisker *et al.*, 1997; Romão *et al.*, 1997). The position of the

molybdenum cofactor so deep within the enzyme would suggest that the cofactor cannot be inserted after the protein has folded into its final configuration. It is suggested that the protein (as a tetramer) would maintain an open configuration until the molybdenum cofactor has been inserted with the assistance of XdhC. Such assistance could take the form of a carrier protein for the molybdenum cofactor itself or as a chaperone that helps to fold the Xdh protein as the molybdenum cofactor is inserted once the attachment of the cofactor is complete. Another example of a molybdenum cofactor containing enzyme is the nitrate reductase A of *E. coli*. This enzyme is composed of three subunits, NarG, NarH and NarI the latter containing the molybdenum co-factor. In addition, there is an anotherl gene involved, *narJ* that is believed to encode a chaperone that is involved in inserting the molybdenum co-factor into the NarG subunit. Yet another example of this type of molybdenum co-factor containing enzyme requiring the action of a chaperone (TorD) to form the final active enzyme is found in the case of trimethylamine *N*-oxide reductase in *E. coli* (Pommier *et al.*, 1998; Liu and DeMoss, 1997; Blasco *et al.*, 1992, 1989; Sodergren and DeMoss, 1988).

In the case of *xdhC*, the mutant was created using the suicide vector pK19*mob* containing a fragment of *xdhC* that was transferred into *S. meliloti* via tri-parental mating. Upon entry into the cell, the plasmid underwent recombination into the *S. meliloti* genome resulting in a truncated (mutated) copy of the *xdhC* gene. Using the newly created strain to inoculate M9 minimal media, containing either hypoxanthine or xanthine as the carbon and nitrogen source, demonstrated a phenotype whereby the mutant strain was unable to grow (Fig. 4.11) in either of these media while the wild-type strain SmP110 showed obvious growth. The lack of growth of the mutant was accompanied by no obvious reduction in the presence of hypoxanthine or xanthine crystals in solution unlike the wild type where

the hypoxanthine or xanthine crystals could be observed to vanish as the culture grew indicating that the wild-type strain was using these crystals as a nutrient source. What was interesting about these observations was the fact that in creating the mutant, only the *xdhC* gene was targeted for mutation while the remaining upstream genes as well as the *xdhA2* and *xdhB2* genes found in the *bdhAxdhA2xdhB2* operon were untouched. This indicated that the XdhC is required for the production of active Xdh regardless of which operon is actually coding for the XdhA and XdhB subunits.

Assaying cell extract from SmP110 and the *xdhC* mutant strains supported this conclusion. Examination of the non-denaturing polyacrylamide gel on which cell extract had been run and stained for Xdh activity revealed that cell extract from the *xdhC* mutant strain lacked any detectable Xdh activity. All bands normally associated with the presence of active Xdh were absent. The wild type strain showed bands of three sizes corresponding to the high molecular weight enzyme formed by the combination of XdhA1 and XdhB1, the intermediate sized band made up of XdhA1/XdhB2 and XdhA2/XdhB1 subunits and the lower molecular weight band composed of XdhA2 and XdhB2 subunits. It was clear that the assay was working correctly as the small spot corresponding to the activity of an unknown oxidase not associated with the either Xdh coding operon is present both in the assay using wild-type cell extract as well as that of the mutant strain. This band acts as a useful internal control for the success of the assay. This again supports the results seen for the growth assay. The complete lack of Xdh activity explains the inability of the mutant strain to grow on this media and again supports the idea that the XdhC protein (i.e. a functional copy of the *xdhC* gene) is required for the formation of active Xdh from the XdhA and XdhB subunits from either operon. Complementation of the *xdhC* mutant restores the ability to grow on hypoxanthine and xanthine thus confirming the requirement

of a functional copy of this gene being necessary to form an active Xdh.

4.6.4 The Impact of Mutating SMb21291, a Putative LysR Transcriptional Regulator

Mutating the SMb21291 gene in the SmP110 strain resulted in a distinct phenotype including the inability to grow on purines such as hypoxanthine or xanthine as the only source of carbon and nitrogen (Fig. 4.12) along with an alteration in the activity pattern observed on the native gel assay (Fig. 4.13). Mutating the SMb21291 gene created an obvious phenotypic impact upon expression in the *xdhA1* fusion strain. Transduction of the *lacZ* fusion into the SMb21291 mutant genetic background resulted in the levels of β-galactosidase activity dropping significantly from those seen in the fusion strain alone when grown in the same media (Fig. 4.14). This result was consistent with that observed on the native protein gel. On the gel, it was quite apparent that the level of staining was greatly reduced for the band known to contain the functional protein composed of XdhA1 and XdhB1 subunits (Aneja and Charles, 1999). This is the band with the highest molecular weight and thus appears as the highest of the three bands normally found on gels using this assay. It was interesting to note that while there was a large reduction in staining of this band it did show that there was still some activity detectable which was consistent with the levels of expression seen in the *lacZ* assays where the levels of expression were greatly reduced but still above the background level observed for the fusion strain lacking the SMb21291 mutation grown in media containing hypoxanthine or xanthine or the SmP110 wild-type strain grown in the same media. Consistent with this observation was the large reduction in activity associated with the intermediate band. This band is

composed of a combination of XdhA1, XdhB2, XdhA2 and XdhB1 hence the position of this band between the other two on the gel. Since there were fewer XdhA1 and XdhB1 subunits available to combine with XdhA2 and XdhB2 subunits the reduction in the activity associated with the intermediately sized protein is a logical observation. There was simply less of this particular variant of the enzyme being produced than in the wild-type. The combination of the *lacZ* expression results and the native gel results for the SMb21291 mutant strain Rm11507 (Fig.4.13) would seem to indicate that the protein coded for by SMb21291 is not acting as a simple on/off switch for expression of *xdhA1*. It would be expected that all expression would be lost if this were the case due to not only the complete loss of expression of *xdhA1* and *xdhB1* but of *xdhC* as it is part of the same operon and encoded on the same transcript. This clearly is not what is happening. It seems the observation of a reduced expression level in *xdhA1* is more consistent with the idea that the protein coded for by SMb21291 acts as a positive regulator or inducer of expression of the operon containing these genes and in the absence of this protein there is an attenuation of expression of *xdhA1* to a lower level. This is again supported by the complementation data where the strain combining the SMb21291 mutation with a *lacZ* fusion to *xdhA1* showed a restoration of a higher level of β-galactosidase activity upon the introduction of the pMG1 plasmid (cosmid) containing an intact copy of SMb21291. This result further reinforces the involvement of this gene in regulating the level of expression of *xdhA1* i.e. it increased expression to levels resembling the fusion strain in which the SMb21291 gene was not mutated.

Transferring the *lacZ* fusion to SMb21285 (a protein of unknown function) into the SMb21291 mutant background produced results (Fig. 4.15) very similar to those found when the fusion to *xdhA1* was transduced into the same mutant. Since both of these genes

are predicted to be found in the same operon (see http://biocyc.org) it would be expected that this upstream region would be common for both of these genes as well as other genes in the same operon (*xdhC* for example). It would also be expected that this intergenic region would probably be the shared promoter containing region for the common transcript generated by this operon. Indeed the expression results in both cases are quite similar as would be expected if both these fusions shared a common promoter. In both cases, the level of expression in M9 minimal media containing hypoxanthine or xanthine were similar and there was clearly a significantly elevated level of *lacZ* expression when the strain was grown in M9 minimal media containing one of these purines. In the SMb21291 mutant background expression from the fusion was significantly reduced as was the case for the *xdhA1* fusion in the same background. It seems that providing an alternate source of nitrogen and carbon resulted in reduced expression though the levels remained higher than in the background strain for either the fusion strain or the SMb21291 mutant (Rm11507) when grown in the same media. This seems to indicate that the SMb21291 mutation resulted in an attenuation of expression from the promoter region so that the expression of genes that utilize this promoter were substantially reduced but not completely switched off. As was the case with the *xdhA1*, the data for the fusion to SMb21285 suggested that the protein coded for by SMb21291 acts as positive regulator of expression in this case. The similar levels of expression in the M9 minimal media containing hypoxanthine or xanthine and the similar magnitude of reduced expression in the SMb21291 mutant background suggested that the promoter of each gene has virtually identical characteristics. This, combined with the organization of these genes and their predicted placement in a common operon all imply that in both cases it is the same promoter region that is driving the expression of these genes. As was the case with the *xdhA1* fusion the transfer of the fusion into the SMb21291 mutant strain did not result in a change in the phenotype of the mutant as the mutant still wouldn't grow on M9 minimal media where the only

source of carbon or nitrogen is xanthine or hypoxanthine so expression in this media could not be tested. Addition of the pMG1 plasmid containing an intact copy of the SMb21291 gene restored *lacZ* expression to levels similar to those found in the fusion strain that did not contain a mutation in SMb21291. Again this supported the idea that it is the product of this gene that is responsible for the observed changes in expression.

The putative uricase-like protein coded for by the SMb21284 gene shares an intergenic region with the operon containing the *xdhA1*, *xdhB1* and *xdhC* genes (Toulouse annotation database, http://iant.toulouse.inra.fr/bacteria/annotation/cgi/rhime.cgi). Unlike this operon, SMb21284 runs in the opposite orientation and is encoded on the opposite strand of DNA. The predicted substrate for this gene based on the annotation would be uric acid (Nygaard, 1983; Vogels and van der Drift, 1976) and indeed when grown on M9 minimal media containing uric acid there was a significantly elevated level of expression as was previously discussed. Introduction of the SMb21284*lacZ* fusion into the SMb21291 mutant strain resulted in a sharp drop in *lacZ* expression when grown in M9 minimal media containing either uric acid, hypoxanthine or xanthine (Fig. 4.17). In this case, the drop in expression was quite dramatic as the expression levels were reduced essentially to background levels suggesting that the expression from the promoter region was turned off or at least very sharply attenuated. From this data, it would seem that the small intergenic region (167 bp) shared by these genes may play a role in regulating the expression of both of these coding regions via the protein coded for by SMb21291 as expression of both operons is impacted by this mutation. It is also interesting to note that the operons containing the genes coding for the smaller Xdh subunits as well as the required XdhC protein are located next to the genes encoding enzymes predicted to encode the enzymes involved in the following steps in the putative purine degradation pathway (i.e the conver-

sion of uric acid to allantoin). Not only are these operons located next to each other in the *S. meliloti* genome, they also share a common intergenic region suggesting that they may share some common regulatory mechanism of which the SMb21291 protein would be a plausible candidate given the observed changes in expression. Complementation with the pMG1 cosmid containing an intact copy of SMb21291 again, as was seen for the *xdhA1* fusion/SMb21291 mutant, restored the expression to a level similar to the fusion strain in which the SMb21291 gene was not mutated.

The gene annotated as *guaD1* exists as the second gene in an operon that also includes the gene SMb21292 predicted to encode a hypothetical protein (Toulouse annotation database and Biocyc.org at http://biocyc.org). Mutation of SMb21292 had been previously been shown to produce a phenotype in which the mutant *S. meliloti* strain was unable to utilize hypoxanthine as a sole source of carbon and nitrogen. This mutant also showed an alteration in the activity of the Xdh protein detected on the native gel assay in which the activity associated with the Xdh composed of the XdhA1 and XdhB1 subunits was elevated (Capstick, 2004). It would be expected that the mutation in SMb21292 would also be polar on *guaD1* as they are predicted to be part of the same transcriptional unit. This operon is also located close to SMb21291, the putative LysR transcriptional regulator. The *guaD1* gene is predicted to code for a guanine deaminase that catalyzes the conversion of the purine base guanine into xanthine which can then be further degraded to ammonia via the putative purine degradation pathway. The conversion of guanine to xanthine is predicted to be a reversible reaction so it should be, in theory, possible for the cell to produce additional guanine from xanthine as part of this pathway as well. Examining the results of the *lacZ* expression assays, it is clear that the presence of guanine in the media resulted in significantly increased expression though again this level of expression was diminished if an

alternative source of nitrogen and carbon was provided. The lower level of expression was still significantly higher than the background level observed in the SmP110 strain grown in the same media or in the background strain grown in media lacking guanine. Upon transducing the *lacZ* fusion into the SMb21291 mutant strain and examining expression in guanine containing M9 minimal media (Fig. 4.16), little difference between the two strains was observed. This suggests that the SMb21291 protein does not play a role in regulating the expression of *guaD1*. The same situation was seen in M9 minimal media containing hypoxanthine or xanthine. While the presence of these purines in the media used to grow the *guaD1* fusion strain resulted in an elevated level of expression there was no change in expression observed when the fusion was transduced into the SMb21291 mutant and grown under the same conditions. This again suggests that the expression of *guaD1* is not under the regulation of the protein coded for by SMb21292. Again in all cases the mutation in SMb21291 did not result in any change in the background level of expression in the SmP110 strain as the level of expression in this strain was similar to the mutant strain (Rm11507) when grown in the presence of purines. This demonstrated that the SMb21291 gene was not playing any role in expression of the *lacZ* genes found in the wild-type strain. The elevated level of expression seen for the *guaD1* fusion grown in media containing hypoxanthine or xanthine is interesting. There was an elevated level of expression in the presence of xanthine suggesting that when grown on these purines there was some induction of *guaD1* indicating that the cell uses some of the available hypoxanthine or xanthine to synthesize guanine bases although the level of expression on this substrate is much lower than it is for guanine under these growth conditions. The addition of pMG1, containing an intact copy of the SMb21291 gene produced similar levels of expression as those seen in both the *guaD1* fusion strain and the strain in which the fusion was transduced into the the SMb21291 mutant strain. Again this result was as expected if the SMb21291 gene product has no impact on expression of *guaD1*. The location of the *guaD1* gene in the same

predicted operon as the SMb21292 gene raises the possibility that it is the SMb21292 gene product that may have the direct impact upon the SMb21284-85 intergenic region. If this were the case then the impact of the SMb21291 mutation would be an indirect effect upon the expression of *xdhA1B1C* and the uricase like protein. In this scenario, the SMb21291 encoded protein would act as the regulator of SMb21292 which would act directly on the SMb21284-85 intergenic region. However the fact that the mutation of SMb21291 does not show any impact upon the expression of *guaD1* which is in the same operon as SMb21292 does not support this model. It would be expected that SMb21292 and *guaD1* would be transcribed together if they are on the same transcript. This does not however exclude the possibility that the SMb21292 protein plays some role in allowing the SMb21291 protein to act as a transcriptional regulator. Such a scenario could possibly explain the altered pattern of Xdh activity and hypoxanthine utilization phenotype seen when SMb21292 is mutated.

The expression results for the *guaD2* fusion strain (Fig. 4.9) were similar to those seen for the *guaD1* fusion. It seems clear from the results that guanine was the strongest inducer of expression from the promoter of each of these genes as expected. As with *guaD1* the presence of xanthine or hypoxanthine in the media did result in a higher level of expression from the promoter for reasons, one would expect, similar to that seen with the *guaD1* fusion. Unlike *guaD1*, *guaD2* is not predicted to be part of an operon with another gene.

The *lacZ* fusion to the promoter of the gene annotated as a permease showed no detectable change in expression when grown in M9 minimal media containing hypoxanthine or xanthine (Fig. 4.7). The mutation of this gene did not impact the growth of the mutant strain on hypoxanthine or xanthine compared with the wild-type strain. This result calls

into question the annotation of this gene as a permease involved in transporting xanthine. The transduction of the *lacZ* fusion into the SMb21291 mutant strain showed no detectable change in *lacZ* expression resulting from the mutation in the transcriptional regulator. Expression levels in the transduced strain were similar to those of the *lacZ* fusion strain when grown in the same media (Fig. 4.7) indicating that this gene is not induced by either of these purines. This does not mean that the gene does not code for a permease only that there is no experimental evidence that it is involved in the transport of these particular purines or that its expression increases when these purines are used as the only carbon and nitrogen sources.

The expression results for the *xdhA2* fusion strain didn't show any change in expression relative to SmP110 when grown in media containing hypoxanthine or xanthine (Fig. 4.8). This initial result may seem unexpected given the levels of increased induction observed for the fusion to the *xdhA1* gene. However the operon containing *xdhA2* has previously been demonstrated to be inducible by biotin and contains an upstream promoter region different from that of the *xdhA1B1C* containing operon as well as the first gene in the operon being *bdhA*, a gene involved in the PHB cycle. The region upstream of this operon containing *xdhA2* had previously been shown to contain a promoter region very similar to the promoter of *bioS*, a gene in *S. meliloti* shown to be biotin inducible (Hofmann *et al.*, 2000a). Given this it does not seem surprising that this operon is regulated in a manner different from the operon containing *xdhA1B1C*. This idea was further supported by the results observed when the *xdhA2lacZ* fusion was transduced into the SMb21291 mutant. Unlike with the *xdhA1* or SMb21285 fusions the LysR mutant showed no impact upon *xdhA2* expression when grown in media containing purines (Fig. 4.19). This suggested that the *bdhAxdhA2xdhB2* operon is regulated via a different transcriptional regulator so

it is not surprising that this transcriptional regulator would use a inducer molecule other than a purine, in this case, biotin. The fact that XdhC is still required for the activity of the XdhA2/XdhB2 subunits does show that this operon is still dependent upon the expression of the *xdhC* gene in the other operon though presumably the *bdhA* gene could still be expressed and form a functional enzyme even if active XdhA2 and XdhB2 are not formed.

The SMb21291 gene is annotated in the *S. meliloti* Toulouse annotation database as a LysR type transcriptional regulator. BLASTn analysis (expect threshold = 10, word size = 28, match/mismatch scores = 1,-2, gap costs = linear) (NCBI http://blast.ncbi.nlm.nih.gov) shows it most closely matches the similarly annotated sequences in *Sinorhizobium medicae* WSM419 (e-score = 0, identity = 92%), *Rhizobium* species NGR234 and is also similar to sequences from several species of *Brucella* including *B. suis*, *B. abortus* S19 (annotated as a transcriptional regulator, LysR family) and *B. canis* which all have similar e-scores and identity values (76%). The results of the BLAST analysis show that sequences somewhat similar to that of SMb21291 can be found not just in the α-proteobacteria but also in β-proteobacteria (*Polaromonas naphthalenivoran* and *Variovorax paradoxus* for example) as well as the γ-proteobacteria (such as *Xanthomonas axonopodis*) (Fig.4.20). Note that in the tree generated for the sequence comparisons with SMb21291 the sequences from some species such as *Polaromonas naphthalenivoran* are predicted to be a closer match to SMb21291 than is suggested by their BLAST alignment scores (between 45-50) as the tree is generated by using the balanced minimum evolution method described by Desper and Gascuel (Desper and Gascuel, 2004) and not simply a direct translation of the BLAST scores. Again, it must be noted that this annotation is not based on any experimental evidence in any of these species but purely on sequence similarity. The *S. meliloti* gene

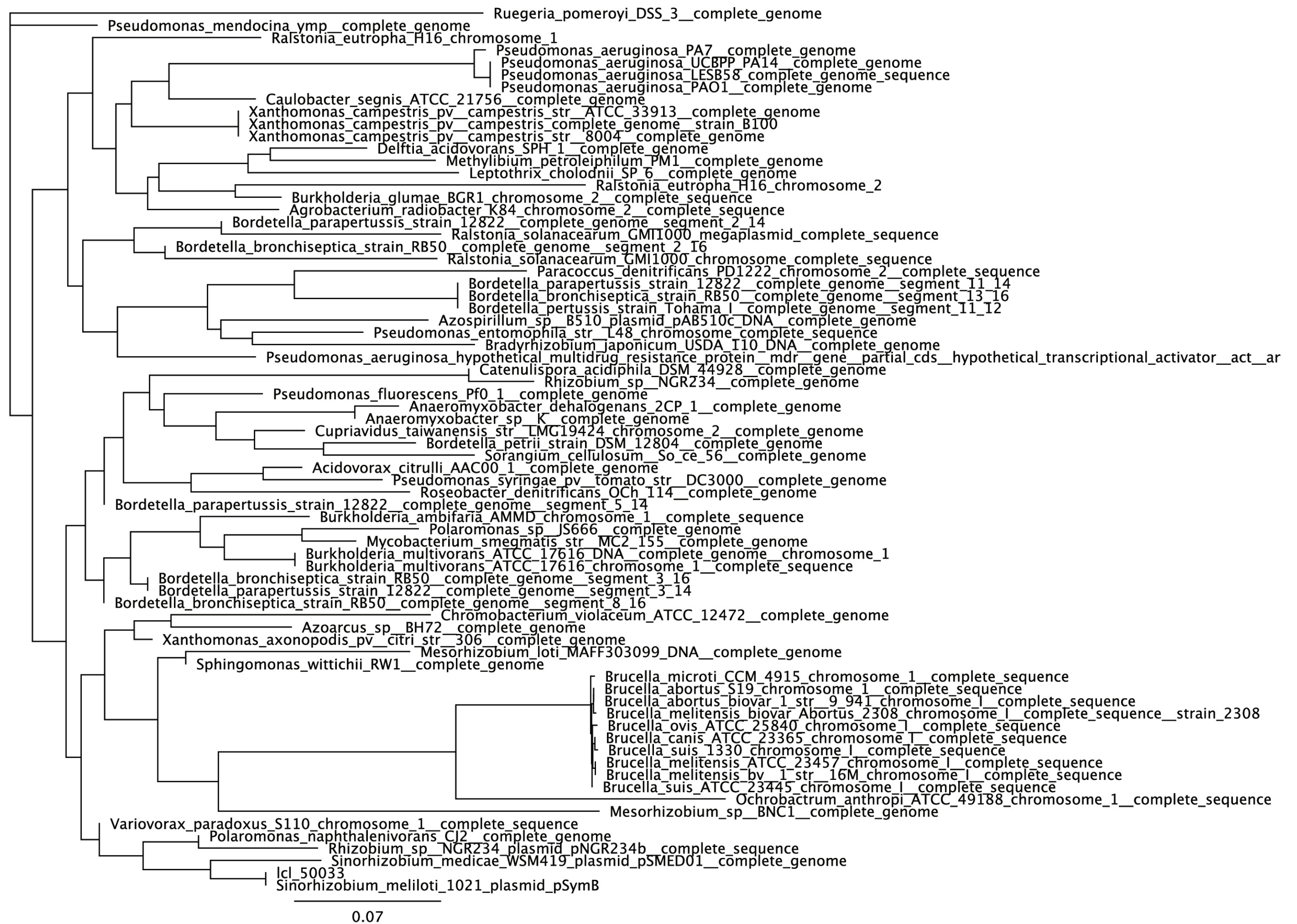

Figure 4.20: Distance tree comparing the similarity of the *S. meliloti* SMb21291 sequence with sequences in other bacterial genomes. The tree was generated by the balanced minimum evolution method of Desper and Gascuel (2004).

SMb20847 is annotated as a probable transcriptional regulator. BLAST analysis shows it is most similar to two other genes. One is a putative LysR regulator in *S. medicae* WSM419 (e-score = 0, identity = 80%) and another putative LysR regulator in *Rhizobium* species NGR234 with an e-score of 1x10^{-155} and an identity of 78%. In most of the cases for the genes of interest examined in this work there has been little or no direct experimental evidence to support the annotations assigned to these genes.

The SMb21291 gene is predicted to code for a protein consisting of 315 amino acids with a predicted molecular weight of 35.4 kDa and a theoretical isoelectric point of 6.47 when calculated using the Compute pI/Mw tool from the Swiss Institute of Bioinformatics (SIB Lausanne, Switzerland, http://www.isb-sib.ch/index.php). Analysis of the amino acid sequence of SMb21291 using the Pfam database (http://pfam.sanger.ac.uk/) reveals two predicted domains of interest. On the N terminal side of the amino acid sequence is a predicted helix-turn-helix (HTH) domain. HTH domains are known to be regions that bind with DNA by interacting with grooves formed in the DNA double helical structure. Typically HTH domains consist of a roughly 20 amino acids forming a pair of α helices crossing at a roughly 120° angle. While these features of HTH motifs are common these domains otherwise show a high degree of variability in structure and tend to be stable only when they exist as a subunit of a larger protein. HTH motifs are known to contain a DNA base pair binding region known as a recognition helix. X-ray and NMR analyses of proteins containing HTH motifs interacting with DNA have shown that the HTH motif interacts with DNA via side chains that extend outward from the second helix of the motif. It is these side chains that make up the recognition helix. The interaction of HTH residues with DNA is known to involve interactions that are more complex than the motif simply recognizing a particular pattern of DNA bases with corresponding patterns of hydrogen

bonds which it interacts with. In reality the HTH residues can bind with short sequences of DNA by having a shape that conforms to the shape of the surface of the DNA and interacts with more than a single groove in the DNA (usually successive grooves). Once in close proximity to the DNA surface the HTH domain interacts with the DNA bases via elaborate interactions with the DNA base pairs and sugar phosphate groups forming "networks of hydrogen bonds, salt bridges and van der Waals contacts" (Voet and Voet, 1995b). The exact patterns of interaction between the HTH motif and a particular target region of DNA are believed to be unique and no other side chain can interact with that particular DNA sequence in the same manner. In some cases, it has been shown that a HTH motif interacts with its target DNA without making direct nonpolar bonding or hydrogen bonds. Such is the case with the *trp* repressor of *E. coli.* In this case X-ray diffraction studies suggest that the HTH motif interacts with successive major grooves in the target DNA region via hydrogen bonds with oxygens belonging to the phosphate groups. None of these interactions however can explain the specificity of the HTH motif for its target DNA. It seems that hydrogen bonding occurs indirectly via water molecules that bridge the gap between the HTH motif and the bases of the target DNA. Thus in such a case it is unclear how exactly the HTH targets a specific stretch of DNA or how a specific region of DNA could be the exclusive target for a particular HTH motif. It has been suggested that this may be a case where the interaction between the *trp* repressor's HTH may simply be the most energetically favourable when it interacts with this region of DNA among all the other possible proteins that could possibly interact with this particular piece of DNA and is thus the one that results in the altered DNA conformation which causes the repression of expression. The bottom line is that there are no simple rules which allow one to predict if or how a particular HTH motif will interact with a given region of DNA. The interactions can be as simple as direct covalent or hydrogen bonds but are just as likely to involve far more subtle interactions such as van der Wall's forces or bridging molecules (Voet and

Voet, 1995b). Given the vast number of proteins with HTH motifs relative to the number that have been studied in any detail it seems obvious that there is still much to be learned about these types of interactions.

The second domain in the SMb21291 amino acid sequence identified through Pfam analysis is designated as a LysR substrate binding domain. This region is located towards the C-terminal end of the predicted protein. It is a domain known to be found in LysR type transcriptional regulators. According to this analysis this region is similar to that found for CysB, tetrameric LysR composed of four 36 kDa subunits from *Klebsiella aerogenes*. In the case of CysB analysis of a crystallized portion of the protein revealed that it is composed of two alpha and two beta domains connected by two short polypeptide segments (Tyrrell *et al.*, 1997). The two domains surround a cavity whose inner surface is lined with polar sidechains. This cavity is predicted via computer modelling to form the binding region for an inducer molecule. The sequence similarity of this region to that of the predicted domain in SMb21291 not only suggests that this region contains the inducer molecule binding domain but that the overall LysR type transcriptional regulator structure may consist of two or more of these subunits together making up the structure of the final regulatory protein. Such a scenario would fit with what is known about several of the LysR structures, though knowledge of LysR structures is admittedly limited to the handful of such proteins that have been purified and subject to detailed structural analysis by X-ray crystallography (Maddocks and Oyston, 2008). The fact that the SMb21291 amino acid sequence contains both a HTH domain as well as a predicted substrate binding domain does however suggest that entire functional protein is probably encoded within this single DNA open reading frame but does not exclude the possibility that the final polypeptides may join together to form dimers, tetramers etc in the final higher level protein structure.

There is one other report of the SMb21291 gene being mutated that can be found in the literature (Luo *et al.*, 2005). This involved the mutating of genes in a screen for genes that have an obvious negative impact upon symbiosis. The authors did not identify a phenotypic impact from the mutation when the strain was used to inoculate plants but this was the extent of their phenotypic testing. In limiting their investigation to just the ability of the mutant to establish symbiosis (indicated by green leaves and pink nodules on the inoculated plants) they may have missed the actual phenotype of the mutant that is shown in the work presented here. We repeated the symbiotic analysis and also found that a symbiosis was established but due to our interest in the purine degradation cycle we were able to establish the actual phenotypic impact of this mutation upon purine utilization and the expression of genes involved in the purine degradation cycle as discussed above.

The intergenic region between SMb21284 and SMb21285 consists of 167 bp. It would be desirable to be able to identify the binding site of a regulator such as the SMb21291 protein within this region. Without being able to isolate the actual protein itself for use in direct binding assays such as DNA footprinting however, this becomes problematic. One of the best studied examples of a binding site for a LysR-type transcriptional regulator is the NOD box associated with NodD, the protein involved in regulating the synthesis of nod factors in rhizobiales in response to chemical signals from plant roots. In this system, there is a clearly defined sequence consisting of two repeats of ATC-N_9-GAT with the ATC and GAT forming a palindromic sequence. The stems of the stem and loop structure formed by the palindromic sequences are spaced closely together (7 bp apart) and the loops are quite small being formed by just 9 bp. The duplication of this pattern is due to the fact that NodD binds to this site as a dimer (Schlaman *et al.*, 1998; Tyrrell *et al.*, 1997;

```
40      gtagcgag      47
        |||||| |
106     catcgccc      99

43      gcgagaat      50
        | | ||||
69      ccccctta      62

44      cgagaattcc      53
        ||||| || |
89      gctctgaaag      80

45      gagaattc      52
        ||| || |
133     ctcatagg     126

47      gaattccg      54
        ||||  ||
65      cttaccgc      58

66      cccccgaaa      74
        || ||||||
158     ggaggcttt     150

77      tttgaaagtc      86
        |  |||||||
156     aggctttcag     147

81      aaagtctc      88
        | ||| ||
134     tctcatag     127

83      agtctcga      90
        ||  ||||
146     tcgtagct     139

87      tcgaaggt      94
        |||| | |
142     agctccta     135

101     cgctacac     108
        |||| | |
162     gcgaggag     155

118     cgcgcttc     125
        ||| | ||
162     gcgaggag     155

132     tctatcctc     140
        ||  |||||
163     agcgaggag     155

139     tcgatgct     146
        ||| | ||
163     agcgagga     156
```

Figure 4.21: Subset of possible palindromic sequences predicted by EMBOSS for the *S. meliloti* genomic region between SMb21284 and SMb21285. Numbers next to the sequences represent the positions within the 167 bp intergenic region where the given portion of the sequence begins and ends.

Fisher and Long, 1993; Goethals *et al.*, 1992; Kondorosi *et al.*, 1989; Fisher *et al.*, 1988; Rostas *et al.*, 1986). While this well known example shows a clearly delineated form of LysR binding site, the number of different LysR coding genes is huge (in the thousands) with this group of regulators forming one of the largest protein families in prokaryotes. Of these thousands of genes the actual protein products have only been purified in a handful of cases(Maddocks and Oyston, 2008). So it seems that making generalizations regarding the potential target binding sites would be premature. Examination of the SMb21284-85 intergenic region shows the difficulty of trying to identify a possible binding site based on sequence data alone. Searching the 167 bp for a Nod box like configuration reveals no exact match. This intergenic region does however have numerous predicted possible palindromic sequences that could form stem and loop DNA structures. Examining this region using EMBOSS explorer (http://emboss.bioinformatics.nl/) reveals a wide variety of such sequences. There are some 109 sites along the 167 bp stretch of DNA that contain potential palindromes (Fig. 4.21). The vast majority of these sites contain 3 bp palindromes like that found in the NOD box sequences but there are many differences as well. There are many that are quite unlike the NOD box configuration in that the loop portion of the possible palindromes stretch over 2/3's or more of the intergenic region for example from bases 1 to 3 paired with bases 104 to 102 thus being 99 bases apart. This would be expected to form a very large loop structure which is quite unlike that seen in the NOD box. There are many other predicted examples that form much smaller loops ranging from 1 to 23 bases as well. Many of the three base pair predicted palindromes are in fact repeated again further along the sequence as seen in the NOD box but the actual configurations are quite different. As just one of numerous examples the palindrome formed by the sequence AGC pairing with TCG is found seven times in the intergenic region. The first five of these are found just before the start of SMb21284 and these predicted loops overlap so only one of the predicted palindromes could actually be present at a given time. The next closest

duplicate is 40 bp away and unlike what was seen in the case of the NOD box the sizes of the predicted loops formed by the palindromic sequence are quite different. The loops predicted near the start of SMb21284 that could be present along with the (assuming there can't be loops within loops as well) palindrome 40 bp away are quite different in size. Near the start of SMb21284 the predicted loop sizes are either 2 to 5 bases or 96 to 99 bases. The same palindrome downstream is predicted to form a loop of some 100 bp.

4.6.5 Growth and Expression of *lacZ* Fusion Strains on Carbon Sources

Growth and expression analysis of the fusion strains on different carbon sources showed that significantly elevated levels of *lacZ* expression and distinct growth phenotypes were only observed when purines were in the media (Table 4.1). No compounds from other classes such as pyrimidines or amino acids were observed to produce levels of expression that differed significantly from the wild-type control strain (Table 4.2). These results support the idea that the *lacZ* fusions in the test strains are in fact showing elevated levels of expression due to being induced by the purines in the media and that the promoter regions to which the *lacZ* genes are fused are responding only to the presence of the purines and not other classes of compounds.

Chapter 5

Attempts to express the SMb21291 LysR protein

One of our final goals for this project was to overexpress and purify SMb21291 gene product. Despite considerable effort involving several different approaches this effort did not prove to be successful. Our initial approach involved cloning the complete gene into the pET30 expression vector (Novagen / EMD Darmstadt, Germany) which allows expression of a protein encoding gene in *E. coli* via a T7 promoter system. Amplifying the SMb21291 DNA via PCR was trivial with excellent amplification but the cloning of the gene sequence into pET30 proved surprisingly difficult. Repeated efforts to isolate a clone proved unsuccessful. After attempting about ten rounds of cloning a single successful clone was isolated. This single clone was sequenced and found to be in frame and in the correct orientation. The plasmid containing the successful clone was transformed into *E. coli* BL21(DE3)pLysS to test for His•Tag-SMb21291 expression. The *E. coli* BL21 culture was grown overnight using an autoinduction medium (6 g/L Na_2HPO_4 3 g/L KH_2PO_4 , 20 g/L

tryptone, 5 g/L yeast extract, 5 g/L NaCl) . The pH of this solution was adjusted to 7.2 with NaOH and autoclaved. 0.2% lactose, 0.6% glycerol, and 0.05% glucose were added to the medium just prior to the inoculation using stock solutions of 8%, 60%, and 10% respectively) with kanamycin (Km) and chloramphenicol (Cm). Cells were harvested, 6X SDS-PAGE sample buffer was added and the cells were boiled for 10 minutes then run on a 10% SDS-polyacrylamide gel (SDS-PAGE). Following electrophoresis one gel was stained with standard Coomassie blue. A second gel was used for a Western blot performed using a His•Tag monoclonal antibody from mouse (Novagen/ EMD) and Alexa Fluor 488 goat anti-mouse IgG (H+L) (Invitrogen, Molecular Probes) as the secondary antibody. The antibody probed membrane was visualized for Alexa Fluor 488 fluorescence with a Typhoon 9400. Purified ChvI protein was used as a control and was consistently observed on Western blots. Attempts to use the clone to produce the protein however failed repeatedly. Different times for harvesting cells post induction with IPTG did not yield any detectable protein on Western Blots nor did growing cells to different culture densities. Attempts to grow cells at various temperatures (20-37°C) likewise did not succeed. Attempts were also made to assay the periplasmic cellular fraction as well as the culture media to see if the protein was exported to the periplasm or from the cell into the media but no evidence of expression was found in either of these fractions. After repeated unsuccessful attempts to overexpress the protein with this initial approach it became apparent that this effort was not going to work. We were uncertain of the reason for this but we speculated that there could be some problem with the insertion of the amplified DNA in the vector or with the vector we were using. We could also not exclude the possibility that the issue could be due to the nature of the gene we were attempting to overexpress. Since the SMb21291 protein is predicted to be a transcriptional regulator, and indeed does impact gene expression in our fusion strain assays we wondered if expressing this gene in *E. coli* had some detrimental impact on the very cells we were attempting to recover the clone from or to overexpress

the protein in. If the protein was able to be expressed once the vector containing the amplified DNA fragment was transformed into the host *E. coli* it might explain why it was so difficult to recover a clone. Normally in this system expression from the pET vector is not supposed to occur in the absence of an inducer such as IPTG or growth in an autoinducing media but given our unusual problems with clone recovery such a possibility could not be ruled out. If this were in fact the case it raised the possibility that the clone we were able to isolate may be somehow defective in expression and this could explain why we were able to recover this single clone as well as the lack of protein expression from this construct. The fact that the insert was in frame and the amplified sequence was correct suggested our clone was correct but we still could not completely rule out this possibility. It was also possible that the induction of expression caused the synthesis of enough protein to negatively impact the cells viability or that any cells expressing the protein were simply unable to grow. It was observed that *E. coli* containing the expression vector had no problem growing in either LB or autoinducing media. The cells also continued growing after the addition of IPTG and in the auto induction media suggesting that the cells were simply unable to express the protein or perhaps rejected the vector if it was causing metabolic problems once expression occurred.

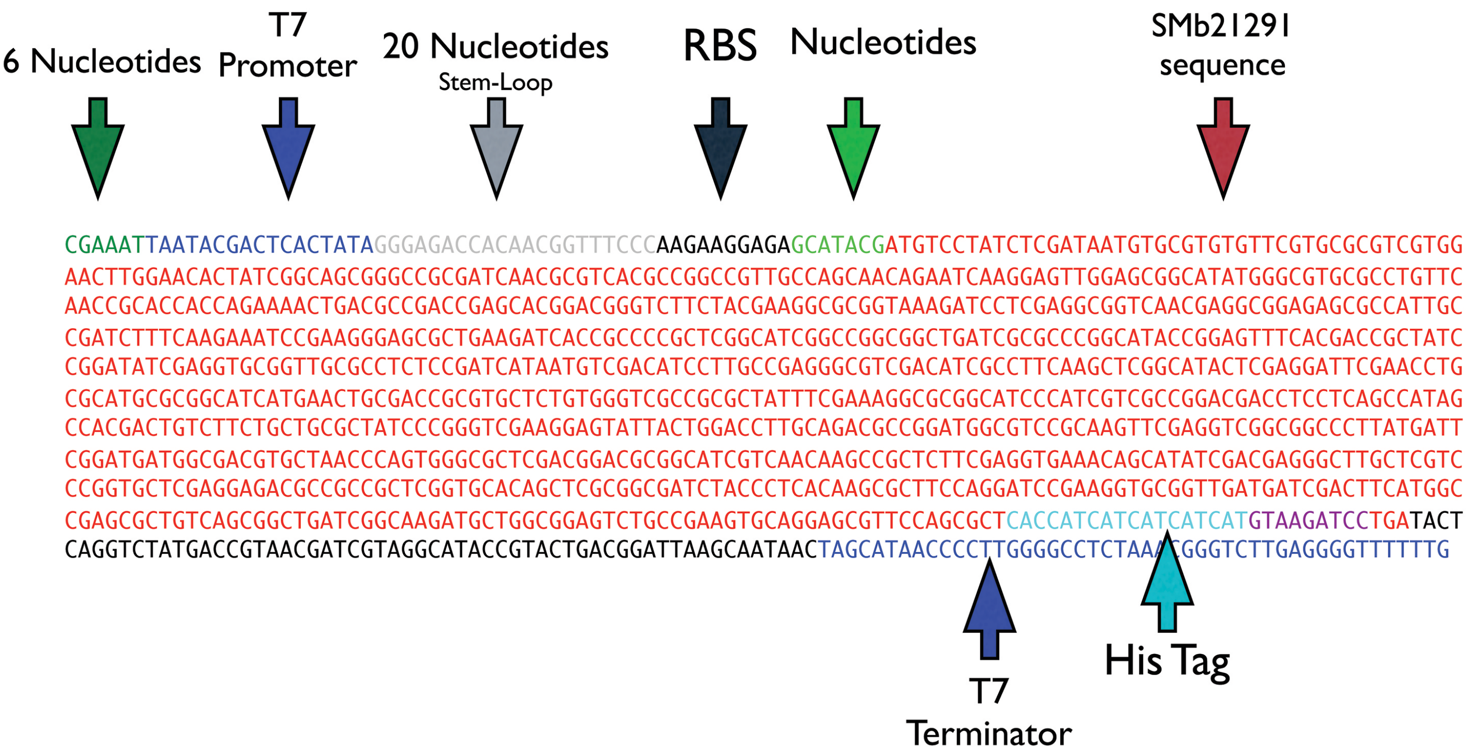

Figure 5.1: Sequence of SMb21291 construct cloned into the pSMART vector showing position of T7 promoter, RBS, SMb21291 sequence, His tag and T7 terminator.

To attempt to get around these possibilities we decided to take a different approach. This involved actually designing the insert including the upstream promoter region, ribosome binding site, start codon, his-tag and stop codon and having the entire construct synthesized instead of doing PCR and ligating the amplified gene into an expression vector or to a cloned promoter region. We reasoned that doing gene synthesis would allow us to have a greater degree of control over the construct we were trying to express and in doing so we would get everything in exactly the correct position and that our construct would be precisely what we wanted by design. This would allow us to overcome possible problems with the cloning, the particular copy of the vector used in the initial cloning or with *in vivo* expression. This construct was designed to be optimal for protein expression in cell-free expression systems and was designed following the guidelines provided by Invitrogen for their Expressway Cell-Free *E. coli* Expression System (Invitrogen, Carlsbad, CA USA) and from Studier *et al.* (1990). The construct was synthesized by Integrated DNA technologies (Coralville, IA, USA) and placed in IDTs pSMART vector (kan^r). We decided to utilize two commercially available cell free expression systems. These systems provide the required cellular machinery and enzymes for protein expression *in vitro*. These required components have been purified from cells and thus allow protein overexpression in reactions that are somewhat analogous to carrying out a PCR in that the amino acids, cell extract, buffer and DNA template are combined in a reaction tube and incubated. Using both these approaches we could test the issue of a possible problem with the isolated clone we obtained as well as bypassing the need for cells to be viable when expressing the protein. Two such commercially available systems were utilized, one from Invitrogen, the Expressway Cell-Free *E. coli* Expression System mentioned above and another from Promega, the S30 T7 High Yield Protein Expression System (Promega Madison, WI, USA). Both systems operate on a very similar principle. Included are all the components that should be required to express a protein from the vector supplied by the user. The kits typically include cell-free

T7 RNA polymerase extract from *E. coli* and amino acids required for protein synthesis. The value of these kits is that they provide a means of expressing a protein that is difficult to express or can't be expressed *in vivo.* The protocol for the Invitrogen expression system consisted of the following (per 100 μl of reaction in a 1.5 ml tube), 20 μl of *E. coli slyD* extract, 20 μl of 2.5X IVPS *E.coli* reaction buffer, 1.25 μl of 50 mM amino acid mix (-methionine), 1 μl of 75 mM methionine, 1 μl T7 enzyme mix, 1 μg of the DNA template and DNase/RNase-free distilled water to a final volume of 50 μl. The mix was incubated for 30 minutes at 300 rpm and 30^oC. Following the 30 minute incubation, 50 μl of feed buffer solution was added (25 μl of 2X IVPS feed buffer, 1.25 μl of 50 mM amino acids (-methionine), 1 μl of 75 mM methionine, and of DNase/RNase-free distilled water to a final volume of 50 μl). The incubation was then continued as before for up to 6 h. The samples were then placed on ice and analyzed as described above. Again the expression assays were carried out using both the original clone and the synthetic construct with each of the cell free systems. All assays were carried out at a range of temperatures (20 to 37 ^oC) and reaction times. Unfortunately no matter which system was used no matter which combination of vector, expression system, temperature or incubation time was tried there was never any evidence of protein expression when the samples were run on polyacrylamide gels stained with Coomassie blue or assayed using anti-His fluorescent antibodies (Western blots).

Chapter 6

General Discussion

The goal of the research outlined in this thesis was to expand our understanding of the purine degradation cycle in *S. meliloti* by continuing the work begun in our lab on the *bdhAxdhA2xdhB2* mixed function operon (Aneja and Charles, 1999) and the random screening of the *S.meliloti* genome for hypoxanthine utilization mutants by Tn5 mutagenesis (Capstick, 2004). The characterization of the mixed function operon had established an interesting connection between the PHB cycle and purine degradation by linking the degradation of PHB with the degradation of hypoxanthine or xanthine. This operon thus allowed for the development of a model whereby *S. meliloti* could concurrently obtain carbon/energy by degrading PHB while degrading purine bases present in the soil environment as a source of nitrogen. The screening of the *S. meliloti* genome by Tn5 mutagenesis generated some additional interesting results. Unexpectedly this work did not isolate any insertions in genes predicted to have functions related to the purine degradation. Instead, mutants were isolated in genes whose only annotation was as hypothetical proteins or a gene involved in the formation of cytochrome *c* (*ccmC*). These unexpected results re-

vealed a couple of points regarding purine degradation. One is that there are genes that had never previously been associated with purine degradation without whose function *S. meliloti* loses the ability to degrade hypoxanthine (a key step in the purine degradation pathway) and use it as a sole carbon and / or nitrogen source. Though it should be noted that the function of these genes could be related to another step in the pathway resulting in the observed phenotype or to some other pathway with out which the ability to utilize this purine is lost. A second point is that the mutation in the *ccmC* gene resulting in the observed hypoxanthine phenotype suggesting a link between the electron transport chain and purine degradation never previously anticipated.

These results from earlier work in our lab began to fill in some of the missing links in purine degradation in *S. meliloti* but left much territory uncharted. Logically for this thesis we began by examining the symbiotic phenotype of the Tn5 mutants. From this it was determined that for the strains with mutations in genes predicted to encode hypothetical proteins while not able to utilize hypoxanthine as a sole nitrogen and / or carbon source were still able to form a symbiosis with the host plant and fix atmospheric nitrogen. The story is not so cut and dry however as was revealed in the competition assay. These results demonstrated that strains with these mutations did in fact suffer some impairment relative to the wild-type strain in terms of actually getting into the nodule (plant). These results suggested that purine utilization may, in fact, play a role in allowing *S. meliloti* to efficiently colonize a host plant. It would seem logical that the colonizing bacteria containing these mutations could be impaired in their ability to obtain enough nitrogen to synthesize proteins vital to this process. The levels of nitrogenase activity seen with these mutant strains were similar for nodules infected with the wild-type strain lending further support to the idea that the mutations in these strains impair the efficiency with which the bacteria

colonize the plant but is not related to the nitrogen fixation process itself.

The results with the *ccmC* mutant strain supported the idea that the fixing of atmospheric nitrogen is a very energy demanding process. By mutating this key gene in the functioning of the electron transport chain the ability to fix atmospheric nitrogen is lost. This would presumably be due to the inability of produce enough ATP to power the process. These results were similar to those found by (Pobigaylo *et al.*, 2008) though we were able to add an additional piece of evidence for the loss of nitrogen fixation via acetylene reduction assays.

The analysis of the mixed function operon as well as the identification of the three genes whose mutation resulted in a hypoxanthine utilization deficiency opened up new insights into the degradation of purines by *S. meliloti*. However an examination of the annotation database suggested that there was still much to be discovered. The previous projects had concentrated on only one of the regions of the genome with a cluster of genes whose sequences suggested roles in purine degradation or identified genes from genomic regions without an obvious link to this process. The one exception was the identification of SMb21292, a gene coding for a hypothetical protein, that when mutated resulted in a hypoxanthine utilization phenotype. This gene is found in a genomic region with a concentration of genes whose sequences suggested a role in purine degradation or whose function was unknown.

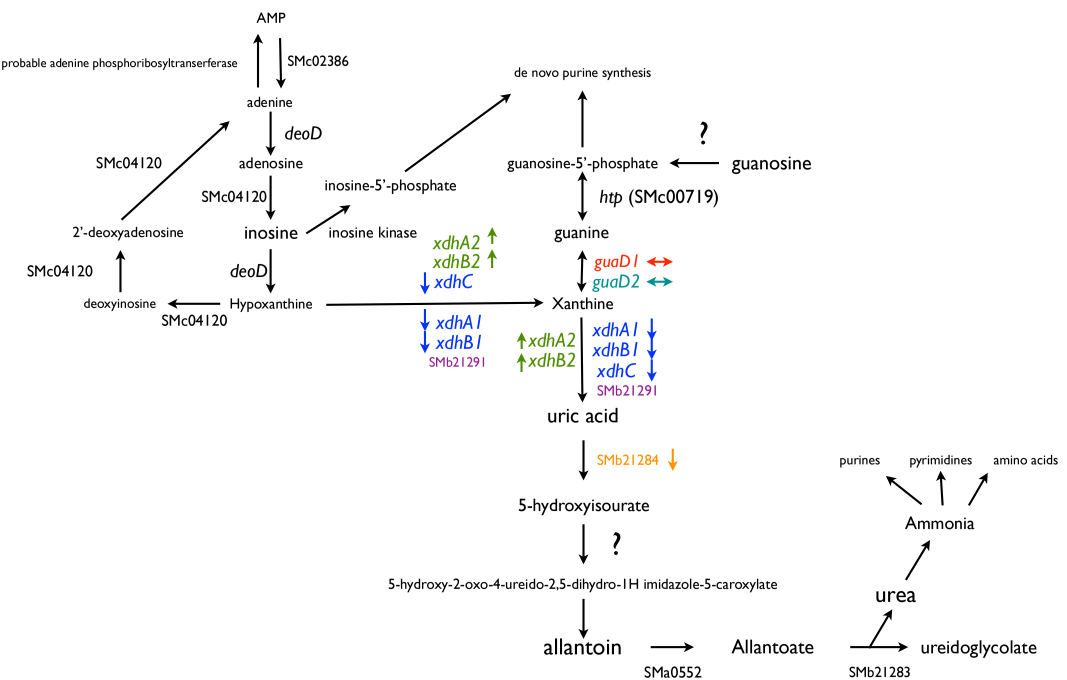

Figure 6.1: Purine degradation pathway showing position of reactions catalyzed or regulated by genes analyzed in this study. Coloured text (other than black) = genes analyzed in this study. Gene names written using the same colour are located in the same operon or not found in an operon. Coloured arrows indicate change in expression of gene when SMb21291 is mutated. *xdh* = xanthine dehydrogenase, *gua* = guanine deaminase.

We decided to focus our attention on this region of the genome by selecting genes whose annotation suggested they could play central roles in a purine degradation pathway (Fig. 6.1). In addition there were genes annotated as having potential transport or regulatory functions that could also be involved in purine degradation. To approach trying to understand the role of these genes in purine degradation, we decided to take advantage of the pTH1703 plasmid since this vector could be used to create reporter fusions to the promoter region of a gene as well as creating a mutation in the same gene. Using this approach, we were able to provide the first experimental evidence supporting the genomic annotation of several of these genes by demonstrating that they, in fact, show significantly elevated levels of expression when grown in the presence of their predicted substrates. This was the case for genes such as *xdhA1*, *guaD1*, SMb21284 (a predicted uricase) as well as *xdhA2* and *guaD2*. We also were able to demonstrate the essential nature of the *xdhC* gene to a complete purine degradation cycle. By mutating this gene, we were able to eliminate the ability of the mutant strain to grow on hypoxanthine or xanthine as a sole source of carbon and nitrogen. The native gel assay demonstrated that the mutation of this gene eliminated Xdh activity. This finding showed that *xdhC* is essential to the formation of a functional Xdh even though there are two separate operons (*bdhAxdhA2xdhB2* and the *xdhA1xdhB1xdhC* containing operon) that code for both of the subunits needed to form the enzyme. The only copy of *xdhC* in the *S. meliloti* genome resides in the same operon as *xdhA1* and *xdhB1* and must be used to create the functional enzyme no matter which operon was used to create the Xdh subunits. In this way, *xdhC* creates a link between the two Xdh encoding operons and also links the *xdhA1xdhB1xdhC* operon to the purine and PHB degradation model proposed earlier from our understanding of the mixed functional nature of *bdhAxdhA2xdhB2*. The pathways are not shown to be linked in the sense that the product of a reaction in one pathway flows directly into a reaction in the other pathway. Instead, the situation is one where in the presence of an inducer (presumably biotin

from the plant) the *bdaAxdhA2xdhB2* operon is induced. The concurrent induction of the *xdhA1xdhB1xdhC* operon due to the presence of purines such as xanthine or hypoxanthine in the soil environment ensures the synthesis of XdhC required to produce functional Xdh using subunits from both operons. This raises the possibility that the presence of a signal from the plant not only leads to the degradation of PHB as a carbon source and purines as a nitrogen source but that induction of the second mixed function operon results in the synthesis of additional Xdh. Thus could ensure that the amount of nitrogen provided to *S. meliloti* is optimized and / or maximized from purine bases during the colonization of the plant.

We also focused our attention on three additional genes. One whose annotation as a permease and whose physical location in the genome suggested a possible role in purine (xanthine in particular) transport. Our assays were not able to confirm any such role as growth in media containing purines did not stimulate increased expression from the promoter region of this gene nor did the mutation of this gene result in any discernible purine phenotype. The second of these genes was SMb20847 a putative LysR transcriptional regulator located just after the *bdhAxdhA2xdhB2* operon. Again, its position and putative function made it intriguing as a possible regulator of purine degradation genes. Mutating this gene, however, did not result in any notable change in the strains growth phenotype on purines nor did the native gel assay results show any difference compared with the wild-type strain. In both cases all three bands of active Xdh could be observed and the pattern of intensity or staining looked indistinguishable between the strains.

The third such gene examined however did lead to very interesting results. SMb21291, a putative LysR transcriptional regulator located in the same region of the genome as the

xdhA1B1C containing operon, the putative uricase SMb21284 as well as *guaD1* lost the ability to grow on hypoxanthine or xanthine when mutated. Examination of the sequence of this gene showed that it has features characteristic of LysRs such as a helix-turn-helix DNA binding domain. This clearly suggested some role for this regulatory gene in purine degradation in *S. meliloti*. Examination of the native gel activity assay results reinforced this idea. On the native gel, it was clearly observed that the activity from the bands representing the XdhA1 and XdhB1 subunits was greatly reduced compared with the wild-type strain. These results suggested a role for SMb21291 in regulating the operon containing *xdhA1B1C*. This was not a completely satisfying result however. It still did not fully explain the inability of the mutant strain to grow on hypoxanthine or xanthine since the native gel showed that there was still Xdh activity. The activity of the XdhA1 / XdhB1 form of the enzyme was greatly reduced but there was still some activity. Furthermore, the activity of the XdhA2 / XdhB2 form of the enzyme was actually higher in the mutant. To gain additional insight into this situation we decided to take advantage of the pTH1703 fusion strains we created. By transducing the reporter fusions into the SMb21291 mutant strain we were able to assess the impact of the LysR mutation on the expression of individual genes. As was seen with the native gel assay the level of expression in *xdhA1* was greatly reduced but remained above the background level. However, the same *lacZ* assay when performed on the SMb21284 fusion showed a very sharp reduction of expression to virtually control levels. This result suggested that the LysR mutation was shutting off the purine degradation pathway at the uricase step (i.e. the next step in the pathway after the conversion of xanthine to uric acid) which could explain the growth phenotype of the mutant. It seems plausible that the SMb21291 protein binds to DNA in the intergenic region between the operons containing SMb21284 and the XdhA1/XdhB1 coding genes. In doing so it seems to act as a positive regulator as the mutation results in a sharp drop in expression from genes in each of these operons. Whether the LysR is acting directly as a

regulator of both operons simultaneously is unclear. It is possible that the downregulation of a gene for one step in the pathway results in reduced expression of the gene coding for the enzyme in the following or preceding steps. Another possibility could be that the absence of the binding regulator results in a distortion in the DNA that is unfavourable to expression. Since both operons share the same intergenic region such a localized distortion could impact expression from both operons. Further understanding of the interaction of the LysR with the intergenic region could be gained though biochemical analysis of the interaction of the protein with the DNA.

References

Akakura, R. and Winans, S.C. (2002a). Constitutive mutations of the OccR regulatory protein affect DNA bending in response to metabolites released from plant tumors. *J Biol Chem*, **277**, 5866–74.

Akakura, R. and Winans, S.C. (2002b). Mutations in the *occQ* operator that decrease OccR-induced dna bending do not cause constitutive promoter activity. *J Biol Chem*, **277**, 15773–80.

Amaya, Y., Yamazaki, K., Soto, M., Noda, K., Nishino, T. and Nishino, T. (1990). Proteolytic conversion of xanthine dehydrogenase from the NAD-dependent type to the O_2-dependent type. Amino acid sequence of rat liver xanthine dehydrogenase and identification of the cleavage sites of the enzyme protein during irreversible conversion by trypsin. *J. Biol. Chem.*, **265**, 14170– 14175.

Anderson, A.J. and Dawes, E.A. (1990). Occurrence, metabolism, metabolic role, and industrial uses of bacterial polyhydroxyalkanoates. *Microbiol Rev*, **54**, 450–72.

Aneja, P. (1999). *Molecular genetic characterization of polyhydroxyalkanoate metabolism in Rhizobium (Sinorhizobium) meliloti*. Ph.D. thesis, University of Waterloo, Waterloo, Ont., Canada.

Aneja, P. and Charles, T.C. (1999). Poly-3-hydroxybutyrate degradation in *Rhizobium* (*Sinorhizobium*) *meliloti*: isolation and characterization of a gene encoding 3-hydroxybutyrate dehydrogenase. *J Bacteriol*, **181**, 849–57.

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

Anraku, Y. (1988). Bacterial electron transport chains. *Annu Rev Biochem*, **57**, 101–32.

Aravind, L., Anantharaman, V., Balaji, S., Babu, M.M. and Iyer, L.M. (2005). The many faces of the helix-turn-helix domain: transcription regulation and beyond. *FEMS Microbiol Rev*, **29**, 231–62.

Ardourel, M., Demont, N., Debellé, F., Maillet, F., de Billy, F., Promé, J.C., Dénarié, J. and Truchet, G. (1994). *Rhizobium meliloti* lipooligosaccharide nodulation factors: different structural requirements for bacterial entry into target root hair cells and induction of plant symbiotic developmental responses. *Plant Cell*, **6**, 1357–74.

Arima, K. and Nose, K. (1968). Studies on bacterial urate:oxygen oxidoreductase. I. Purification and properties of the enzyme. *Biochim Biophys Acta*, **151**, 54–62.

Arwas, R., McKay, I.A., Rowney, F.R.P., Dilworth, M.J. and Glenn, A.R. (1985). Properties of organic acid utilization mutants of *Rhizobium leguminosarum* strain 300. *J. Gen. Microbiol*, **131**, 2059–2066.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1992). *Short protocols in molecular biology*. Greene Publishing Associates and John Wiley & Sons, 2nd edn.

Barbour, W.M., Hattermann, D.R. and Stacey, G. (1991). Chemotaxis of *Bradyrhizobium japonicum* to soybean exudates. *Appl Environ Microbiol*, **57**, 2635–9.

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

Bassam, B.J., Djordjevic, M.A., Redmond, J.W., Batley, M. and Rolfe, B.G. (1988). Identification of a *nodD*-dependent locus in the *Rhizobium* strain NGR234 activated by phenolic factors secreted by soybeans and other legumes. *Mol Plant Microbe Interact*, **1**, 161–8.

Baum, H., Hubseher, G. and Mahler, H.R. (1956). Studies on uricase. II. The enzyme-substrate complex. *Biochim Biophys Acta*, **22**, 514–27.

Belitsky, B.R., Janssen, P.J. and Sonenshein, A.L. (1995). Sites required for GltC-dependent regulation of *Bacillus subtilis* glutamate synthase expression. *J Bacteriol*, **177**, 5686–95.

Ben-Hamida, F. and Schlessinger, D. (1966). Synthesis and breakdown of ribonucleic acid in *Escherichia coli* starving for nitrogen. *Biochim Biophys Acta*, **119**, 183–91.

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

Beringer, J.E. (1974). R factor transfer in *Rhizobium leguminosarum*. *J Gen Microbiol*, **84**, 188–98.

Bernstein, L., Roy, J., Delhotal, K.C., Harnisch, J., Matsuhashi, R., Price, L., Tanaka, K., Worrell, E., Yamba, F. and Fengqi, Z. (2007). *Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.

Birnboim, H.C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res*, **7**, 1513–23.

Blasco, F., Iobbi, C., Giordano, G., Chippaux, M. and Bonnefoy, V. (1989). Nitrate reductase of *Escherichia coli*: completion of the nucleotide sequence of the *nar* operon and reassessment of the role of the alpha and beta subunits in iron binding and electron transfer. *Mol Gen Genet*, **218**, 249–56.

Blasco, F., Pommier, J., Augier, V., Chippaux, M. and Giordano, G. (1992). Involvement of the *narJ* or *narW* gene product in the formation of active nitrate reductase in *Escherichia coli*. *Mol Microbiol*, **6**, 221–30.

Burn, J., Rossen, L. and Johnston., A.W.B. (1987). Four classes of mutations in the *nodD* gene of *Rhizobium leguminosarum* biovar *viciae* that affect its ability to autoregulate and/or activate other *nod* genes in the presence of flavonoid inducers. *Genes Dev.*, **1**, 456–464.

Burn, J.E., Hamilton, W.D., Wootton, J.C. and Johnston, A.W. (1989). Single and multiple mutations affecting properties of the regulatory gene *nodD* of *Rhizobium*. *Mol Microbiol*, **3**, 1567–77.

Burton, K. (1994). Adenine transport in *Escherichia coli*. *Proc. R. Soc. Lond. Ser. B. Biol. Sci.*, **255**, 153–157.

Byrne, G.A., Russell, D.A., Chen, X. and Meijer, W.G. (2007). Transcriptional regulation of the *virR* operon of the intracellular pathogen *Rhodococcus equi*. *J Bacteriol*, **189**, 5082–9.

Byrom, D. (1987). Polymer synthesis by micro-organisms: technology and economics. *Trends Biotechnol.*, **5**, 246–250.

Cao, H., Krishnan, G., Goumnerov, B., Tsongalis, J., Tompkins, R. and Rahme, L.G. (2001). A quorum sensing-associated virulence gene of *Pseudomonas aeruginosa* encodes a LysR-like transcription regulator with a unique self-regulatory mechanism. *Proc Natl Acad Sci U S A*, **98**, 14613–8.

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

Capstick, D.A. (2004). *Genetic Characterization of Hypoxanthine Utilization in Sinorhizobium meliloti*. Master's thesis, University of Waterloo, Waterloo, Ont., Canada.

Cebolla, A., Sousa, C. and de Lorenzo, V. (1997). Effector specificity mutants of the transcriptional activator NahR of naphthalene degrading *Pseudomonas* define protein sites involved in binding of aromatic inducers. *J Biol Chem*, **272**, 3986–92.

Celis, R.T. (1999). Repression and activation of arginine transport genes in *Escherichia coli* K12 by the ArgP protein. *J Mol Biol*, **294**, 1087–95.

Charles, T.C. (2002). *Nitrogen fixation: Global Perspectives*. CABI Publishing, New York, NY.

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

Charles, T.C. and Nester, E.W. (1993). A chromosomally encoded two-component sensory transduction system is required for virulence of *Agrobacterium tumefaciens*. *J Bacteriol*, **175**, 6614–25.

Charles, T.C., Cai, G.Q. and Aneja, P. (1997). Megaplasmid and chromosomal loci for the PHB degradation pathway in *Rhizobium* (*Sinorhizobium*) *meliloti*. *Genetics*, **146**, 1211–20.

Chen, T.H.H. and Murata, N. (2002). Enhancement of tolerance of abiotic stress by metabolic engineering of betaines and other compatible solutes. *Curr Opin Plant Biol*, **5**, 250–7.

Cheng, H.P. and Walker, G.C. (1998). Succinoglycan production by *Rhizobium meliloti* is regulated through the ExoS-ChvI two-component regulatory system. *J Bacteriol*, **180**, 20–6.

Choi, H., Kim, S., Mukhopadhyay, P., Cho, S., Woo, J., Storz, G. and Ryu, S.E. (2001). Structural basis of the redox switch in the OxyR transcription factor. *Cell*, **105**, 103–13.

Christiansen, L.C., Schou, S., Nygaard, P. and Saxild, H.H. (1997). Xanthine metabolism in *Bacillus subtilis*: characterization of the *xpt-pbuX* operon and evidence for purine- and nitrogen-controlled expression of genes involved in xanthine salvage and catabolism. *J Bacteriol*, **179**, 2540–50.

Coronado, C., Zuanazzi, J.A.S., Sallaud, C., Quirion, J.C., Esnault, R., Husson, H.P., Kondorosi, A. and Ratet, P. (1995). Alfalfa root flavonoid production is nitrogen regulated. *Plant Physiol*, **108**, 533–542.

Corte, E.D. and Stirpe, F. (1968). Regulation of xanthine oxidase in rat liver: modifications of the enzyme activity of rat liver supernatant on storage at 20 degrees. *Biochem J*, **108**, 349–51.

Corte, E.D. and Stirpe, F. (1972). The regulation of rat liver xanthine oxidase. Involvement of thiol groups in the conversion of the enzyme activity from dehydrogenase (type D) into oxidase (type O) and purification of the enzyme. *Biochem J*, **126**, 739–45.

Cowie, A., Cheng, J., Sibley, C.D., Fong, Y., Zaheer, R., Patten, C.L., Morton, R.M., Golding, G.B. and Finan, T.M. (2006). An integrated approach to functional genomics: construction of a novel reporter gene fusion library for *Sinorhizobium meliloti*. *Appl Environ Microbiol*, **72**, 7156–67.

Cren, M., Kondorosi, A. and Kondorosi, E. (1995). NolR controls expression of the *Rhizobium meliloti* nodulation genes involved in the core Nod factor synthesis. *Mol Microbiol*, **15**, 733–47.

Cullimore, J.V., Ranjeva, R. and Bono, J.J. (2001). Perception of lipo-chitooligosaccharidic Nod factors in legumes. *Trends Plant Sci*, **6**, 24–30.

Cunningham, S., Kollmeyer, W.D. and Stacey, G. (1991). Chemical control of interstrain competition for soybean nodulation by *Bradyrhizobium japonicum*. *Appl Environ Microbiol*, **57**, 1886–92.

Davis, E.O. and Johnston, A.W. (1990). Regulatory functions of the three *nodD* genes of *Rhizobium leguminosarum* biovar *phaseoli*. *Mol Microbiol*, **4**, 933–41.

Dawes, E.A. (1965). *Microbial energetics*. Blackie & Son Ltd.

Dawes, E.A. and Ribbons, D.W. (1965). Studies on the endogenous metabolism of *Escherichia coli*. *Biochem J*, **95**, 332–43.

Deghmane, A.E., Petit, S., Topilko, A., Pereira, Y., Giorgini, D., Larribe, M. and Taha, M.K. (2000). Intimate adhesion of *Neisseria meningitidis* to human epithelial cells is under the control of the *crgA* gene, a novel LysR-type transcriptional regulator. *EMBO J*, **19**, 1068–78.

Deghmane, A.E., Giorgini, D., Larribe, M., Alonso, J.M. and Taha, M.K. (2002). Down-regulation of pili and capsule of *Neisseria meningitidis* upon contact with epithelial cells is mediated by CrgA regulatory protein. *Mol Microbiol*, **43**, 1555–64.

Delgado, M.J., Bedmar, E.J. and Downie, J.A. (1998). Genes involved in the formation and assembly of rhizobial cytochromes and their role in symbiotic nitrogen fixation. *Adv Microb Physiol*, **40**, 191–231.

Demont, N., Debellé, F., Aurelle, H., Dénarié, J. and Promé, J.C. (1993). Role of the *Rhizobium meliloti nodF* and *nodE* genes in the biosynthesis of lipo-oligosaccharidic nodulation factors. *J Biol Chem*, **268**, 20134–42.

Dénarié, J., Debellé, F. and Rosenberg, C. (1992). Signaling and host range variation in nodulation. *Annu Rev Microbiol*, **46**, 497–531.

Desper, R. and Gascuel, O. (2004). Theoretical foundation of the balanced minimum evolution method of phylogenetic inference and its relationship to weighted least-squares tree fitting. *Mol Biol Evol*, **21**, 587–98.

Downie, J.A. (1998). *The Rhizobiaceae*. Kluwer Academic Publishers.

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

Dunn, M.F. (1998). Tricarboxylic acid cycle and anaplerotic enzymes in rhizobia. *FEMS Microbiol. Rev.*, **22**, 105–123.

El-Din, A.K.Y.G. (1992). A succinate transport mutant of *Bradyrhizobium japonicum* forms ineffective nodules on soybeans. *Can. J. Microbiol.*, **38**, 230– 234.

Encarnacion, S., Osorio, J.C., Mendoza, G., Dunn, M., Contreras, S. and Mora, J. (1998). *Biological Nitrogen Fixation for the 21st Century*. Kluwer Academic Publishers, Dordrecht.

Engelke, T., Jagadish, M.N. and Pühler, A. (1987). Biochemical and genetical analysis of *Rhizobium meliloti* mutants defective in C4-dicarboxylate transport. *J. Gen. Microbiol*, **133:**, 3019–3029.

Feng, J., Qing Li, F., Li, Q., Liang Hu, H. and Fan Hong, G. (2002). Expression and purification of *Rhizobium leguminosarum* NodD. *Protein Expr Purif*, **26**, 321–8.

Finan, T.M., Wood, J.M. and Jordan, D.C. (1983). Symbiotic properties of C4-dicarboxylic acid transport mutants of *Rhizobium leguminosarum*. *J Bacteriol*, **154**, 1403–13.

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**, 120–4.

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., Vorhölter, F.J., Hernandez-Lucas, I., Becker, A., Cowie, A., Gouzy, J., Golding, B. and Pühler, A. (2001). The complete sequence of the 1,683-kb pSymB megaplasmid from the N_2-fixing endosymbiont *Sinorhizobium meliloti*. *Proc Natl Acad Sci U S A*, **98**, 9889–94.

Firmin, J.L., Wilson, K.E., Rossen, L. and Johnston, A.W.B. (1986). Flavonoid activation of nodulation genes in *Rhizobium* reversed by other compounds present in plants. *Nature*, **324**, 90–92.

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

Fisher, R.F., Egelhoff, T.T., Mulligan, J.T. and Long, S.R. (1988). Specific binding of proteins from *Rhizobium meliloti* cell-free extracts containing NodD to DNA sequences upstream of inducible nodulation genes. *Genes Dev*, **2**, 282–93.

Gagnon, H., and Ibrahim, R.K. (1998). Aldonic acids: a novel family of *nod* gene inducers of *Mesorhizobium loti*, *Rhizobium lupini*, and *Sinorhizobium meliloti*. *Mol Plant Microbe Interact*, **11**, 988–998.

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

K.C. and Batut, J. (2001). The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science*, **293**, 668–72.

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

Goethals, K., Gao, M., Tomekpe, K., Van Montagu, M. and Holsters, M. (1989). Common *nodABC* genes in Nod locus 1 of *Azorhizobium caulinodans*: nucleotide sequence and plant-inducible expression. *Mol Gen Genet*, **219**, 289–98.

Goethals, K., Van Montagu, M. and Holsters, M. (1992). Conserved motifs in a divergent nod box of *Azorhizobium caulinodans* ORS571 reveal a common structure in promoters regulated by LysR-type proteins. *Proc Natl Acad Sci U S A*, **89**, 1646–50.

Hanin, M., Jabbouri, S., Broughton, W.J. and Fellay, R. (1998). SyrM1 of *Rhizobium* sp. NGR234 activates transcription of symbiotic loci and controls the level of sulfated Nod factors. *Mol Plant Microbe Interact*, **11**, 343–350.

Harborne, J.B. and Williams, C.A. (2000). Advances in flavonoid research since 1992. *Phytochemistry*, **55**, 481–504.

Harborne, J.B. and Williams, C.A. (2001). Anthocyanins and other flavonoids. *Nat Prod Rep*, **18**, 310–33.

Heidstra, R., Geurts, R., Franssen, H., Spaink, H.P., Van Kammen, A. and Bisseling, T. (1994). Root hair deformation activity of nodulation factors and their fate on *Vicia sativa*. *Plant Physiol*, **105**, 787–797.

Heinz, E.B., Phillips, D.A. and Streit, W.R. (1999). BioS, a biotin-induced, stationary-phase, and possible LysR-type regulator in*Sinorhizobium meliloti*. *Mol Plant Microbe Interact*, **12**, 803–12.

Henikoff, S., Haughn, G.W., Calvo, J.M. and Wallace, J.C. (1988). A large family of bacterial activator proteins. *Proc Natl Acad Sci U S A*, **85**, 6602–6.

Hernández-Lucas, I., Gallego-Hernández, A.L., Encarnación, S., Fernández-Mora, M., Martínez-Batallar, A.G., Salgado, H., Oropeza, R. and Calva, E. (2008). The LysR-type transcriptional regulator LeuO controls expression of several genes in *Salmonella enterica* serovar *Typhi*. *J Bacteriol*, **190**, 1658–70.

Heroven, A.K. and Dersch, P. (2006). RovM, a novel LysR-type regulator of the virulence activator gene *rovA*, controls cell invasion, virulence and motility of *Yersinia pseudotuberculosis*. *Mol Microbiol*, **62**, 1469–83.

Hirsch, A.M., Long, S.R., Bang, M., Haskins, N. and Ausubel, F.M. (1982). Structural studies of alfalfa roots infected with nodulation mutants of *Rhizobium meliloti*. *J Bacteriol*, **151**, 411–9.

Hirsch, A.M., Bang, M. and Ausubel, F.M. (1983). Ultrastructural analysis of ineffective alfalfa nodules formed by *nif::*Tn5 mutants of *Rhizobium meliloti*. *J Bacteriol*, **155**, 367–80.

Hofmann, K., Heinz, E.B., Charles, T.C., Hoppert, M., Liebl, W., and Streit., W.R. (2000a). *Sinorhizobium meliloti* strain 1021 *bioS* and *bdhA* gene transcriptions are both affected by biotin available in defined medium. *FEMS Microbiol. Lett.*, **182**, 41–44.

Hofmann, K., Heinz, E.B., Charles, T.C., Hoppert, M., Liebl, W. and Streit, W.R. (2000b).

Sinorhizobium meliloti strain 1021 *bioS* and *bdhA* gene transcriptions are both affected by biotin available in defined medium. *FEMS Microbiol Lett*, **182**, 41–4.

Hong, G.F., Burn, J.E. and Johnston, A.W. (1987). Evidence that DNA involved in the expression of nodulation (*nod*) genes in *Rhizobium* binds to the product of the regulatory gene *nodD*. *Nucleic Acids Res*, **15**, 9677–90.

Horvath, B., Bachem, C.W., Schell, J. and Kondorosi, A. (1987). Host-specific regulation of nodulation genes in *Rhizobium* is mediated by a plant-signal, interacting with the *nodD* gene product. *EMBO J*, **6**, 841–8.

Huffman, J.L. and Brennan, R.G. (2002). Prokaryotic transcription regulators: more than just the helix-turn-helix motif. *Curr Opin Struct Biol*, **12**, 98–106.

Humbeck, C. and Werner, D. (1989). Delayed nodule development in a succinate transport mutant of *Bradyrhizobium japonicum*. *J. Plant Physiol.*, **134**, 276–283.

Jackson, F.A. and Dawes, E.A. (1976). Regulation of the tricarboxylic acid cycle and poly-beta-hydroxybutyrate metabolism in *Azotobacter beijerinckii* grown under nitrogen or oxygen limitation. *J Gen Microbiol*, **97**, 303–12.

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

Jensen, H.L. (1942). Nitrogen fixation in leguminous plants. I. general characters of root-nodule bacteria isolated from species of *Medicago* and *Trifolium* in Australia. *Proc Linn Soc N S W*, **66**, 98–108.

Jørgensen, C. and Dandanell, G. (1999). Isolation and characterization of mutations in the *Escherichia coli* regulatory protein XapR. *J Bacteriol*, **181**, 4397–403.

Jörnvall, H., Persson, B., Krook, M., Atrian, S., Gonzàlez-Duarte, R., Jeffery, J. and Ghosh, D. (1995). Short-chain dehydrogenases/reductases (SDR). *Biochemistry*, **34**, 6003–13.

Kadouri, D., Jurkevitch, E. and Okon, Y. (2003). Involvement of the reserve material poly-beta-hydroxybutyrate in *Azospirillum brasilense* stress endurance and root colonization. *Appl Environ Microbiol*, **69**, 3244–50.

Kallberg, Y., Oppermann, U., Jörnvall, H. and Persson, B. (2002). Short-chain dehydrogenases/reductases (SDRs). *Eur J Biochem*, **269**, 4409–17.

Kape, R., Parniske, M. and Werner, D. (1991). Chemotaxis and *nod* gene activity of *Bradyrhizobium japonicum* in response to hydroxycinnamic acids and isoflavonoids. *Appl Environ Microbiol*, **57**, 316–9.

Kaplan, R. and Apirion, D. (1975). The fate of ribosomes in *Escherichia coli* cells starved for a carbon source. *J Biol Chem*, **250**, 1854–63.

Kim, J., Kim, J.G., Kang, Y., Jang, J.Y., Jog, G.J., Lim, J.Y., Kim, S., Suga, H., Nagamatsu, T. and Hwang, I. (2004). Quorum sensing and the LysR-type transcriptional activator ToxR regulate toxoflavin biosynthesis and transport in *Burkholderia glumae*. *Mol Microbiol*, **54**, 921–34.

Kisker, C., Schindelin, H. and Rees, D.C. (1997). Molybdenum-cofactor-containing enzymes: structure and mechanism. *Annu Rev Biochem*, **66**, 233–67.

Kondorosi, E., Banfalvi, Z. and Kondorosi., A. (1984). Physical and genetic analysis of a symbiotic region of *Rhizobium meliloti*: identification of nodulation genes. *Mol. Gen. Genet.*, **193**, 445–452.

Kondorosi, E., Gyuris, J., Schmidt, J., John, M., Duda, E., Hoffmann, B., Schell, J. and Kondorosi, A. (1989). Positive and negative control of *nod* gene expression in *Rhizobium meliloti* is required for optimal nodulation. *EMBO J*, **8**, 1331–40.

Kondorosi, E., Pierre, M., Cren, M., Haumann, U., Buiré, M., Hoffmann, B., Schell, J. and Kondorosi, A. (1991). Identification of NolR, a negative transacting factor controlling the nod regulon in *Rhizobium meliloti*. *J Mol Biol*, **222**, 885–96.

Kovacikova, G. and Skorupski, K. (1999). A *Vibrio cholerae* LysR homolog, AphB, cooperates with AphA at the tcpPH promoter to activate expression of the ToxR virulence cascade. *J Bacteriol*, **181**, 4250–6.

Kurkdjian, A.C. (1995). Role of the differentiation of root epidermal cells in Nod factor (from *Rhizobium meliloti*)-induced root-hair depolarization of *Medicago sativa*. *Plant Physiol*, **107**, 783–790.

Kuwabara, Y., Nishino, T., Okamoto, K., Matsumura, T., Eger, B.T., Pai, E.F. and Nishino, T. (2003). Unique amino acids cluster for switching from the dehydrogenase to oxidase form of xanthine oxidoreductase. *Proc Natl Acad Sci U S A*, **100**, 8170–5.

Lafontaine, P.J., LaFrenière, C. and Antoun., H. (1989). Some properties of carbohydrate and C4-dicarboxylic acid utilization negative mutants of *Rhizobium leguminosarum* biovar *phaseoli* strain P121. *Plant Sci.*, **120**, 195–201.

Leimkuhler, S., Kern, M., Solomon, P.S., McEwan, A.G., Schwarz, G., Mendel, R.R. and Klipp, W. (1998). Xanthine dehydrogenase from the phototrophic purple bacterium *Rhodobacter capsulatus* is more similar to its eukaryotic counterparts that to prokaryotic molybdenum enzymes. *Mol. Microbiol.*, **27**, 853– 869.

Lerouge, P., Roche, P., Faucher, C., Maillet, F., Truchet, G., Promé, J.C. and Dénarié, J. (1990). Symbiotic host-specificity of *Rhizobium meliloti* is determined by a sulphated and acylated glucosamine oligosaccharide signal. *Nature*, **344**, 781–4.

Limpens, E., Franken, C., Smit, P., Willemse, J., Bisseling, T. and Geurts, R. (2003). LysM domain receptor kinases regulating rhizobial Nod factor-induced infection. *Science*, **302**, 630–3.

Lindquist, S., Lindberg, F. and Normark, S. (1989). Binding of the *Citrobacter freundii* AmpR regulator to a single DNA site provides both autoregulation and activation of the inducible *ampC* beta-lactamase gene. *J Bacteriol*, **171**, 3746–53.

Liu, X. and DeMoss, J.A. (1997). Characterization of NarJ, a system-specific chaperone required for nitrate reductase biogenesis in *Escherichia coli*. *J Biol Chem*, **272**, 24266–71.

Lochowska, A., Iwanicka-Nowicka, R., Plochocka, D. and Hryniewicz, M.M. (2001). Functional dissection of the LysR-type CysB transcriptional regulator. regions important for DNA binding, inducer response, oligomerization, and positive control. *J Biol Chem*, **276**, 2098–107.

Loh, K.D., Gyaneshwar, P., Markenscoff Papadimitriou, E., Fong, R., Kim, K.S., Parales, R., Zhou, Z., Inwood, W. and Kustu, S. (2006). A previously undescribed pathway for pyrimidine catabolism. *Proc Natl Acad Sci U S A*, **103**, 5114–9.

Lopez, N.I., Floccari, M.E., Garcia, A.F., Steinbuchel, A. and Mendez, B.S. (1995). Effect of poly(3-hydroxybutyrate) (PHB) content on the starvation survival of bacteria in natural waters. *FEMS Microbiol. Ecol.*, **16**, 95–102.

Lu, Z., Takeuchi, M. and Sato, T. (2007). The LysR-type transcriptional regulator YofA

controls cell division through the regulation of expression of *ftsW* in *Bacillus subtilis*. *J Bacteriol*, **189**, 5642–51.

Luo, L., Yao, S.Y., Becker, A., Rüberg, S., Yu, G.Q., Zhu, J.B. and Cheng, H.P. (2005). Two new *Sinorhizobium meliloti* LysR-type transcriptional regulators required for nodulation. *J Bacteriol*, **187**, 4562–72.

MacLean, A.M., Anstey, M.I. and Finan, T.M. (2008). Binding site determinants for the LysR-type transcriptional regulator PcaQ in the legume endosymbiont *Sinorhizobium meliloti*. *J Bacteriol*, **190**, 1237–46.

Maddocks, S.E. and Oyston, P.C.F. (2008). Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. *Microbiology*, **154**, 3609–23.

Madsen, E.B., Madsen, L.H., Radutoiu, S., Olbryt, M., Rakwalska, M., Szczyglowski, K., Sato, S., Kaneko, T., Tabata, S., Sandal, N. and Stougaard, J. (2003). A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals. *Nature*, **425**, 637–40.

Magasanik, B. (1996). *Escherichia coli and Salmonella: cellular and molecular biology*. American Society for Microbiology, Washington D. C.

Mandelstam, J. (1960). The intracellular turnover of protein and nucleic acids and its role in biochemical differentiation. *Bacteriol. Rev.*, **24**, 289–30.

Mandelstam, J. and Halvorson, H. (1960). Turnover of proteins and nucleic acid in soluble and ribosome fractions of non-growing *Escherichia coli*. *Biochim. Biophys. Acta*, **40**, 43–49.

Martinez, E., Romero, D. and Palacios, R. (1990). The *Rhizobium* genome. *Crit. Rev. Plant Sci.*, **9**, 59–93.

Mauchline, T.H., Fowler, J.E., East, A.K., Sartor, A.L., Zaheer, R., Hosie, A.H.F., Poole, P.S. and Finan, T.M. (2006). Mapping the *Sinorhizobium meliloti* 1021 solute-binding protein-dependent transportome. *Proc Natl Acad Sci U S A*, **103**, 17933–8.

McDermott, T.R., Griffith, S.M., Vance, C.P. and Graham, P.H. (1989). Carbon metabolism in *Bradyrhizobium japonicum* bacteroids. *FEMS Microbiol Rev*, **63**, 327–340.

McIver, J., Djordjevic, M.A., Weinman, J.J., Bender, G.L. and Rolfe, B.G. (1989). Extension of host range of *Rhizobium leguminosarum* bv. *trifolii* caused by point mutations in *nodD* that result in alterations in regulatory function and recognition of inducer molecules. *Mol Plant Microbe Interact*, **2**, 97–106.

Mergaert, P., Van Montagu, M., Promé, J.C. and Holsters, M. (1993). Three unusual modifications, a D-arabinosyl, an N-methyl, and a carbamoyl group, are present on the Nod factors of *Azorhizobium caulinodans* strain ORS571. *Proc Natl Acad Sci U S A*, **90**, 1551–5.

Mergaert, P., Van Montagu, M. and Holsters, M. (1997). Molecular mechanisms of Nod factor diversity. *Mol Microbiol*, **25**, 811–7.

Michiels, J., De Wilde, P. and Vanderleyden, J. (1993). Sequence of the *Rhizobium leguminosarum* biovar *phaseoli syrM* gene. *Nucleic Acids Res*, **21**, 3893.

Michiels, J., Pelemans, H., Vlassak, K., Verreth, C. and Vanderleyden, J. (1995). Identification and characterization of a *Rhizobium leguminosarum* bv. *phaseoli* gene that is important for nodulation competitiveness and shows structural homology to a *Rhizobium fredii* host-inducible gene. *Mol Plant Microbe Interact*, **8**, 468–72.

Miller, J.H. (1972). *Experiments in molecular genetics*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

MingSheng, Q., Luo, L., Cheng, H., Zhu, J. and Yu, G. (2009). Identification and characterization of two *gcvA* genes in *Sinorhizobium meliloti*. *Chinese Science Bulletin*.

Miyake, M., Kataoka, K., Shirai, M. and Asada, Y. (1997). Control of poly-beta-hydroxybutyrate synthase mediated by acetyl phosphate in cyanobacteria. *J Bacteriol*, **179**, 5009–13.

Mosier, A.R. (2008). *Nitrogen in the Enviornment. Sources, Problems and Management.*, chap. 13, 444–459. Elsevier Inc, San Diego USA, 2nd edn.

Mulligan, J.T. and Long, S.R. (1985). Induction of *Rhizobium meliloti nodC* expression by plant exudate requires *nodD*. *Proc Natl Acad Sci U S A*, **82**, 6609–13.

Mulligan, J.T. and Long, S.R. (1989). A family of activator genes regulates expression of *Rhizobium meliloti* nodulation genes. *Genetics*, **122**, 7–18.

Muraoka, S., Okumura, R., Ogawa, N., Nonaka, T., Miyashita, K. and Senda, T. (2003a). Crystal structure of a full-length LysR-type transcriptional regulator, CbnR: unusual combination of two subunit forms and molecular bases for causing and changing DNA bend. *J Mol Biol*, **328**, 555–66.

Muraoka, S., Okumura, R., Uragami, Y., Nonaka, T., Ogawa, N., Miyashita, K. and Senda, T. (2003b). Purification and crystallization of a LysR-type transcriptional regulator CbnR from *Ralstonia eutropha* NH9. *Protein Pept Lett*, **10**, 325–9.

Nath, K. and Koch, A.L. (1971). Protein degradation in *Escherichia coli*. II. strain differences in the degradation of protein and nucleic acid resulting from starvation. *J Biol Chem*, **246**, 6956–67.

Nishino, T. (1994). The conversion of xanthine dehydrogenase to xanthine oxidase and the role of the enzyme in reperfusion injury. *J Biochem*, **116**, 1–6.

Nygaard, P. (1983). *Metabolism of nucleotides, nucleosides and nucleobases in microorganisms.*. Academic Press Ltd., London, United Kingdom.

Ogawa, J. and Long, S.R. (1995). The *Rhizobium meliloti groELc* locus is required for regulation of early nod genes by the transcription activator NodD. *Genes Dev*, **9**, 714–29.

Okamura, S., Maruyama, H.B. and Yanagita, T. (1973). Ribosome degradation and degradation products in starved *Escherichia coli*. VI. prolonged culture during glucose starvation. *J Biochem*, **73**, 915–22.

Okker, R.J., Spaink, H., Hille, J., van Brussel, T.A., Lugtenberg, B. and Schilperoort, R.A. (1984). Plant-inducible virulence promoter of the *Agrobacterium tumefaciens* Ti plasmid. *Nature*, **312**, 564–6.

Paau, A.S., Bloch, C.B. and Brill, W.J. (1980). Developmental fate of *Rhizobium meliloti* bacteroids in alfalfa nodules. *J Bacteriol*, **143**, 1480–90.

Paau, A.S., Leps, W.T. and Brill, W.J. (1985). Regulation of nodulation by *Rhizobium meliloti* 102F15 on its mutant which forms an unusually high number of nodules on alfalfa. *Appl Environ Microbiol*, **50**, 1118–22.

Pabo, C.O. and Sauer, R.T. (1984). Protein-DNA recognition. *Annu Rev Biochem*, **53**, 293–321.

Pacios Bras, C., Jordá, M.A., Wijfjes, A.H., Harteveld, M., Stuurman, N., Thomas-Oates, J.E. and Spaink, H.P. (2000). A *Lotus japonicus* nodulation system based on heterologous

expression of the fucosyl transferase NodZ and the acetyl transferase NoIL in *Rhizobium leguminosarum*. *Mol Plant Microbe Interact*, **13**, 475–9.

Parniske, M. and Downie, J.A. (2003). Plant biology: locks, keys and symbioses. *Nature*, **425**, 569–70.

Parsek, M.R., McFall, S.M., Shinabarger, D.L. and Chakrabarty, A.M. (1994). Interaction of two LysR-type regulatory proteins CatR and ClcR with heterologous promoters: functional and evolutionary implications. *Proc Natl Acad Sci U S A*, **91**, 12393–7.

Pérez-Rueda, E. and Collado-Vides, J. (2001). Common history at the origin of the position-function correlation in transcriptional regulators in archaea and bacteria. *J Mol Evol*, **53**, 172–9.

Persson, B., Hedlund, J. and Jörnvall, H. (2008). Medium- and short-chain dehydrogenase/reductase gene and protein families : the mdr superfamily. *Cell Mol Life Sci*, **65**, 3879–94.

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

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

Phillips, D.A., Joseph, C.M. and Maxwell, C.A. (1992). Trigonelline and stachydrine released from alfalfa seeds activate NodD2 protein in *Rhizobium meliloti*. *Plant Physiol*, **99**, 1526–31.

Picossi, S., Belitsky, B.R. and Sonenshein, A.L. (2007). Molecular mechanism of the regulation of *Bacillus subtilis gltAB* expression by GltC. *J Mol Biol*, **365**, 1298–313.

Pobigaylo, N., Szymczak, S., Nattkemper, T.W. and Becker, A. (2008). Identification of genes relevant to symbiosis and competitiveness in *Sinorhizobium meliloti* using signature-tagged mutants. *Mol Plant Microbe Interact*, **21**, 219–31.

Pommier, J., Méjean, V., Giordano, G. and Iobbi-Nivol, C. (1998). TorD, a cytoplasmic chaperone that interacts with the unfolded trimethylamine N-oxide reductase enzyme (TorA) in *Escherichia coli*. *J Biol Chem*, **273**, 16615–20.

Poole, R.K. (1983). Bacterial cytochrome oxidases. a structurally and functionally diverse group of electron-transfer proteins. *Biochim Biophys Acta*, **726**, 205–43.

Porrúa, O., García-Jaramillo, M., Santero, E. and Govantes, F. (2007). The LysR-type regulator AtzR binding site: DNA sequences involved in activation, repression and cyanuric acid-dependent repositioning. *Mol Microbiol*, **66**, 410–27.

Povolo, S., Tombolini, R., Morea, A., Anderson, A.J., Casella, S. and Nuti, M.P. (1994). Isolation and characterization of mutants of *Rhizobium meliloti* unable to synthesize poly-β-hydroxybutyrate (PHB). *Can. J. Microbiol*, **40**, 823– 829.

Price, N.P., Relić, B., Talmont, F., Lewin, A., Promé, D., Pueppke, S.G., Maillet, F., Dénarié, J., Promé, J.C. and Broughton, W.J. (1992). Broad-host-range *Rhizobium* species strain NGR234 secretes a family of carbamoylated, and fucosylated, nodulation signals that are O-acetylated or sulphated. *Mol Microbiol*, **6**, 3575–84.

Radutoiu, S., Madsen, L.H., Madsen, E.B., Felle, H.H., Umehara, Y., Grønlund, M., Sato, S., Nakamura, Y., Tabata, S., Sandal, N. and Stougaard, J. (2003). Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. *Nature*, **425**, 585–92.

Ratcliff, W.C., Kadam, S.V. and Denison, R.F. (2008). Poly-3-hydroxybutyrate (phb)

supports survival and reproduction in starving rhizobia. *FEMS Microbiol Ecol*, **65**, 391–9.

Relić, B., Talmont, F., Kopcinska, J., Golinowski, W., Promé, J.C. and Broughton, W.J. (1993). Biological activity of *Rhizobium* sp. NGR234 Nod-factors on *Macroptilium atropurpureum*. *Mol Plant Microbe Interact*, **6**, 764–74.

Roche, P., Lerouge, P., Ponthus, C. and Promé, J.C. (1991). Structural determination of bacterial nodulation factors involved in the *Rhizobium meliloti*-alfalfa symbiosis. *J Biol Chem*, **266**, 10933–40.

Rolfe, B.G. (1988). Flavones and isoflavones as inducing substances of legume nodulation. *Biofactors*, **1**, 3–10.

Romão, M.J., Knäblein, J., Huber, R. and Moura, J.J. (1997). Structure and function of molybdopterin containing enzymes. *Prog Biophys Mol Biol*, **68**, 121–44.

Ronson, C.W., Lyttleton, P. and Robertson, J.G. (1981). C(4)-dicarboxylate transport mutants of *Rhizobium trifolii* form ineffective nodules on *Trifolium repens*. *Proc Natl Acad Sci U S A*, **78**, 4284–4288.

Rostas, K., Kondorosi, E., Horvath, B., Simoncsits, A. and Kondorosi, A. (1986). Conservation of extended promoter regions of nodulation genes in *Rhizobium*. *Proc Natl Acad Sci U S A*, **83**, 1757–61.

Ruiz, J.A., López, N.I., Fernández, R.O. and Méndez, B.S. (2001). Polyhydroxyalkanoate degradation is associated with nucleotide accumulation and enhances stress resistance and survival of *Pseudomonas oleovorans* in natural water microcosms. *Appl Environ Microbiol*, **67**, 225–30.

Russell, D.A., Byrne, G.A., O'Connell, E.P., Boland, C.A. and Meijer, W.G. (2004). The LysR-type transcriptional regulator VirR is required for expression of the virulence gene *vapA* of *Rhodococcus equi* ATCC 33701. *J Bacteriol*, **186**, 5576–84.

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

Schell, M.A. (1993). Molecular biology of the LysR family of transcriptional regulators. *Annu Rev Microbiol*, **47**, 597–626.

Schlaman, H.R.M., Phillips, D.A. and Kondorosi, E. (1998). *The Rhizobiaceae*. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Schultz, A.C., Nygaard, P. and Saxild, H.H. (2001). Functional analysis of 14 genes that constitute the purine catabolic pathway in *Bacillus subtilis* and evidence for a novel regulon controlled by the PucR transcription activator. *J Bacteriol*, **183**, 3293–302.

Schultze, M. and Kondorosi, A. (1998). Regulation of symbiotic root nodule development. *Annu Rev Genet*, **32**, 33–57.

Schultze, M., Quiclet-Sire, B., Kondorosi, E., Virelizer, H., Glushka, J.N., Endre, G., Géro, S.D. and Kondorosi, A. (1992). *Rhizobium meliloti* produces a family of sulfated lipooligosaccharides exhibiting different degrees of plant host specificity. *Proc Natl Acad Sci U S A*, **89**, 192–6.

Schwedock, J. and Long, S.R. (1989). Nucleotide sequence and protein products of two new nodulation genes of *Rhizobium meliloti*, *nodP* and *nodQ*. *Mol Plant Microbe Interact*, **2**, 181–94.

Selander, R.K., Caugant, D.A., Ochman, H., Musser, J.M., Gilmour, M.N. and Whittam, T.S. (1986). Methods of multilocus enzyme electrophoresis for bacterial population genetics and systematics. *Appl Environ Microbiol*, **51**, 873–84.

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

Senior, P.J. and Dawes., E.A. (1973). The regulation of poly-β-hydroxybutyrate metabolism in *Azobacter beijerinckii. Biochem. J.*, **134**, 225–248.

Senior, P.J., Beech, G.A., Ritchie, G.A. and Dawes, E.A. (1972). The role of oxygen limitation in the formation of poly-β-hydroxybutyrate during batch and continuous culture of *Azotobacter beijerinckii. Biochem J*, **128**, 1193–201.

Sierra, S.B., Rodelas, M.B., Martinez-Toledo, C.P. and Gonzalez-Lopez, J. (1999). Production of B-group vitamins by two *Rhizobium* strains in chemically defined media. *J. Appl. Microbiol.*, **86**, 851–858.

Smirnova, I.A., Dian, C., Leonard, G.A., McSweeney, S., Birse, D. and Brzezinski, P. (2004). Development of a bacterial biosensor for nitrotoluenes: the crystal structure of the transcriptional regulator DntR. *J Mol Biol*, **340**, 405–18.

Sodergren, E.J. and DeMoss, J.A. (1988). *narI* region of the *Escherichia coli* nitrate reductase (*nar*) operon contains two genes. *J Bacteriol*, **170**, 1721–9.

Spaink, H.P., Wijffelman, C.A., Pees, E., Okker, R.J.H. and Lugtenberg, B.J.J. (1987). *Rhizobium* nodulation gene *nodD* as a determinant of host specificity. *Nature*, **328**, 337–340.

Spaink, H.P., Wijffelman, C.A., Okker, R.J.H., and Lugtenberg, B.E.J. (1989). Localization of functional regions of the *Rhizobium nodD* product using hybrid *nodD* genes. *Plant Mol. Biol*, **12**, 59–73.

Spaink, H.P., Sheeley, D.M., van Brussel, A.A., Glushka, J., York, W.S., Tak, T., Geiger, O., Kennedy, E.P., Reinhold, V.N. and Lugtenberg, B.J. (1991). A novel highly unsaturated fatty acid moiety of lipo-oligosaccharide signals determines host specificity of *Rhizobium*. *Nature*, **354**, 125–30.

Sperandio, B., Gautier, C., McGovern, S., Ehrlich, D.S., Renault, P., Martin-Verstraete, I. and Guédon, E. (2007). Control of methionine synthesis and uptake by MetR and homocysteine in *Streptococcus mutans*. *J Bacteriol*, **189**, 7032–44.

Stacey, G., Luka, S., Sanjuan, J., Banfalvi, Z., Nieuwkoop, A.J., Chun, J.Y., Forsberg, L.S. and Carlson, R. (1994). *nodZ*, a unique host-specific nodulation gene, is involved in the fucosylation of the lipooligosaccharide nodulation signal of *Bradyrhizobium japonicum*. *J Bacteriol*, **176**, 620–33.

Stec, E., Witkowska-Zimny, M., Hryniewicz, M.M., Neumann, P., Wilkinson, A.J., Brzozowski, A.M., Verma, C.S., Zaim, J., Wysocki, S. and Bujacz, G.D. (2006). Structural basis of the sulphate starvation response in *E. coli*: crystal structure and mutational analysis of the cofactor-binding domain of the Cbl transcriptional regulator. *J Mol Biol*, **364**, 309–22.

Stirpe, F. and Della Corte, E. (1969). The regulation of rat liver xanthine oxidase. conversion *in vitro* of the enzyme activity from dehydrogenase (type D) to oxidase (type O). *J Biol Chem*, **244**, 3855–63.

Stragier, P. and Patte, J.C. (1983). Regulation of diaminopimelate decarboxylase synthesis

in *Escherichia coli*. III. nucleotide sequence and regulation of the *lysR* gene. *J Mol Biol*, **168**, 333–50.

Stragier, P., Richaud, F., Borne, F. and Patte, J.C. (1983). Regulation of diaminopimelate decarboxylase synthesis in *Escherichia coli*. I. identification of a *lysR* gene encoding an activator of the *lysA* gene. *J Mol Biol*, **168**, 307–20.

Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol*, **185**, 60–89.

Sun, J. and Klein, A. (2004). A LysR-type regulator is involved in the negative regulation of genes encoding selenium-free hydrogenases in the archaeon *Methanococcus voltae*. *Mol Microbiol*, **52**, 563–71.

Swart, S., Lugtenberg, B.J., Smit, G. and Kijne, J.W. (1994). Rhicadhesin-mediated attachment and virulence of an *Agrobacterium tumefaciens chvB* mutant can be restored by growth in a highly osmotic medium. *J Bacteriol*, **176**, 3816–9.

Tal, S. and Okon, Y. (1985). Production of the reserve material poly-β-hydroxybutyrate and its function in *Azospirillum brasilense*. *Can. J. Microbiol.*, **31**, 608–613.

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

Trainer, M.A. and Charles, T.C. (2006). The role of phb metabolism in the symbiosis of rhizobia with legumes. *Appl Microbiol Biotechnol*, **71**, 377–86.

Tropel, D. and van der Meer, J.R. (2004). Bacterial transcriptional regulators for degradation pathways of aromatic compounds. *Microbiol Mol Biol Rev*, **68**, 474–500.

Tyrrell, R., Verschueren, K.H., Dodson, E.J., Murshudov, G.N., Addy, C. and Wilkinson, A.J. (1997). The structure of the cofactor-binding fragment of the LysR family member, CysB: a familiar fold with a surprising subunit arrangement. *Structure*, **5**, 1017–32.

van Berkum, P. and Eardly, B.D. (1998). *Molecular evolutionary systematics of the Rhizobiaceae. In H. Spaink, A. Kondorosi, and P. Hooykaas (ed.).*. Springer.

van Elsas, J.D. and van Overbeek, L.S. (1993). *Starvation in bacteria*. Plenum Press, New York.

van Keulen, G., Girbal, L., van den Bergh, E.R., Dijkhuizen, L. and Meijer, W.G. (1998). The LysR-type transcriptional regulator CbbR controlling autotrophic CO_2 fixation by *Xanthobacter flavus* is an NADPH sensor. *J Bacteriol*, **180**, 1411–7.

van Keulen, G., Ridder, A.N.J.A., Dijkhuizen, L. and Meijer, W.G. (2003). Analysis of DNA binding and transcriptional activation by the LysR-type transcriptional regulator CbbR of *Xanthobacter flavus. J Bacteriol*, **185**, 1245–52.

van Rhijn, P. and Vanderleyden, J. (1995). The *Rhizobium*-plant symbiosis. *Microbiol Rev*, **59**, 124–42.

van Rhijn, P.J.S., Desair, J., Vlassak, K. and Vanderleyden, J. (1994). Functional analysis of *nodD* genes of *Rhizobium tropici* CIAT899. *Mol Plant Microbe Interact*, **7**, 666–676.

van Slooten, J.C., Cervantes, E., Broughton, W.J., Wong, C.H. and Stanley, J. (1990). Sequence and analysis of the *rpoN* sigma factor gene of *Rhizobium* sp. strain NGR234, a primary coregulator of symbiosis. *J Bacteriol*, **172**, 5563–74.

Viale, A.M., Kobayashi, H., Akazawa, T. and Henikoff, S. (1991). *rbcR* [correction of *rcbR*], a gene coding for a member of the LysR family of transcriptional regulators, is located

upstream of the expressed set of ribulose 1,5-bisphosphate carboxylase/oxygenase genes in the photosynthetic bacterium *Chromatium vinosum*. *J Bacteriol*, **173**, 5224–9.

Viswanathan, P., Ueki, T., Inouye, S. and Kroos, L. (2007). Combinatorial regulation of genes essential for *Myxococcus xanthus* development involves a response regulator and a LysR-type regulator. *Proc Natl Acad Sci U S A*, **104**, 7969–74.

Voet, D. and Voet, J.G. (1995a). *Biochemistry Second Edition*. John Wiley and Sons.

Voet, D. and Voet, J.G. (1995b). *Biochemistry Second Edition*. John Wiley and Sons.

Vogels, G.D. and van der Drift, C. (1976). Degradation of purines and pyrimidines by microorganisms. *Bacteriol. Rev.*, **40**, 403–469.

Walshaw, D.L., Wilkinson, A., Mundy, M., Smith, M. and Poole, P.S. (1997). Regulation of the TCA cycle and the general amino acid permease by overflow metabolism in *Rhizobium leguminosarum*. *Microbiology*, **143 (Pt 7)**, 2209–21.

Wang, C., Saldanha, M., Sheng, X., Shelswell, K.J., Walsh, K.T., Sobral, B.W.S. and Charles, T.C. (2007). Roles of poly-3-hydroxybutyrate (PHB) and glycogen in symbiosis of *Sinorhizobium meliloti* with *Medicago* sp. *Microbiology*, **153**, 388–98.

Wang, L.X., Wang, Y., Pellock, B. and Walker, G.C. (1999). Structural characterization of the symbiotically important low-molecular-weight succinoglycan of *Sinorhizobium meliloti*. *J Bacteriol*, **181**, 6788–96.

West, P.M. and Wilson, P.W. (1939). Growth factor requirements of the root nodule bacteria. *J Bacteriol*, **37**, 161–85.

Willis, L.B. and Walker, G.C. (1998). The *phbC* (poly-beta-hydroxybutyrate synthase)

gene of *Rhizobium* (*Sinorhizobium*) *meliloti* and characterization of *phbC* mutants. *Can J Microbiol*, **44**, 554–64.

Wilson, R.L., Urbanowski, M.L. and Stauffer, G.V. (1995). DNA binding sites of the LysR-type regulator GcvA in the *gcv* and *gcvA* control regions of *Escherichia coli*. *J Bacteriol*, **177**, 4940–6.

Witty, J.F., Minchin, F.R., Skøt, L. and Sheehy, J.E. (1986). *Oxford surveys of plant molecular and cell biology*, vol. 3. Oxford University Press, Oxford.

Xi, H., Schneider, B.L. and Reitzer, L. (2000). Purine catabolism in *Escherichia coli* and function of xanthine dehydrogenase in purine salvage. *J Bacteriol*, **182**, 5332–41.

Yeh, K.C., Peck, M.C. and Long, S.R. (2002). Luteolin and GroESL modulate *in vitro* activity of NodD. *J Bacteriol*, **184**, 525–30.

Yurgel, S.N., Berrocal, J., Wilson, C. and Kahn, M.L. (2007). Pleiotropic effects of mutations that alter the *Sinorhizobium meliloti* cytochrome c respiratory system. *Microbiology*, **153**, 399–410.

Zaat, S.A., Wijffelman, C.A., Spaink, H.P., van Brussel, A.A., Okker, R.J. and Lugtenberg, B.J. (1987). Induction of the *nodA* promoter of *Rhizobium leguminosarum* Sym plasmid pRL1JI by plant flavanones and flavones. *J Bacteriol*, **169**, 198–204.

Zuanazzi, J., Clergeot, P.H., Quirion, J.C., Husson, H.P., Kondorosi, A. and Ratet., P. (1998). Production of *Sinorhizobium meliloti nod* gene activator and repressor flavonoids from *Medicago sativa* roots. *Mol Plant Microbe Interact*, **11**, 784–794.