

Characterization of the Rhizobiaceae Protein RhaK

By

Damien M.R. Rivers

A Thesis Submitted to the Faculty of Graduate Studies of
The University of Manitoba
In Partial fulfillment of the Requirements for the Degree of

Doctor of Philosophy

Department of Microbiology
University of Manitoba
Winnipeg, Manitoba, Canada

Copyright © 2015 Damien M.R. Rivers

Abstract

In *Rhizobium leguminosarum* the ABC transporter responsible for rhamnose transport is dependent on RhaK, a sugar kinase that is necessary for the catabolism of rhamnose. It was hypothesized that RhaK has two separate functions; phosphorylation of rhamnose, and an unknown interaction with the rhamnose ABC transporter.

To address this hypothesis a linker-scanning mutagenesis of *rhaK* was carried out. Two generated variants (RhaK72 and RhaK73) were found to maintain kinase activity, but were severely impaired in rhamnose transport function. Structural modelling suggested that both RhaK72 and RhaK73 affect surface exposed residues in two distinct regions localized to one face of the protein. This suggests that this proteins face may play a role in a protein-protein interaction that affects rhamnose transport.

Using a two-hybrid system, an N-terminal and a C-terminal fragment of RhaK were both shown to interact with the N-terminal fragment of RhaT. These fragments span the regions that contain the *rhaK73* and *rhaK72* inserts respectively. When the *rhaK72* and *rhaK73* insert alleles were cloned and assayed using the two-hybrid system, these they were unable to interact with the RhaT fragment, suggesting these inserts abolish transport by interfering with a physical interaction between RhaT and RhaK.

A phylogeny was generated based on the amino acid sequence of RhaK like proteins found in syntenous operons. To gain insight into what residues may constitute a binding domain a PRALINE alignment of the orthologous kinases was combined with secondary structure analysis, known informative mutations, and functional residue predictions. A putative 12 amino acid binding site was identified using this method. An alanine scanning mutagenesis and

subsequent two-hybrid analysis was carried out on this region. The substitution of any of these residues greatly affected the interaction between RhaT and RhaK.

Although heterologous complementation of RhaK is possible, cosmid complementation anomalies and phylogenetic analysis of RhaK indicates the *R. leguminosarum* and *S. meliloti* kinases are different. Through a series of heterologous complementation experiments, enzyme assays, gene fusions, and transport experiments we show that the *R. leguminosarum* kinase is capable of directly phosphorylating rhamnose and rhamnulose, whereas the *Sinorhizobium meliloti* kinase does not have rhamnose kinase activity.

Acknowledgements

I would like to thank everyone who has helped me throughout my education. Above all I would like to thank my Ph.D. supervisor Dr. Ivan J. Oresnik . Thanks for rolling the dice and taking a chance on me. Whether you were giving me some much needed advice and guidance, or allowing me to learn on my own; it was very appreciated. Thank you to my graduate studies committee Dr. Deborah Court, Dr. Silvia T. Cardona, and Dr. Jörg Stetefeld, for all of the invaluable research advice. Thanks to my external examiner Dr. Turlough M. Finan. Your questions made me consider my work from a perspective I would never have otherwise.

Thanks to all the staff and students from the University of Manitoba department of Microbiology. I would like to specifically thank Mark Miller-Williams, Brad Pickering, Barney Geddes, Harry Yudistira, Justin Hawkins, Mac Kohlmeier, and Jackie Donogh of the Oresnik lab not only for all the help over the years, but for making coming to work such a pleasure. In particular thank you to Barney Geddes, not only for all the invaluable conversations about our projects, but for all the great times, and for being such a good friend.

Finally I would like to thank all my friends and family for the support. Whether it was playing music or hockey, or just sharing a meal, the ability to get away and clear my head made more difference than you know.

Dedication

This thesis is dedicated to my wife Bernice Rivers, and my parents Ron and Camilla Rivers.

Thank you to my parents. You have supported me throughout my education in more ways than I can count. Thanks most of all for starting me out on this path by instilling curiosity in me, and then supporting me no questions asked, whatever I chose to do with my life.

Thanks most of all to my wife. Thanks for supporting me unconditionally, even though you knew it would be a long hard path. Thanks for letting me be stressed or grumpy or whatever else I needed. You are the only one who knows how to handle me, and you always gave me just what I needed without me ever needing to ask . Thank you. I love you.

Table of Contents

ABSTRACT.....	ii
ACKNOWLEDGEMENTS.....	iv
DEDICATION.....	v
TABLE OF CONTENTS.....	vi
LIST OF TABLES.....	xi
LIST OF FIGURES.....	xii
LIST OF COPYRIGHTED MATERIAL.....	xiv
LIST OF ABBREVIATIONS AND DEFINITIONS.....	xv
CHAPTER 1: LITERATURE REVIEW.....	1
<i>1.1 Rhizobaceae</i>	2
1.1.1 Introduction to Rhizobia.....	2
1.1.2 Taxonomy of Rhizobium.....	2
1.1.3 Rhizobium genetic make up.....	3
1.1.4 <i>Sinorhizobium meliloti</i> Rm1021.....	4
1.1.5 <i>Rhizobium leguminosarum</i> biovar <i>trifoli</i> Rlt100.....	5
<i>1.2 Rhizobium symbiosis</i>	6
1.2.1 Biological nitrogen fixation.....	6
1.2.2 Root invasion and nodule formation.....	6
1.2.2.1 Determinate nodules.....	10
1.2.2.2 Indeterminate nodules.....	11
1.2.2.2.1 Symbiosomes and bacterioids in indeterminate nodule.....	11
<i>1.3 Rhamnose catabolism</i>	15
1.3.1 Plasmid curing experiments in <i>R. leguminosarum</i>	15
1.3.2 Canonical rhamnose catabolism in bacteria.....	16

1.3.2.1	<i>Transport of rhamnose by the bacterial cell.....</i>	17
1.3.2.2	<i>Isomerisation of rhamnose to rhamnulose</i>	19
1.3.2.3	<i>Phosphorylation of rhamnulose to rhamnulose-1-P.....</i>	20
1.3.2.4	<i>Metabolism of rhamnulose-1-P.....</i>	22
1.3.2.5	<i>Regulation of l-rhamnose catabolic regulons.....</i>	24
1.3.3	<i>A locus required for rhamnose catabolism in <u>R. leguminosarum</u></i>	
	<i>Rlt100</i>	25
1.3.3.1	<i><u>rhaI</u> encodes a <u>P. stutzeri</u> type rhamnose isomerase.....</i>	25
1.3.3.2	<i><u>rhaD</u> encodes a predicted dehydrogenase/ aldolase.....</i>	26
1.3.3.3	<i><u>rhaR</u> encodes a DeoR type negative regulator.....</i>	26
1.3.3.4	<i><u>rhaSTPQ</u> encodes the components of a rhamnose ABC transporter.....</i>	27
1.3.3.5	<i><u>rhaU</u> encodes a mutarotase.....</i>	30
1.3.3.6	<i><u>rhaK</u> encodes a sugar kinase.....</i>	31
1.3.3.7	<i>Basic regulation of the Rlt100 rhamnose operon.....</i>	31
1.3.3.8	<i><u>R. leguminosarum</u> appears to have a rhamnose catabolic pathway that is different than the canonical phosphorylated rhamnose pathway.....</i>	32
1.3.3.9	<i>RhaK is required for the transport of rhamnose by <u>R. leguminosarum</u>.....</i>	33
1.4	<i>Transport across the inner membrane.....</i>	35
1.4.1	<i>Introduction.....</i>	35
1.4.2	<i>Phospho-enol-pyruvate-dependent carbohydrate phosphotransferase system (PTS).....</i>	36
1.4.3	<i>ABC transporters.....</i>	40
1.4.3.1	<i>Introduction.....</i>	40
1.4.3.2	<i>The core structure of an ABC transporter.....</i>	40
1.4.3.3	<i>Evolution and spread of ABC transporters.....</i>	42
1.4.3.4	<i>Discovery of ABC-type importers in bacteria.....</i>	43
1.4.3.5	<i>Prokaryotic importers.....</i>	44
1.4.3.6	<i>Gram positive and Archea substrate binding protein dependant importers.....</i>	44
1.4.3.7	<i><u>Carbohydrate Uptake Transporters 1 and 2 (CUT 1 and 2)</u>.....</i>	45
1.4.3.8	<i><u>E. coli</u> Maltose transporter as a model for the mechanism of transport.....</i>	47
1.4.3.9	<i>Variations of the ABC importer architecture.....</i>	50
1.5	<i>Protein-protein interactions.....</i>	52
1.5.1	<i>Basic factors responsible for protein-protein interactions.....</i>	52
1.5.2	<i>Protein binding hot spots.....</i>	53

1.5.3	<i>Protein-Protein interactions involving transport systems and sugar kinases.....</i>	54
1.5.3.1	<i>Glucose PTS component EIIA^{glc} interacts with LacY and MelB.....</i>	54
1.5.3.2	<i>MalK physically interacts with MalT.....</i>	55
1.5.3.3	<i>EIIA^{Glc} interacts with MalK.....</i>	55
1.6	Thesis objectives.....	56
CHAPTER 2: CARBOHYDRATE KINASE (RHA K) DEPENDANT ABC-TRANSPORT OF RHAMNOSE IN RHIZOBIUM LEGUMINOSARUM: GENETIC SEPARATION OF KINASE AND TRANSPORT ACTIVITIES.....		
2.1	Introduction.....	60
2.2	Materials and Methods.....	62
2.2.1	<i>Bacterial strains and culture conditions.....</i>	62
2.2.2	<i>Genetic manipulations.....</i>	65
2.2.3	<i>DNA manipulations, constructions.....</i>	65
2.2.4	<i>Construction of pDR1.....</i>	65
2.2.5	<i>Generation of Tn7 in frame insertional <u>rhaK</u> alleles.....</i>	66
2.2.6	<i>Rhamnose transport assay.....</i>	66
2.2.7	<i>Biochemical assays.....</i>	67
2.2.8	<i>Isolation of cell lysate membrane fraction.....</i>	67
2.2.9	<i>Western blots.....</i>	68
2.2.10	<i>Bioinformatic analysis and molecular modelling RhaK.....</i>	68
2.3	Results.....	68
2.3.1	<i>RhaK can associate with the cytoplasmic membrane.....</i>	68
2.3.2	<i>Generation of <u>rhaK</u> mutant alleles.....</i>	72
2.3.3	<i>Screening of linker scanning <u>rhaK</u> alleles.....</i>	73
2.3.4	<i>Biochemical properties of linker scanning alleles.....</i>	77
2.3.5	<i>Alleles that uncouple transport and kinase activity have different complementation phenotypes.....</i>	79
2.3.6	<i>Molecular architecture of a RhaK.....</i>	81
2.4	Discussion.....	87

CHAPTER 3: THE SUGAR KINASE (RHA_K) THAT IS NECESSARY FOR THE CATABOLISM AND TRANSPORT OF RHAMNOSE IN *RHIZOBIUM LEGUMINOSARUM* DIRECTLY INTERACTS WITH THE ABC TRANSPORTER COMPONENT (RHA_T)..... 91

3.1 Introduction.....	92
3.2 Materials and Methods.....	94
3.2.1 Bacterial strains, plasmids, and media.....	94
3.2.2 Clonings, and genetic manipulations	97
3.2.3 Bacterial two-hybrid analysis.....	97
3.2.4 Generation of motif variants.....	97
3.2.5 Beta-galactosidase assays.....	98
3.2.6 Phylogenetic and bioinformatic analysis.....	98
3.3 Results.....	99
3.3.1 Two hybrid analysis of Rha _K and Rha _T	99
3.3.2 The <i>rhaK72</i> and <i>rhaK73</i> alleles disrupt two hybrid interactions.....	103
3.3.3. Identification of putative motifs in the C- and N- terminal fragments of Rha _K	104
3.3.4. Identification of a putative interacting domain.....	108
3.3.5 Characterization of the N-terminal motif.....	110
3.4 Discussion.....	114

CHAPTER 4: INABILITY TO UTILIZE RHAMNOSE BY *SINORHIZOBIUM MELILOTI* RM1021AFFECTS COMPETITION FOR NODULE OCCUPANCY... 117

4.1 Introduction.....	118
4.2 Materials and Methods.....	120
4.2.1 Bacterial strains, plasmids, and media.....	120
4.2.2 Genetic techniques.....	123
4.2.3 DNA manipulations.....	123
4.2.4 Rhamnose transport assay.....	124

4.2.5 Enzyme assays.....	124
4.2.6 β -Galactosidase assays.....	124
4.2.7 Plant assays.....	125
4.3 Results.....	125
4.3.1 Identification of a rhamnose catabolic operon in <i>S. meliloti</i>	125
4.3.2 Complementation of <i>S. meliloti</i> rhamnose mutants with <i>R. leguminosarum</i> rhamnose genes.....	126
4.3.3 <i>S. meliloti</i> rhaDI mutants are not sensitive on rhamnose/glycerol Media.....	127
4.3.4 <i>S. meliloti</i> rhaK (<i>rhaK_{Sm}</i>) and rhaI (<i>rhaI_{Sm}</i>) can complement <i>R. leguminosarum</i> mutants.....	128
4.3.5 <i>RhaK_{Sm}</i> does not have measurable rhamnose kinase activity.....	131
4.3.6 The rhamnose transporter RhaSTPQ is required for growth on rhamnose.....	133
4.3.7 Cosmids carrying the <i>R. leguminosarum</i> rhamnose locus show reduced expression in Rm1021	135
4.3.8 <i>S. meliloti</i> rhamnose mutants are uncompetitive for nodule occupancy on alfalfa.....	139
4.4 Discussion.....	141
CHAPTER 5: GENERAL THESIS CONCLUSIONS.....	145
5.1 Thesis conclusions and observations.....	146
5.1.1 Project history.....	146
5.1.2 The kinase activity and transport role of RhaK can be uncoupled....	146
5.1.3 RhaK physically interacts with RhaT.....	147
5.1.4 A putative binding domain has been identified in the N-terminal region of RhaK that is required for its interaction with RhaT.....	148
5.1.5 Some RhaK alleles have intermediate transport phenotypes.....	149
5.1.6 Developing a model for RhaK dependant transport.....	150
5.1.7 <i>RhaK_{Sm}</i> and <i>RhaK_{RI}</i> differ in catabolic function.....	153
5.2. Future work on RhaK dependant rhamnose transport.....	155
REFERENCES.....	158

List of Tables

Table 2.1. Strains and Plasmids.....	63
Table 2.2. Complementation analysis of <i>rhaK50</i> using <i>rhaK</i> linker scanning alleles.....	75
Table 2.3. <i>rhaK72</i> and <i>rhaK73</i> alleles confer different growth rates when expressed in Rlt144 (<i>rhaK50::Tn5</i>).....	80
Table 3.1. Strains and plasmids.....	95
Table 4.1. Strains and Plasmids.....	121
Table 4.2. <i>S. meliloti</i> displays different conditional growth phenotypes than <i>R. leguminosarum</i>	130
Table 4.3. Plasmid encoded <i>rhaK</i> and <i>rhaI</i> can heterologously complement <i>rhaK</i> and <i>rhaI</i> mutations respectively.....	132
Table 4.4. Kinase and isomerase activity of <i>R. leguminosarum</i> and <i>S. meliloti</i> strains.....	134
Table 4.5. Induction of <i>R. leguminosarum</i> cosmid fusions in either <i>R. leguminosarum</i> or <i>S. meliloti</i> backgrounds.....	138

List of Figures

Figure 1.1. Development of the infection thread.....	9
Figure 1.2. The nodule structure of indeterminate (A) and determinate (B) nodule types.	12
Figure 1.3. Summary of the canonical l-rhamnose utilization pathway in bacteria.....	18
Figure 1.4. A) The oligomeric structure of <i>E. coli</i> xylulose kinase, and the allosterically-inhibited <i>E. coli</i> glycerol kinase tetramer.....	23
Figure 1.5 The rhamnose region of <i>R. leguminosarum</i> bv. <i>trifolii</i> train Rlt100.....	28
Figure 1.6. The predicted amino acid conservation of the RhaI proteins from <i>R.leguminosarum</i> Rlt100, <i>S. meliloti</i> Rm1021, and <i>P. stutzeri</i>	29
Figure 1.7 A diagram of the basic phospho-transferase system used by bacteria for the uptake of carbohydrates.....	39
Figure 1.8 The current model for the ATP dependant ABC transport mechanism.....	49
Figure 2.1. Western blot analysis of RhaK in cellular fractions.....	71
Figure 2.2. Linker scanning mutagenesis.....	74
Figure 2.3. Transport and kinase rates of representative <i>rhaK</i> alleles.....	78
Figure 2.4. Accumulation of rhamnose in Rlt144 carrying either <i>rhaK72</i> or <i>rhaK73</i>	82
Figure 2.5. Predicted structure of wild type RhaK and variants.....	86
Figure 3.1. A diagram indicating which residues make up each RhaK and RhaT fragment translationally fused to Cya fragments T25 and T18.....	101
Figure 3.2. Beta-galactosidase activity of BTH101 cells co-transformed with the RhaK and the RhaT fragment indicated.....	102
Figure 3.3. Beta-galactosidase activity of BTH101 cells co-transformed with the variant RhaK fragment and the RhaT fragment indicated.....	105
Figure 3.4. A phylogenetic tree of RhaK-like proteins from α -proteobacteria.....	107

Figure 3.5. The amino acid conservation of 23 RhaK like proteins.....	109
Figure 3.6. (A) The consensus sequences of the putative binding motif as generated using WebLogo. (B) Beta-galactosidase activity of BTH101 cells co-transformed with a RhaKF1 variant fragment, and RhaTF1.....	112
Figure 4.1. The rhamnose catabolic regions in <i>S. meliloti</i> strain Rm1021 (A), and <i>R. leguminosarum</i> <i>bv. trifolii</i> strain Rlt100 (B).....	129
Figure 4.2. Rhamnose transport assays.....	136
Figure 4.3. Competition for nodule occupancy assays between Rm1021 (wild-type) and the rhamnose catabolic mutant SRmA145 (<i>rhaT</i> ::Tn5).....	137
Figure 5.1. RhaK and RhaT interactions consistent with two-hybrid data.....	152

List of Copyrighted Materials

- Figure 1.1.** Development of the infection thread..... 9
Taken From: Oldroyd, G. E. D., et al. (2011). "The rules of engagement in the legume-rhizobial symbiosis." *Annual review of genetics* 45: 119-144.
Reproduced with permission obtained from: <http://support.copyright.com>
- Figure 1.2.** The nodule structure of indeterminate (A) and determinate (B) nodule types. 12
Taken From: Luyten, E. and J. Vanderleyden (2000). "Survey of genes identified in *Sinorhizobium meliloti* spp., necessary for the development of an efficient symbiosis." *European Journal of Soil Biology* 36 (1): 1-26.
Reproduced with permission obtained from: <http://support.copyright.com>
- Figure 1.3.** Summary of the canonical l-rhamnose utilization pathway in bacteria..... 18
Taken from: Rodionova, I. A., et al. (2013). "Comparative genomics and functional analysis of rhamnose catabolic pathways and regulons in bacteria." *Front Microbiol.*(4)
Reproduced with permission obtained from: <http://support.copyright.com>
- Figure 1.4.** A) The oligomeric structure of *E. coli* xylulose kinase, and the allosterically-inhibited *E. coli* glycerol kinase tetramer..... 23
Modified from: Di Luccio, E., et al. (2007). "Structural and kinetic studies of induced fit of xylulose kinase from *Escherichia coli*." *J. Mol. Biol.* (365).
Reproduced with permission obtained from:
<https://s100.copyright.com/AppDispatchServlet?publisherName=ELS&contentID=S002X1000020X&orderBeanReset=true>
- Figure 1.7** A diagram of the basic phospho-transferase system used by bacteria for the uptake of carbohydrates..... 39
Modified from: Pflüger-Grau, K. and B. Görke (2010). "Regulatory roles of the bacterial nitrogen-related phosphotransferase system." *Trends in microbiology* 18(5): 205-214.
Reproduced with permission obtained from:
<https://s100.copyright.com/AppDispatchServlet?publisherName=ELS&contentID=S0966842X1000020X&orderBeanReset=true>
- Figure 1.8.** the current model for the ATP dependant ABC transport mechanism..... 49
Taken from: Bordignon, E., et al. (2010). "The maltose ATP_binding cassette transporter in the 21st century—towards a structural dynamic perspective on its mode of action." *Molecular Microbiology* 77(6): 1354-1366.
Reproduced with permission obtained from: <http://support.copyright.com>

Please note: All of the above Journals indicated materials may be republished in a thesis / dissertation without obtaining additional permission, providing that the author and the original source of publication are fully acknowledged.

List of Abbreviations

Å = angstrom

°C = degrees Celsius

ASEdb = Alanine Scanning Energetics database

ABC = ATP binding cassette

Amp = ampicillin

ATP = adenosine triphosphate

BLAST = basic local alignment search tool

bp = base pair

bv. = biovar

CCaMK = calcium/calmodulin-dependent protein kinase

Cm = chloramphenicol

CUT 1 = Carbohydrate uptake transporters class 1

CUT 2 = Carbohydrate uptake transporters class 2

DNA = deoxyribonucleic acid

GlcNAc = β -1,4-linked N-acetyl-D168 glucosamine

Gm = gentamycin

HPr = heat-stable protein

IMG = Integrated Microbial Genomes

IPTG = isopropyl β -D-1-thiogalactopyranoside

Kan = kanamycin

kcal = kilocalories

L = litre

LB = Luria-Bertani broth

M = molar units (mol/L)

m = meter

Mb = megabase

mol = mole

NBD = nucleotide binding domain

NCBI = National Center for Biotechnology Information

NCRs = nodule specific cysteine rich antimicrobial like peptides

Nm = neomycin

OD = optical density

ORF = open reading frame

PCR = polymerase chain reaction

PDB = Protein Data Bank

PEP = phosphoenol-pyruvate

PTS = phosphoenol-pyruvate-dependent phosphotransferase system

qPCR = quantitative real-time PCR

Rf = rifampicin

SBP = substrate binding protein

SDS-PAGE = sodium dodecyl sulphate-polyacrylamide gel electrophoresis

Sm = streptomycin

Tc = tetracycline

TCA cycle = tricarboxylic acid cycle

TMD = transmembrane domain

TY = tryptone yeast

VMM = Vincent's minimal media

X-gal = 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside

Chapter 1:
Literature Review

1.1 Rhizobiaceae

1.1.1 Introduction to Rhizobia

The word rhizobia is derived from the Greek rhiza, meaning "root" and bios, meaning "life". Rhizobia are Gram-negative soil dwelling, motile rods that are members of the α -proteobacteria. They can interact with legumes, forming a symbiotic relationship where the bacteria fix nitrogen from N_2 to ammonium in exchange for organic acids (carbon source) in a microoxic environment called the nodule. Biological nitrogen fixation utilizes the enzyme nitrogenase, which is irreversibly inactivated by oxygen (Burgess 1984).

Rhizobium leguminosarum was identified in 1889 and represents the first *Rhizobium* species identified. Subsequently, all root nodulating, nitrogen fixing bacteria were placed in the genus named *Rhizobium*, with the species name often indicative of the host they were isolated from. Modern analysis of this genus however makes it clear that organisms originally classified as *Rhizobium* based on their biological role actually fall into two classes, and many different genera (Willems 2006).

1.1.2 Taxonomy of Rhizobium

Bacteria classically defined as *Rhizobium* (root nodulating, nitrogen fixing bacteria) can be found in two classes, with the α -proteobacteria making up the majority of all nitrogen fixing nodule forming bacteria (Willems 2006). However, some β -proteobacteria, *Burkholderia* sp. actually make up a relatively common group of nitrogen fixing, nodule forming symbiotic bacteria (Bournaud, de Faria et al. 2013).

The α -proteobacteria Rhizobia are found in the order Rhizobiales, containing the families, *Bradyrhizobiaceae*, *Brucellaceae*, *Hyphomicrobiaceae*, *Phyllobacteriaceae*, *Rhizobiaceae* (Willems 2006). Most organisms classically referred to as Rhizobia are found in the family Rhizobiaceae in either the genus *Rhizobium* or *Sinorhizobium* (Willems 2006). The family *Phyllobacteriaceae* contains the genus *Mesorhizobium*, and the family *Bradyrhizobiaceae* contains the genus *Bradyrhizobium*. The other families contain the more recently reported nodulating, nitrogen fixing legume symbionts such as species of *Methylobacterium* and *Devosia* (Willems 2006).

The Rhizobiales, including the Rhizobiaceae genus *Rhizobium*, contain non-symbiotic, non-nitrogen fixing members (Willems 2006). In fact, due to precedence the *Sinorhizobium* are now more correctly named *Ensifer*, as a non-symbiotic nitrogen fixing bacterium in the genus was the first isolated and named (Willems 2006). It has been shown that non-symbiotic *Rhizobium* strains, which are closely related to the symbiotic strains persist in soils without legumes. These strains are capable of later picking up the genes required for symbiosis from inoculant strains (Sullivan, Eardly et al. 1996). The most well characterized non-symbiotic *Rhizobium* is the plant pathogen and model organism *Agrobacterium tumefaciens*. Based on phylogenetic studies it has been shown that *Agrobacterium*, *Rhizobium*, and *Allorhizobium* are a single genus, *Rhizobium* (Young, Kuykendall et al. 2001).

1.1.3 Rhizobium genetic make up

The genetic complement of Rhizobia vary significantly; however common amongst genera of Rhizobiaceae is the presence of large plasmids in addition to the core genome.

Sinorhizobium meliloti carries genes necessary for symbiosis on two mega-plasmids. However, *Rhizobium leguminosarum* *bv.phaseoli* contains all the genes necessary for nodulation and nitrogen fixation on a single plasmid (Long 1989). In addition, these plasmids contain genes for metabolic functions (Long 1989), and have been shown to be relevant in competition for nodule occupancy (Oresnik, Pacarynuk et al. 1998). Other Rhizobia contain all the genes for nodulation on the chromosome. An example of this is *Mesorhizobium loti*, which has all the nodulation and nitrogen fixation genes located on the chromosome, in what is termed a symbiotic island (Sullivan and Ronson 1998). It is transmissible amongst *Mesorhizobium* and mobile due to an integrase of the phage P4 family (Sullivan and Ronson 1998).

1.1.4 Sinorhizobium meliloti Rm1021

S. meliloti Rm1021 is an alfalfa symbiont that can enter into symbiotic relationships with *Melilotus* and *Trigonella* as well as the model legume *Medicago truncatula*. Its genome consists of a 3.65-Mb chromosome, plus two mega plasmids, the 1.35-Mb pSymA and 1.68-Mb pSymB (Finan, Weidner et al. 2001, Galibert, Finan et al. 2001). The *S. meliloti* Rm1021, *Medicago truncatula* system is used as one of the main models for invasion, nodule genesis, and nitrogen fixation. This interaction is one of the best studied models for prokaryotic intracellular interactions with eukaryotic systems (Terpolilli, Hood et al. 2012). It serves not only as a model for Rhizobium symbiosis, but for α -proteobacteria-eukaryotic intracellular interactions in general. In addition, *S. meliloti* is well studied in regards to its basic biology.

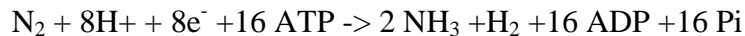
1.1.5 Rhizobium leguminosarum biovar trifoli Rlt100

Rhizobium leguminosarum *bv. trifoli* (initially known as W14-2) is not very well characterized compared to *S. meliloti* Rm1021. It was isolated from a nodule of *Trifolium vesiculosum* Savi in east Texas. It contains four mega plasmids in addition to its chromosome, pRtrW14-2a, pRtrW14-2b, pRtrW14-2c, and pRtrW14-2d. They are 150,170,260,460 MDa respectively (Baldani, Weaver et al. 1992). pRtrW14-2d has been determined to be essential for symbiosis, while pRtrW14-2a is required for full symbiotic efficiency, as its deletion mutants produces fewer nodules (Baldani, Weaver et al. 1992). It does not currently have a sequenced genome; however the rhamnose catabolic region found on pRtrW14-2c is sequenced and deposited in NCBI BLAST database (Richardson, Hynes et al. 2004).

1.2 Rhizobium symbiosis

1.2.1 Biological nitrogen fixation

Atmospheric nitrogen represents the most abundant source of nitrogen. However, it must be converted into a more biologically usable form for assimilation into proteins and nucleic acids. Biological nitrogen fixation is one solution to this problem. N_2 is converted to ammonia, as catalyzed by the prokaryotic enzyme nitrogenase. This is an energetically expensive process, and requires a micro-oxic environment for the stability of nitrogenase (Peters, Fisher et al. 1995).



The majority of nitrogen fixing bacteria are from the α - proteobacteria order Rhizobiales. Forming a symbiotic association with legumes, they fix atmospheric nitrogen in exchange for nutrients and a micro-oxic environment called the root nodule (Terpolilli, Hood et al. 2012).

1.2.2 Root invasion and nodule formation

In the rhizosphere, Rhizobia sense compounds called flavonoids, which are secreted by the host plant. The bacteria respond by inducing the expression of the *nod* genes (Oldroyd, Murray et al. 2011). These are the bacterial genes used to synthesize and export Nod factor, a lipochitooligosaccharide. Fully assembled Nod factor has a four or five residue chitin backbone of β -1,4-linked *N*-acetyl-D glucosamine (GlcNAc) subunits, with an N-

linked acyl tail attached to the non-reducing end (Gage 2004). Decoration of the backbone, and modifications to the lipid tail, lead to species specific variations in Nod factor structure, with some Rhizobial species capable of synthesizing more than one type of Nod Factor (Perret, Staehelin et al. 2000).

Nod factor host specific variation is what determines the host range of a symbiont, as it is the plants ability to recognize Nod factor that is responsible for the initiation of the symbiotic relationship (D’Haeze and Holsters 2002, Oldroyd, Murray et al. 2011). The bacteria also undergo other changes in gene expression. Genes necessary for plant invasion are induced, such as those responsible for resistance to plant released toxins (D’Haeze and Holsters 2002, Gage 2004).

Nod factor causes changes in the host such as root hair deformation, membrane depolarization (Ehrhardt and Atkinson 1992) and calcium oscillation (Ehrhardt, Wais et al. 1996). This results in the initiation of cell division in the root cortex (Gage 2004, Oldroyd, Murray et al. 2011). Nod factor also elicits parallel effects in the root epidermis, where it is important for infection and in the root cortex where it is important for nodule formation.

Nod factor perception in the epidermis leads to two events, one: alterations in gene expression that result from the calcium spiking signal and two: root hair deformation, required for bacterial invasion .The root hair deformation is independent of the calcium spiking signaling (Oldroyd and Downie 2008). Different degrees and specificities of Nod factor recognition are involved at different stages of the rhizobial plant interaction, with root hair deformation being the least stringent perception of Nod factor.

An infection thread is initiated when the bacteria gets trapped inside a deformed root hair that was caused in response to Nod Factor (Esseling, Lhuissier et al. 2003). Invagination

of the cell wall and cell membrane followed by tip growth of the invagination results in initiation of an infection thread. This grows down the inside of the root hair into the epidermal cell. The thread will grow and branch as it enters the nodule primordium cells (Oldroyd, Murray et al. 2011).

More than one bacteria type may become trapped in a curled root hair. However, typically as the infection thread grows, a clonal population is selected by the time the infection thread reaches the nodule (Oldroyd and Downie 2008). Many factors play roles in identifying appropriate bacteria types. High levels of Nod factor are believed to be important in infection thread growth, and the presence of appropriate polysaccharides on the surface of the bacteria are essential for infection (Jones, Kobayashi et al. 2007). They are believed to be vital for the suppression of host defense such as the host production of reactive oxygen species in the infection threads. Bacteria that are not adapted for the symbiosis will not survive, or at least not grow as efficiently as the desired *Rhizobium* species. The cell division of the bacteria takes place at the infection thread tip, and as the infection thread extends, it results in a clonal population entering into the nodule primordium (Oldroyd and Downie 2008, Oldroyd, Murray et al. 2011)(Figure 1.1). Branching of the extending infection thread results in an increased number of places the bacteria can exit the thread into the eventual nodule cells (Gage 2004) .

Bacterial infection can occur even in the absence of nodule organogenesis, and vice versa. However to have an effective symbiosis, a nodule filled with the desired bacteria must result. For this to occur the nodule primordium must be close to the site of bacterial infection.

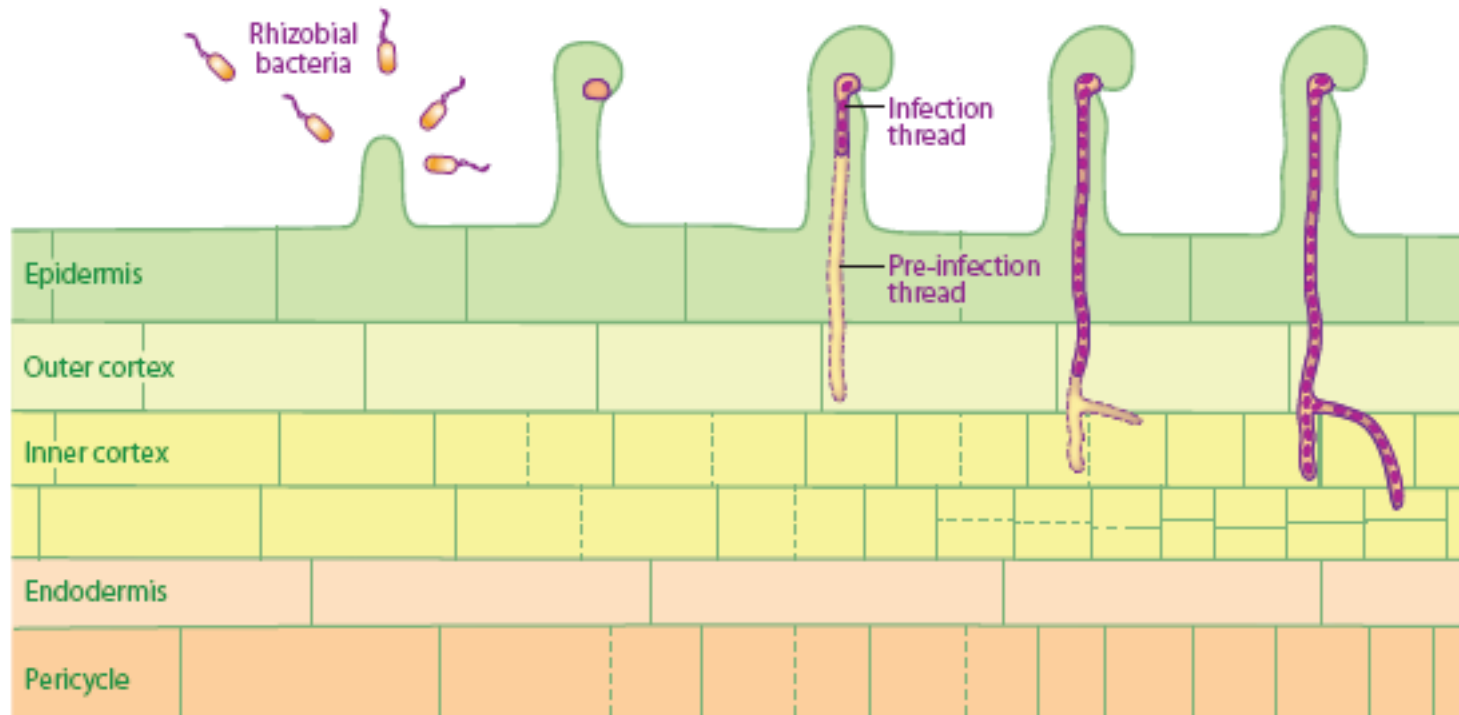


Figure 1.1. Development of the infection thread. Bacteria are entrapped in a curled root hair, and from this site infection threads (ITs) are initiated. The route of the IT is predicted by pre-infection threads that are densely cytoplasmic subdomains with aligned cytoskeleton. ITs progress into the inner cortex where the nodule primordium has formed through a series of cell divisions. From these divided cells, the nodule meristem forms. Taken From: Oldroyd, G. E. D., et al. (2011). "The rules of engagement in the legume-rhizobial symbiosis." Annual review of genetics 45: 119-144.

The nodulation responses in the cortex associated with the formation of the nodule meristem is also initiated in response to Nod factor, and are driven by hormonal changes with cytokinins playing a key role (Gage 2004). This perception of Nod factor is much more stringent than that required for root hair curling (Oldroyd and Downie 2008). Nod factor is perceived by a plant mediated signal transduction. Nod factor perception results in a spike in nuclear calcium levels. This is perceived by a calcium/calmodulin-dependent protein kinase (CCaMK) and transduced to genes that alter cortical cytokinin levels. These alterations in cytokinin levels result in the activation of the genes required for nodule organogenesis (Madsen, Tirichine et al. 2010). As the nodule organogenesis is carried out by plant genes, in rare cases some rhizobia appear to be able to circumvent the need for Nod factor in nodule initiation by directly regulating cytokinin levels (Gage 2004). This is also why nodule organogenesis in the absence of infection can occasionally be seen.

Inside the nodule, the bacteria acquire a plant derived membrane and differentiate into what is termed the bacteroid (Kereszt, Mergaert et al. 2011). This plant derived membrane containing the bacteroid is often referred to as the symbiosome (Kereszt, Mergaert et al. 2011). In the natural environment there is inevitably a competition for nodule occupancy between strains of *Rhizobium* capable of nodulating the same plant.

1.2.2.1 Determinate nodules

There are two main types of nodules associated with the Rhizobia-plant symbiosis, determinate and indeterminate nodules. The type of nodule development is determined by plant factors (Kereszt, Mergaert et al. 2011). The main feature of determinant nodules is that they lose meristematic activity. This means that the growth of the nodule is due to cell

expansion and thus nodules typically have a spherical shape. These types of nodule are associated with *Glycine* (soybean), *phaseolus*, and *vigna* plants (Kereszt, Mergaert et al. 2011). Inside these nodules, the bacteroids are similar to free-living bacteria, and the symbiosomes contain multiple cells (Bisseling, vanden Bos et al. 1977).

1.2.2.2. Indeterminate nodules

Indeterminate nodules are found in the inverted repeat-lacking clade (IRLC) of legumes (Kereszt, Mergaert et al. 2011). This includes the legume hosts of *R. leguminosarum* Rlt100, and *S. meliloti* Rm1021. Indeterminate nodules are defined by an active apical meristem. New cells are formed constantly resulting in a cylindrical shape (Terpolilli, Hood et al. 2012). This also means that this type of nodule develops zones that are in different stages of development. Zone I contains the active meristem that grows and will later differentiate into the other zones. Zone II contains infection threads full of bacteria. In the inter-zone between zone II and III bacteria cells begin to elongate and terminally differentiate. Zone III is the active nitrogen fixation zone filled with differentiated bacteroids. Zone IV is populated with plant cells and bacteroids that are being degraded (Terpolilli, Hood et al. 2012).

1.2.2.2.1 Symbiosomes and Bacteroids in indeterminate nodules.

Inside indeterminate nodules, bacteria differentiate and synthesize proteins required for nitrogen fixation, and maintenance of the mutualistic partnership (Gage 2004). As the bacteria enter the nodule they pick up an additional plant-derived membrane. This means that the symbiosome consists of two membranes, one membrane of bacterial origin, plus the one

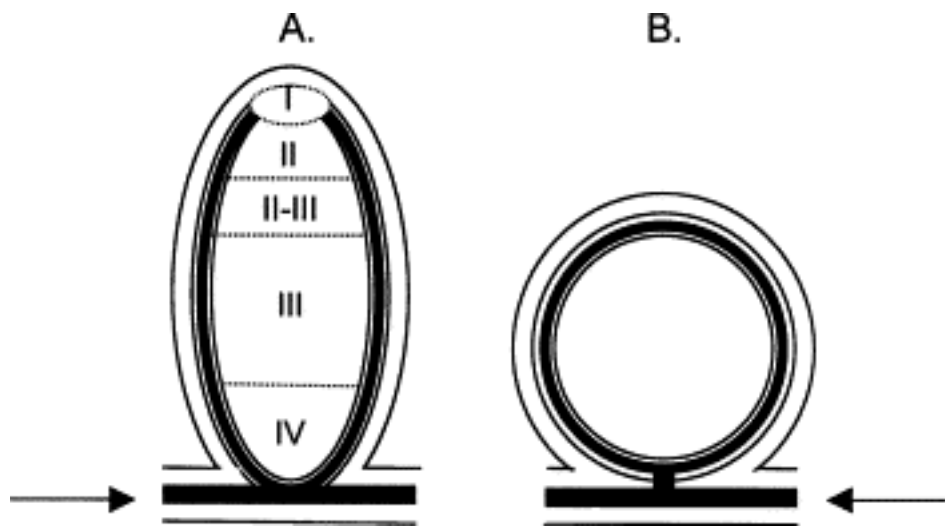


Figure 1.2. The nodule structure of indeterminate (A) and determinate (B) nodule types. The different zones present in indeterminate nodules are indicated. Zone I - the active meristem. Zone II - contains infection threads full of bacteria. Zone III - Actively fixes nitrogen and is filled with differentiated bacteroids. Zone IV - contains plant cells and bacteroids that are dying and being degraded. Taken From: Luyten, E. and J. Vanderleyden (2000). "Survey of genes identified in *Sinorhizobium meliloti* spp., necessary for the development of an efficient symbiosis." European Journal of Soil Biology 36 (1): 1-26.

from the plant (Oldroyd, Murray et al. 2011). The plant-derived membrane has its orientation flipped so its outside face is turned toward the bacteria. This arrangement requires transport across these two membranes for any nutrient exchange between plant and bacteroid to take place. Here in the symbiosome, the bacteria differentiate into the bacteriod (Terpolilli, Hood et al. 2012).

Indeterminate nodules contain symbiosomes with only a single bacteriod. However, these bacteriods are swollen and pleomorphic in shape due to genome multiplication without cell division (Oono and Denison 2010). For example, *S. meliloti* bacteroids found in *Medicago truncatula* nodules can have a chromosome count of up to $24n$ (Terpolilli, Hood et al. 2012). This results in the bacteria being in a terminally differentiated state (Mergaert, Uchiumi et al. 2006).

It was initially unknown if this differentiation was controlled by the plant or the bacteria. To address this, plasmid borne Nod genes were used to extend the host range of *Rhizobium leguminosarum* *bv. viciae* to include *Lotus japonicus*. *Lotus japonicus* normally develops symbiosomes, which contain multiple bacteria cells that are morphologically similar to the free living bacteria. However, *Rhizobium leguminosarum* *bv. viciae* in its native host is pleomorphic with genome multiplication.

When *Rhizobium leguminosarum* *bv. viciae* was modified to nodulate *Lotus japonicas*, this resulted in bacteriods with chromosomes of $1-2n$ and morphology more like that of the free living bacteria. This indicates it is the plant controlling and altering the bacteria cell cycle (Mergaert, Uchiumi et al. 2006, Terpolilli, Hood et al. 2012).

These changes are believed to be mediated by the nodule specific cysteine rich antimicrobial like peptides termed NCRs. NCRs require the bacterially encoded BacA to

function (Oldroyd, Murray et al. 2011) as well as the plant DNF1 protein, which is a component of the plant's signal peptidase complex. In the absence of the DNF1 protein, cells do not differentiate (and thus do not fix N₂). In a DNF1 mutant NCRs are localized to the endoplasmic reticulum suggesting that DNF1 is required for targeting and export of the NCRs to the symbiosome (Wang, Griffiths et al. 2010). *In vitro* *S. meliloti* studies support this, as NCR peptides added to free living cells induced bacteroid-like characteristics, including inducing membrane permeabilization, and endoreduplication (Terpolilli, Hood et al. 2012). These cells were also no longer viable.

Bacteroids behave differently than free-living cells, shutting down most aspects of cell growth and division. Synthesis of ribosomal proteins, and nucleic acids, are both turned down. Other genes, such as those required for the TCA cycle are up regulated (Oldroyd, Murray et al. 2011). Bacteroids and the plant-derived peribacteroid membrane tightly regulate the exchange of metabolites (Udvardi and Poole 2013).

1.3 Rhamnose catabolism

1.3.1 Plasmid curing experiments in Rhizobium leguminosarum

Plasmid curing experiments were carried out in various species of *Rhizobium*. Many of these plasmids contained genes not required for symbiosis. In an effort to determine if some of these plasmid borne genes were relevant to symbiotic fitness, plasmid-cured strains of *Rhizobium* were competed against wild-type strains. Some plasmid-cured derivatives were found to be less competitive for nodule occupancy than their wild-type counterparts (Hynes 1990, Hynes and O'Connell 1990, Brom, De Los Santos et al. 1992). Plasmid-cured derivatives also failed to achieve as high total population numbers as the wild-type when they were co-inoculated with the wild-type (Moënne-Loccoz and Weaver 1995).

Plasmid cured strains of *R. leguminosarum* Rlt100 were tested for growth defects on many carbon sources, and examined for other phenotypic effects. The ability to use rhamnose or sorbitol as sole a carbon source was found to be conferred by the mega-plasmid pRtrW14-2c. The genes required to utilize adonitol were found localized to the mega-plasmid pRtrW14-2b (Baldani, Weaver et al. 1992).

Tn5 insertions in the regions required for the utilization of rhamnose, sorbitol and adonitol were isolated. In an attempt to determine if it was the inability to utilize these sugars that was responsible for the impaired competition phenotype found in the plasmid cured derivatives, the Tn5 mutant strains were competed against the wild-type Rlt100 strains. The inability to use sorbitol or adonitol as a sole carbon source was found to have no effect on competition for nodule occupancy. However, three independent strains unable to use

rhamnose as a sole carbon source were found to be less competitive for nodule occupancy than the wild-type parent strain (Oresnik, Pacarynuk et al. 1998). The genes required for the utilization of rhamnose were also shown to be induced by clover root exudates (Oresnik, Pacarynuk et al. 1998). Rhamnose catabolism in Rlt100 was further studied in an effort to better understand its relationship to competition for nodule occupancy.

1.3.2 Canonical Rhamnose catabolism in bacteria

Rhamnose catabolism in bacteria has been characterized to varying degrees in many bacterial models (Rodionova, Li et al. 2013). The first bacterial rhamnose catabolic pathway elucidated was that of *E. coli*. This is the best characterized bacterial model for rhamnose catabolism, with all the main catabolic enzymes isolated and assayed for activity. The products of these reactions have also been purified to confirm the enzymes functions (Chiu and Feingold 1964, Power 1967). The regulation of the *E. coli* rhamnose regulon has also been extensively studied (Tobin and Schleif 1990) (Egan and Schleif 1994).

Although operon structure and regulation, as well as the phylogeny of the catabolic enzymes vary amongst the systems (Rodionova, Li et al. 2013), the canonical phosphorylated *E. coli* pathway seems to be shared amongst most characterized systems. The only exception is the alternate non-phosphorylated catabolic pathway found in *Azotobacter vinelandii* and some *Sphingomonas* sp. (Watanabe, Saimura et al. 2008, Watanabe and Makino 2009). In *E. coli*, the rhamnose regulon consists of the transporter gene *rhaT*, the two partially overlapping regulator genes, *rhaR* and *rhaS*, and the genes encoding the metabolic enzymes *rhaA*, *rhaB*, and *rhaD* (Power 1967, Moralejo, Egan et al. 1993). A recent comparative

genomics study reveals many different rhamnose catabolic operon structures are found in bacteria, often containing non-orthologous variants of proteins (Rodionova, Li et al. 2013).

In the canonical phosphorylated rhamnose pathway, L-rhamnose is first isomerized to L-rhamnulose. L-rhamnulose is converted to L-rhamnulose-1-phosphate, which is split into L-lactaldehyde and dihydroxyacetone phosphate (Chiu and Feingold 1964, Power 1967). L-lactaldehyde is either oxidized to L-lactate aerobically, or reduced to L-1,2-propanediol in anaerobic conditions (Baldoma and Aguilar 1988) (Figure 1.3).

1.3.2.1 Transport of Rhamnose by the bacterial cell

Rhamnose can be found in nature as a part of complex pectin polysaccharides. Enzymes characterized to date that are able to degrade rhamnogalacturonans and other L-Rha-containing polysaccharides are typically found as extracellular enzymes. In addition to transporters specific for L-rhamnose, many bacteria contain specific transporters for oligosaccharides of rhamnose, or the unsaturated rhamnogalacturonides (Rodionova, Li et al. 2013). The rhamnose oligosaccharides are broken down via α -L-rhamnosidases (RhmA and RamA) in various saprophytic species (Rodionova, Li et al. 2013). Rhamnogalacturonides are broken down by hydrolases (ex. RhiN) to release the constituent rhamnose. However, none of these enzymes are required for growth directly on rhamnose as a sole carbon source, and these enzymes are not encoded by genes found in the same operon that encodes the canonical rhamnose catabolic enzymes (Rodionova, Li et al. 2013).

In *E. coli*, rhamnose transport is dependent on the gene product of *rhaT*. RhaT is a H^+ symporter that catalyzes the uptake of L-mannose, L-lyxose, and L-rhamnose with the

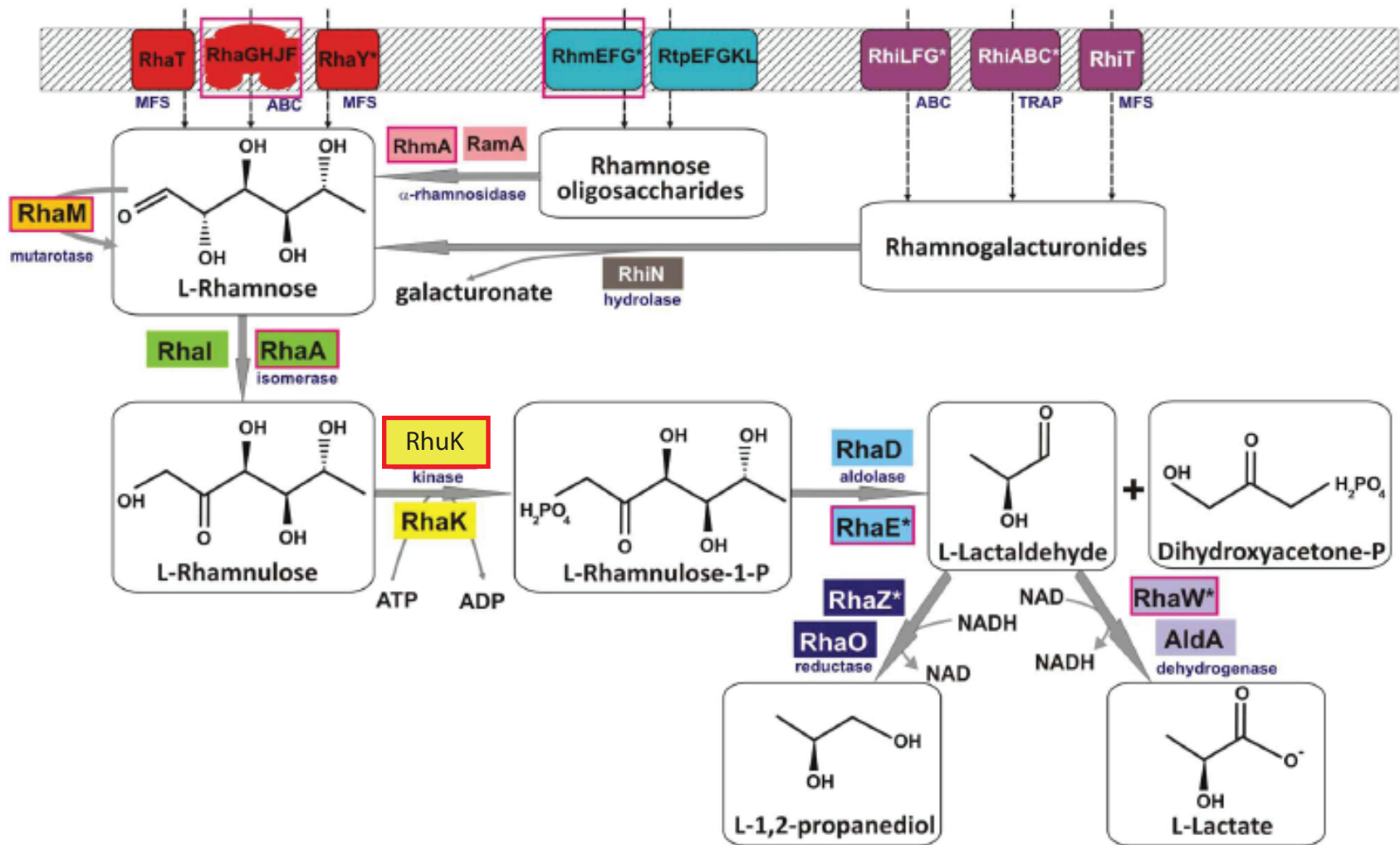


Figure 1.3. Summary of the canonical l-rhamnose utilization pathway in bacteria. Solid gray arrows indicate enzymatic reactions, broken arrows indicate transport. Enzyme classes and families of transporters are shown in blue subscript. Tentatively predicted functional roles are marked by asterisks. Taken from: Rodionova, I. A., et al. (2013). "Comparative genomics and functional analysis of rhamnose catabolic pathways and regulons in bacteria." Front Microbiol. 4.

influx of protons (Tate, Muiry et al. 1992). However, comparative genomics reveals that many types of transport systems are used for the transport of rhamnose in various bacteria. RhaY proteins identified in *Bacillales* and *Actinomycetales* genomes are from the Major Facilitator Superfamily (Rodionova, Li et al. 2013), and ABC transporter systems have been identified in *Chloroflexales*, *Actinomycetales*, and *Enterobacteriales* genomes (Rodionova, Li et al. 2013).

1.3.2.2 Isomerization of rhamnose to rhamnulose

The first catabolic step in the canonical phosphorylated rhamnose pathway is the isomerization of the sugar. In *E. coli* this is carried out by the well characterized enzyme RhaA. Its structure and mechanism have previously been determined (Korndörfer, Fessner et al. 2000). The L-rhamnose isomerase from *Pseudomonas stutzeri* has also been well characterized. It represents another common class of L-rhamnose isomerase that only shares about 17% sequence identity with RhaA (Rodionova, Li et al. 2013).

Based on alignments with the *E. coli* RhaI, it appears that the same residues are involved in the active site of both the *E. coli* and the *P. stutzeri* RhaI (Leang, Takada et al. 2004). However, the *P. stutzeri* class of isomerase is phylogenetically distinct from the one found in the *E. coli* rhamnose regulon (Leang, Takada et al. 2004).

Unlike the *E. coli* L-RhaI, L-RhaI from *Pseudomonas stutzeri* has been demonstrated to catalyze the conversion between aldoses and ketoses for a broad range of substrates, including common and rare sugars in both the D and L forms (Leang, Takada et al. 2004). However, L-rhamnose was the preferred substrate, with a K_m of 11 mM, and a V_{max} of 240

U/mg. L-mannose, L-lyxose, D-ribose, and D-allose were also isomerized at rates well above background (Leang, Takada et al. 2004).

In an effort to gain insight into the broad substrate specificity of *P. stutzeri* L-RhaI, multiple crystal structures have been solved at resolutions between 1.97 Å and 2.0 Å. Structures alone or complexed with either L-rhamnose or D-allose are available (Yoshida, Yamada et al. 2007). L-rhamnose and D-allose were chosen as they differ at C4 and C5. Like the *E. coli* L-RhaI and many D-xylose isomerases, *P. stutzeri* RhaI is found as a homo tetramer. The β 1- α 1 loop (Gly60-Arg76) of *P. stutzeri* L-RhaI is involved in the substrate binding of a neighbouring molecule (Yoshida, Yamada et al. 2007).

The region of the *P. stutzeri* RhaI involved in substrate-binding with the sugar at C1, 2, and 3 is very similar to the equivalent region of *E. coli* L-RhaI. However, at the binding sites for C4, 5, and 6 the *E. coli* L-RhaI β 1- α 1 loop creates a hydrophobic pocket leading to the strict recognition of L-rhamnose. *P. stutzeri* L-RhaI has no corresponding hydrophobic pocket leading to loose substrate recognition at the 4, 5, and 6 positions (Yoshida, Yamada et al. 2007).

1.3.2.3 Phosphorylation of rhamnulose to rhamnulose-1-P

The second major step in the catabolism of rhamnose is the phosphorylation of the sugar. This step is catalysed by a carbohydrate kinase. The most well characterized rhamnulose kinase is the *E.coli* RhuK encoded by *rhaB*. Like all *E. coli* rhamnose catabolic enzymes it has been purified, assayed *in vitro* for activity, and had its products purified to confirm the enzymes activity (Power 1967). This enzyme has been shown to phosphorylate rhamnose and fructose (Grueninger and Schulz 2006).

Carbohydrate kinases belong to the hexokinase-hsp70-actin superfamily (Hurley 1996). Carbohydrate kinases are typically highly conserved in regards to functional domains, and even more so in overall structure (Grueninger and Schulz 2006, Anderson, DeLaBarre et al. 2007, Di Luccio, Petschacher et al. 2007). All members of this superfamily share certain characteristics. They contain 5-stranded beta sheets in the N and C terminal domains, with an ATP-binding site deep inside the cleft between these two domains. This is where the γ -phosphoryl group is transferred to a substrate bound deeper in the cleft (Grueninger and Schulz 2006).

Most carbohydrate kinases, contain the P-loop ATPase. The P-loop is one of the most common protein folds, and it is estimated to be found in 18% of all gene products (Leipe, Wolf et al. 2002). P-loop ATPases consist of a repeating α - β unit where the β strands form a central parallel β -sheet surrounded on both sides by α -helices. This results in the presence of a flexible loop occurring between a strand and a helix (Leipe, Wolf et al. 2002). The P-loop ATPase fold is characterized by the N-terminal Walker A motif defined by the amino acid sequence GXXXXGK (Walker, Saraste et al. 1982),

Most carbohydrate kinase are found as dimers, tetramers, or even higher order oligomers (Grueninger and Schulz 2006, Anderson, DeLaBarre et al. 2007, Di Luccio, Petschacher et al. 2007). Xylulose kinase and Glycerol kinase are two of the most well characterized carbohydrate kinases. They are both active as dimers, with β -strands 18 and 19 making up the dimerization domain. Glycerol kinase can be found as a tetramer, but this is a catalytically inactive form (Anderson, DeLaBarre et al. 2007, Di Luccio, Petschacher et al. 2007). Glycerol kinase also contains a regulatory domain that binds fructose 1,6-bisphosphate along the plane where the tetramer forms (Anderson, DeLaBarre et al. 2007).

This stabilizes the catalytically inactive structure, and is used to regulate activity based on substrate availability. No homologous regulatory domain is found on the *E. coli* xylulose kinase (Di Luccio, Petschacher et al. 2007) (Figure 1.4).

RhuK, like all members of this superfamily consists of two large domains separated by a deep cleft. The N-terminal region is composed of residues 1–236. The C-terminal domain contains residues 237–489 (Grueninger and Schulz 2006). However, RhuK differs from these other kinases in two key ways. 1) RhuK does not contain a P-loop, instead having an alternate ATPase domain and; 2) it is found as a monomer in solution not an oligomer. RhuK appears to contain a residue capable of a disulfide bridge at the position homologous to where the *E. coli* glycerol kinase is regulated by effector binding (Grueninger and Schulz 2006). However, RhuKs activity does not appear to be regulated by disulfide bond formation as the same enzyme activity is observed in both reducing and oxidizing conditions (Grueninger and Schulz 2006).

Comparative genomics have revealed many bacteria contain another class of predicted rhamnulose kinases. They typically share more amino acid identity with *E. coli* glycerol kinase and xylulose kinase than with RhuK, and contain a P-loop for ATP binding (Rodionova, Li et al. 2013).

1.3.2.4 Metabolism of rhamnulose-1-P

In *E. coli*, RhaD, a rhamnulose-1-phosphate aldolase, generates L-lactaldehyde and dihydroxyacetone phosphate from rhamnulose-1-phosphate. Dihydroxyacetone phosphate enters central metabolism, while L-lactaldehyde has two possible fates. In aerobic conditions

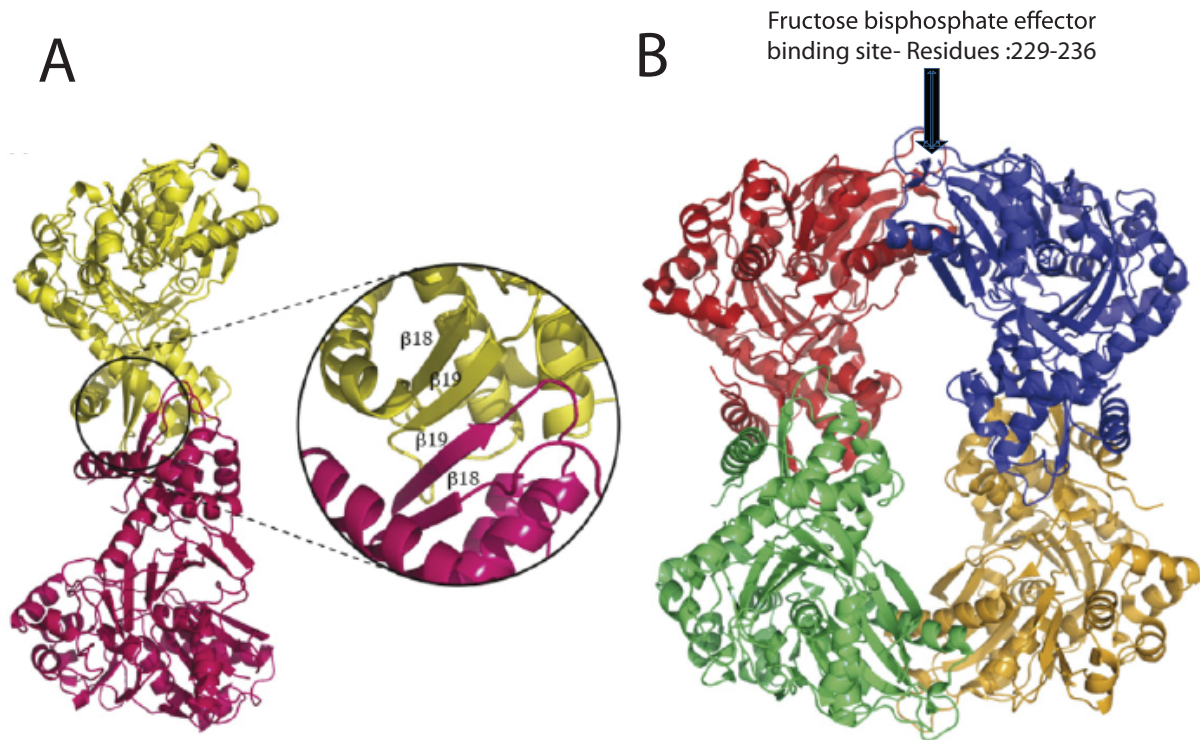


Figure 1.4. A) The oligomeric structure of *E. coli* xylulose kinase, with a magnified view of the dimeric interface between β -sheets responsible for mediating the dimerization. B) The allosterically-inhibited *E. coli* glycerol kinase tetramer. The arrow indicates the location of effector binding responsible for stabilizing the inactive form. Each kinase monomer is illustrated in a different colour. Modified from: Di Luccio, E., et al. (2007). "Structural and kinetic studies of induced fit of xylulose kinase from *Escherichia coli*." J. Mol. Biol. 365

L-lactaldehyde is oxidized to L-lactate by lactaldehyde dehydrogenase (AldA), which appears to be induced by the presence of lactaldehyde. In anaerobic conditions L-lactaldehyde is reduced to L-1,2-propanediol by the enzyme propanediol oxidoreductase (RhaO) (Baldoma and Aguilar 1988, Rodionova, Li et al. 2013). In most organisms, similar biochemical reactions are carried out by various orthologous proteins predicted to have the same functions as the *E. coli* enzymes (Rodionova, Li et al. 2013). However, a novel bi-functional enzyme RhaEW was identified in *B. subtilis* and *C. aurantiacus*. It is a L-rhamnulose-phosphate aldolase (RhaE) and a L-lactaldehyde dehydrogenase (RhaW) fused together. These enzymes are not homologous to previously characterized L-rhamnose catabolic enzymes (Rodionova, Li et al. 2013).

1.3.2.5 Regulation of L-rhamnose catabolic regulons.

In *E. coli*, genes constituting the rhamnose regulon are controlled by RhaR, an AraC type positive regulator that responds to rhamnose and positively regulates transcription of *rhaR* and *rhaS* (Tobin and Schleif 1990). The regulator RhaS in turn regulates the transcription of the structural genes *rhaA,B,D* and *rhaT* (Egan and Schleif 1994). However, other types of regulators are found to regulate the various types of rhamnose operons that have been identified to date. In addition to the well characterized AraC type positive regulators found in the *E. coli* rhamnose regulon, predicted DeoR, and LacI type regulators have been identified in various rhamnose catabolic operons. Consensus sequences for putative bindings sites of these regulators have been predicted (Rodionova, Li et al. 2013).

1.3.3 A locus required for rhamnose catabolism in R. leguminosarum Rlt100

A transposon mutagenesis strategy was used to identify a region from Rlt100 required for growth on rhamnose as a sole carbon source. The rhamnose catabolic region was sequenced, and was shown to span 10,959 bp (Richardson, Hynes et al. 2004). Based on sequence analysis, the region appears to contain the genes required for the canonical phosphorylated rhamnose pathway. It contains: a putative dehydrogenase/aldolase, a predicted rhamnose isomerase, and a putative sugar kinase (Richardson, Hynes et al. 2004). Genes for an ABC-type transporter system, and a DeoR type negative regulator are also found in the region. The genes are arranged in two divergent transcripts. One transcript contains *rhaD* and *rhaI*. The other transcript contains *rhaR* encoding the putative regulator, *rhaSTPQ*, the components of the predicted ABC-transporter; *rhaU* a predicted mutarotase, and *rhaK* a predicted sugar kinase. The region was also shown to be inducible by rhamnose (Richardson, Hynes et al. 2004) (Figure 1.5).

1.3.3.1 rhaI encodes a P. stutzeri type rhamnose isomerase

RhaI is encoded by a 1290-bp gene coding for a 430 amino acid protein. A transposon mutagenesis did not reveal any Tn5 insertions beyond *rhaI* that lead to an inability to use rhamnose as a sole carbon source. Sequencing did not show any predicted ORFs past this gene. This suggests that *rhaI* is the final gene in its transcript (Richardson, Hynes et al. 2004). Sequencing reveals a gap of 193 bp between *rhaI* and the preceding ORF, *rhaD*, in its transcript.

RhaI appears to be a *P. stutzeri* class rhamnose isomerase, and shares high amino acid conservation with the *P. stutzeri* rhamnose isomerase but very little with *E. coli* RhaA (Rodionova, Li et al. 2013) (Figure 1.6).

1.3.3.2 rhaD encodes a predicted dehydrogenase/ aldolase

The ORF *rhaD* is predicted to encode a 698-amino-acid protein with two predicted conserved domains and a putative NAD⁺ binding site. The first domain is predicted to be a short- chain dehydrogenase, and the other a type II aldolase domain . Based on comparative genomics it has been predicted to be a RhaEW class protein, and to have the same biochemical function as RhaEW (Rodionova, Li et al. 2013). However, no biochemical data are available to confirm this function.

1.3.3.3 rhaR encodes a DeoR type negative regulator

RhaR is encoded by the first gene in the *rhaRTSPQUK* transcript. It is predicted that the two divergent promoters regulating the rhamnose operon are located between *rhaR* and *rhaD*. RhaR is predicted to be a 270 amino acid protein, that based on sequence analysis has been identified as a DeoR-type negative regulator (Richardson, Hynes et al. 2004).

The DeoR-type negative regulators are a class of prokaryotic repressors named for the *E. coli* protein DeoR, known to be a repressor of the *deo* operon which encodes nucleotide and deoxyribonucleotide catabolic enzymes (Valentin-Hansen, Højrup et al. 1985). This class of regulator, found in diverse types of bacteria, are typically regulators of sugar and nucleoside metabolic operons.

The DeoR type negative regulator is defined by the DeoR type DNA-binding domain, containing a standard helix-turn-helix (HTH) DNA binding domain of about 50-60 amino acids (Valentin-Hansen, Højrup et al. 1985). This domain is typically N-terminal, with the C-terminal region containing the domains for oligomerization and effector-binding (Valentin-Hansen, Højrup et al. 1985). Some DeoR-like proteins are capable of binding several operators at a time leading to DNA looping (Valentin-Hansen, Albrechtsen et al. 1986, Amouyal, Mortensen et al. 1989).

The effector molecules for DeoR-like regulators are generally phosphorylated intermediates from the metabolic pathway that they regulate (Barriere, Veiga-da-Cunha et al. 2005, Engels and Wendisch 2007, Garces, Fernández et al. 2008). DeoR itself functions as an octamer (Mortensen, Dandanell et al. 1989), though other members of this family have been shown to work as tetramers (Ray and Larson 2004).

The effector molecules for RhaR are unknown; however, putative DNA binding sites for RhaR have been predicted (Rodionova, Li et al. 2013).

1.3.3.4 rhaSTPQ encodes the components of a rhamnose ABC transporter

The genes *rhaSTPQ* encode components of an ABC-type transporter. Between the regulator *rhaR* and *rhaS* there is a 46-bp intergenic space, and there is a 76-bp intergenic region between *rhaS* and *rhaT*. RhaS is the putative substrate binding protein of the ABC transporter, with RhaP and RhaQ making up the membrane spanning permease. RhaT is a CUT 2 type ABC ATPase. The K_m for this transporter was determined to be $13 \pm 2 \mu\text{M}$ by using a tritiated rhamnose uptake assay (Richardson, Hynes et al. 2004). The transporter appears to be specific for rhamnose as other L-sugars added even in 100 fold excess do not

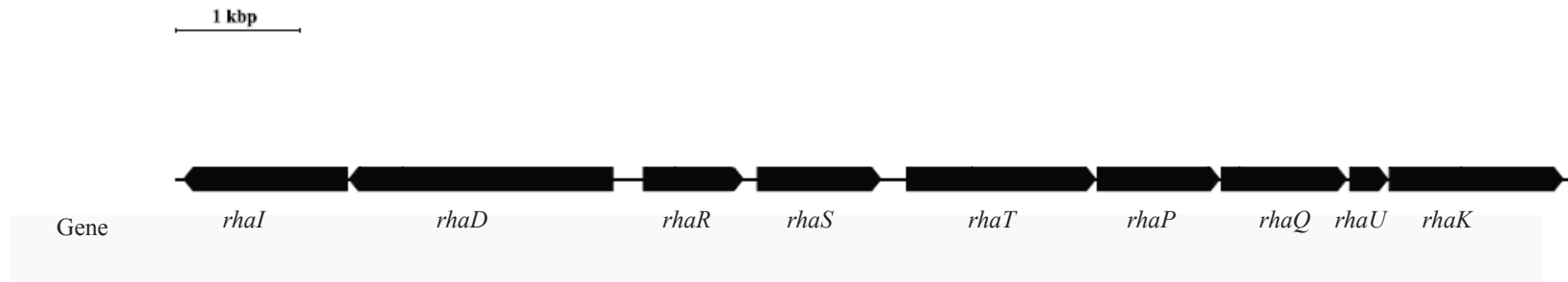


Figure 1.5 The rhamnose region of *R. leguminosarum* *bv. trifolii* strain Rlt100. The solid horizontal line is the Rlt100 genomic DNA. Arrows indicate ORFs. Gene names are under each respective ORF

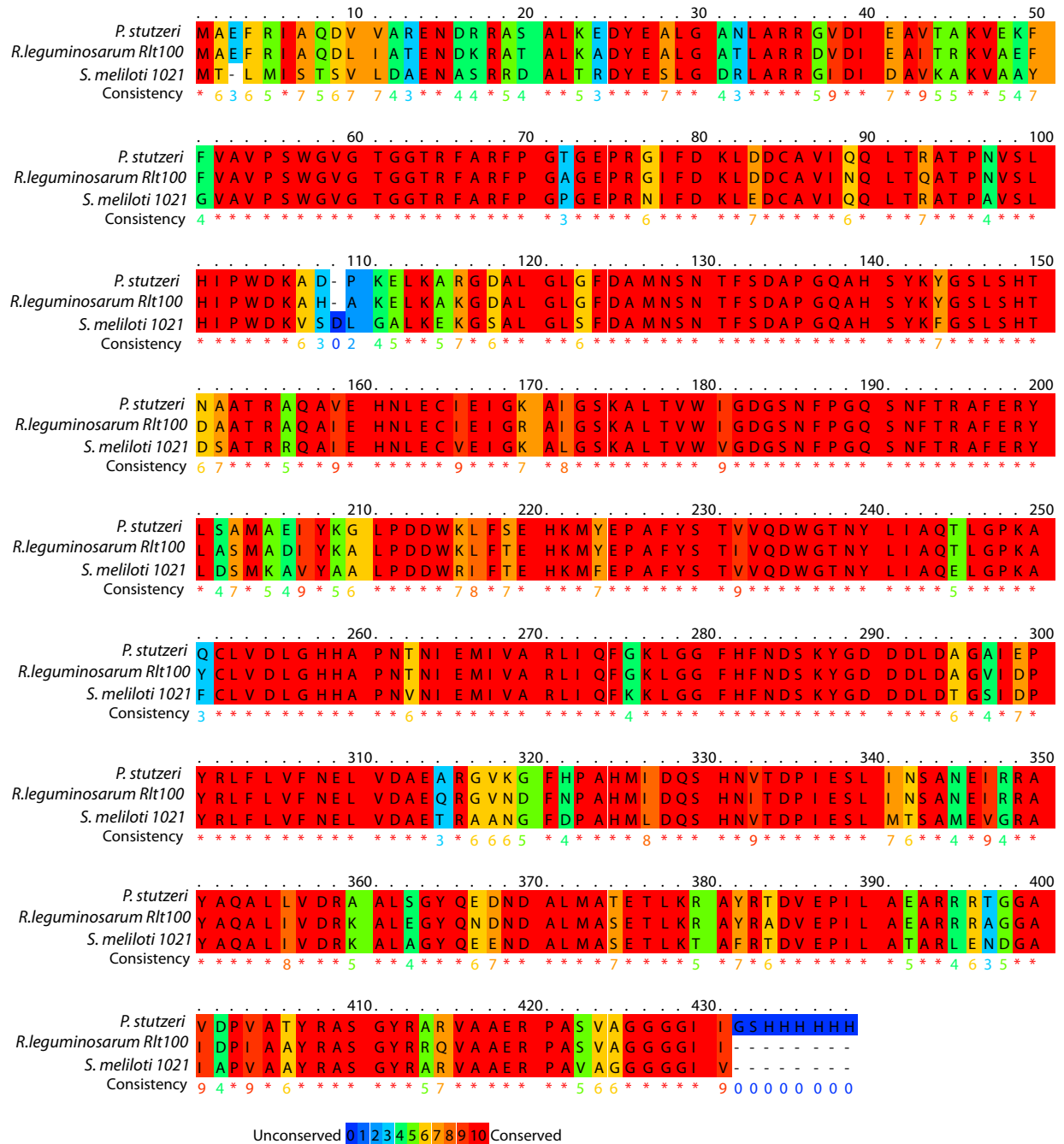


Figure 1.6. The predicted amino acid conservation of the RhaI proteins from *R. leguminosarum* Rlt100, *S. meliloti* Rm1021, and *P. stutzeri*. Analysis was performed by PRALINE, with scoring ranging from 0 (least conserved) to 10 (invariant and denoted by an *)

competitively inhibit the transport of labelled rhamnose (Richardson, Hynes et al. 2004). The transporter encoded by *rhaSTPQ* also appears to be the exclusive rhamnose transporter, as a mutation to any of the transporter genes abolishes the uptake of labelled rhamnose (Richardson, Hynes et al. 2004).

1.3.3.5 rhaU encodes a mutarotase

The gene *rhaU* was first identified as a putative open reading frame, with a start site that was predicted to overlap the stop codon of *rhaQ*. When it was initially identified, its function was unknown (Richardson, Hynes et al. 2004). Strains of *R. leguminosarum* with a mutation to *rhaU* were found to have a slow growth phenotype using rhamnose as a sole carbon source. This phenotype was determined not to be due to impaired rhamnose transport (Richardson, Carpena et al. 2008). To further determine the role of RhaU, a crystal structure was determined and refined to a 1.6 Å resolution. RhaU has a dimeric arrangement, and a structure that is similar to the *E. coli* YiiL. YiiL has been identified as an L-rhamnose mutarotase catalyzing inter-conversion of the anomers β- and α-L-Rhamnose (Ryu, Kim et al. 2005). Consistent with RhaU being a mutarotase, RhaU crystals soaked with L-rhamnose contain β-L-rhamnose in the active site (Richardson, Carpena et al. 2008). This enzyme is only needed for the catabolism of L-rhamnose if it is found in the wrong anomeric form. Typically, rhamnose is found to in a 56:44 ratio of α L-rhamnose: β L-rhamnose as determined by NMR (Ryu, Kim et al. 2005).

1.3.3.6 *rhaK* encodes a sugar kinase

The final gene in the rhamnose operon is *rhaK*. It is predicted to encode a 460 amino acid protein, with its ORF overlapping the stop codon of RhaU (Richardson, Hynes et al. 2004). RhaK is a member of the FGGY family of carbohydrate kinases, and contains multiple conserved domains, most notably the P-loop Nucleotide-Binding Domain (Hurley 1996). Regions necessary for Mg⁺⁺ binding, substrate binding, ATP-binding, and putative active sites have been predicted based on homology to similar carbohydrate kinases (Hurley 1996 , Anderson, DeLaBarre et al. 2007, Di Luccio, Petschacher et al. 2007). Although RhuK from *E. coli* is well characterized, it is not the best model for comparison. BLAST analysis reveals RhaK is more closely related to the *E. coli* xylulose kinase and glycerol kinase than the *E. coli* RhuK.

In vitro assays for RhaKs activity, show that RhaK has activity directly phosphorylating rhamnose (Richardson and Oresnik 2007). This is unlike all other previously identified carbohydrate kinases involved in rhamnose catabolism, as they phosphorylate rhamnulose (Rodionova, Li et al. 2013).

1.3.3.7 Basic regulation of the *Rlt100* Rhamnose operon

The *rhaDI* transcript appears to be induced by rhamnose, and repressed by glucose (Oresnik, Pacarynuk et al. 1998). Likewise, the divergent transcript is also induced by the presence of rhamnose, where as transcription of *rhaP* and *rhaQ* are practically non-existent in non-inducing conditions (Richardson, Hynes et al. 2004). *rhaR* seems to have high transcription levels in non-inducing conditions, but also shows induction in the presence of rhamnose. When *rhaR* was over-expressed on a plasmid in a *rhaR25* (Tn5-B20 fusion) chromosomal background, a 17 fold decrease in expression of *rhaR25* was measured,

consistent with it being a negative regulator of the *rhaRSTPQUK* transcript. However, in the presence of rhamnose, full induction of *rhaR25* was measured even in the presence of constitutively expressed *rhaR* (Richardson, Hynes et al. 2004).

Despite being in the same transcript as *rhaR*, transcription beyond *rhaS* is minimal in non-inducing conditions. This is hypothesized to be due to the predicted hairpin loop between *rhaS* and *rhaT* (Richardson, Hynes et al. 2004).

1.3.3.8 R. leguminosarum appears to have a rhamnose catabolic pathway that is different than the canonical phosphorylated rhamnose pathway

E. coli rhaD mutants show a conditional growth phenotype in the presence of rhamnose when growing on a second (non-repressing) carbon source. This is likely due to a buildup of a toxic phosphorylated rhamnose catabolic intermediate (Power 1967). Similar effects have also been shown previously in the early galactose work (Sundararajan, Rapin et al. 1962).

In *R. leguminosarum*, *rhaD/I* or *rhaI* mutants grown on rhamnose and glycerol media show a conditional growth phenotype. This is hypothesized to also be due to a build up of a phosphorylated intermediate. In a *rhaK*, *rhaD/I* double mutant this conditional growth phenotype is no longer observed. This shows the conditional growth phenotype is dependent on a functional RhaK, and a lesion in the pathway at either RhaD or RhaI. This suggests that RhaK is responsible for the first biochemical step, and that mutants of *rhaD* or *rhaI* both prevent further catabolism of an already phosphorylated intermediate product (Richardson, Hynes et al. 2004). The *E. coli* pathway shows a similar phenotype in only *rhaD* mutants, because its isomerase acts directly on rhamnose prior to the kinase in the pathway (Power

1967). This indicates the rhamnose catabolic pathway in *R. leguminosarum* is different from that of *E. coli*, with the phosphorylation of the sugar as the first catabolic step. This is also supported by the fact that RhaK has an activity directly phosphorylating rhamnose that has been measured *in vitro* (Richardson and Oresnik 2007). This however leaves the enzymatic role of RhaD, and RhaI unknown, and the catabolic pathway after RhaK unknown.

1.3.3.9 RhaK is required for the transport of Rhamnose by Rhizobium leguminosarum.

It has been shown that the ABC transporter encoded by *rhaSTPQ* is necessary for the transport of rhamnose (Richardson, Hynes et al. 2004). However, Rlt100 is unable to transport rhamnose in the absence of the kinase RhaK (Richardson, Hynes et al. 2004). This affect is not seen in the *rhaD* or *rhaI* mutation backgrounds (Richardson, Hynes et al. 2004). RhaK itself does not appear to confer the ability to transport rhamnose in the absence of the transporter, and it has been shown that in the absence of the kinase the ABC transporter components are still being transcribed and translated. The permease component RhaQ has also been shown to correctly localize to the cell membrane (Richardson and Oresnik 2007). This indicates that the inability to transport rhamnose must be due to the absence of the RhaK protein, and is not simply a matter of regulation.

There is also some evidence that it is the biologically active form of RhaK that is required for the transport of rhamnose. K16D and K16M mutations to the RhaK P-loop are predicted to abolish ATP dependent kinase activity. They were observed to not only abolish phosphorylation of the sugar, but also abolished the uptake of labelled rhamnose (Richardson and Oresnik 2007). If rhamnose is not in the correct anomeric form, then RhaU must act on rhamnose prior to the activity of RhaK (Richardson, Carpena et al. 2008). This means that

although RhaK has kinase activity, and is involved in transport, its two functions are temporally separated. This hints at the possibility these are two discrete functions.

1.4 Transport across the inner membrane

1.4.1 Introduction

All cells require a way to import the substrates necessary for growth across their inner membrane. In some cases certain small molecules are abundant in the environment and simple diffusion will be sufficient for their uptake. However, cells require a mechanism to import larger molecules across the inner membrane, both with and against the concentration gradient. This function is fulfilled by various classes of membrane transport proteins. Common to most of these transporters is that a conformational change is required to facilitate the transport process (Nikaido and Saier Jr 1992, Yan 2013).

These membrane transport proteins can be divided into 2 main types based on the energetics of the mechanism, and are termed either passive or active transporters (Nikaido and Saier Jr 1992). Passive transporters simply allow the transport of substrates along the concentration gradient. These are required for molecules that are either too large for simple diffusion, or where uptake rates are too slow to be biologically sufficient. Conversely, active transport must couple an energetically favourable reaction to an energetically unfavourable one to facilitate transport. Typically, the movement of one substrate against its concentration gradient is coupled with the movement of another substrate down its concentration gradient. When molecules move in opposite directions, the transporter is termed an antiporter or exchanger (Marger and Saier Jr 1993). Still other mechanisms exist, and in some cases the transport of a molecule can be linked to the transport of an ion (Marger and Saier Jr 1993). When a transported molecule and a transported ion move in the same direction across the

membrane the transporter is called a symporter (Nikaido and Saier Jr 1992, Marger and Saier Jr 1993).

Cells require the ability to transport many molecules, many of which cannot be coupled to the transport of another substrate down its gradient. Because of this, cells require other classes of transporters to transport substrates across membranes against the concentration gradient. This comes at the expense of energy. Energy supplied by phospho-enol-pyruvate during group translocation (phospho-enol-pyruvate-dependent phosphotransferase system [PTS systems]) (Postma and Lengeler 1985, Deutscher, Francke et al. 2006), or by ATP hydrolysis via an ATP Binding Cassette (ABC) transporters (Berger and Heppel 1974, Higgins and Litton 1992)

1.4.2 Phospho-enol-pyruvate-dependent carbohydrate phosphotransferase system (PTS)

The phospho-enol-pyruvate-dependent carbohydrate phosphotransferase system (PTS) is used to accumulate various carbohydrates termed PTS sugars or PTS carbohydrates (Postma and Lengeler 1985). The transport and phosphorylation of PTS sugars are linked leading to intracellular accumulation of the carbohydrate phosphate. The same proteins that are involved in the sugars transport are responsible for phosphorylation of the sugar (Postma, and Roseman. 1976).

The metabolism of PTS sugars starts with the transport of the sugar through the PTS system, yielding the carbohydrate phosphate. This is then the first intermediate in the catabolism of the sugar. Because of this, the concentration of the carbohydrate taken up is low within the cell unless the sugar phosphate is hydrolyzed (Postma and Lengeler 1985). This also creates a concentration gradient that favours the translocation of sugar.

Several proteins, soluble and membrane bound, are required to catalyze these reactions. The typical PTS system is a phospho-relay system initially believed to be composed of 3 proteins (Postma and Lengeler 1985), two of which, enzyme I and heat-stable protein (HPr) are required for the first steps in the transport and the phosphorylation of all PTS sugars (Postma and Lengeler 1985, Deutscher, Francke et al. 2006). These proteins are phosphorylated at the expense of PEP. The substrate specificity of the system is conferred by Enzyme II. The component known as Enzyme II is actually composed of 3 proteins (Deutscher, Francke et al. 2006). Enzyme II A which is cytosolic, II B which may be cytosolic or membrane bound, and IIC which is membrane bound. Enzyme II A and II B are part of the phospho-relay, where as enzyme II C is the membrane spanning protein (Deutscher, Francke et al. 2006). Typically each enzyme II is able to recognize a series of structurally similar sugars.

Variations to the basic PTS system exist. In a few organisms such as *Rhodospirillum tubrum*, *Rhodopseudomonas sphaeroides*, and *Pseudomonas aeruginosa* enzyme I and HPr are not separate proteins, but both roles are carried out by a single protein (Saier, Feucht et al. 1971, Brouwer, Elferink et al. 1982, Durham and Phibbs 1982).

Translocation of a substrate via the PTS system starts with the phosphoryl group being transferred from PEP through phospho-enzyme I to HPr. Phospho-HPr transfers the phosphoryl group to phospho-enzyme II A. II A transfers the phosphoryl group to phospho-enzyme II B, and then finally the phosphoryl group is transferred to the sugar as it enters the cell through the membrane spanning Enzyme IIC. (Postma and Lengeler 1985, Deutscher, Francke et al. 2006) (Figure 1.7).

PTS systems have been shown to import many carbohydrates. PTS systems primarily for N-acetylglucosamine, glucose, mannose, galactitol, fructose, glucitol, mannitol, l-sorbose dihydroxyacetone, and lactose have been shown in enteric bacteria alone. However, just because a given sugar is transported via a PTS system in one organism does not mean this is the case in other organisms. Lactose for example, is imported utilizing a PTS system in *Staphylococcus aureus* (Morse, Hill et al. 1968) but is transported by proton symport in the *Enterobacteriaceae* (Hengge and Boos 1983).

Initially it was believed that *Rhizobium* lacked a PTS system (Galibert, Finan et al. 2001). However, it has now been shown that certain *Rhizobia* contain some of the components of a PTS systems (Pinedo, Bringhurst et al. 2008, Prell, Mulley et al. 2012). *S. meliloti* has been shown to contain genes for HPr, and an EIIA^{man}-type enzyme.

HPr and EIIA have been shown to regulate carbon utilization in other bacteria. This is done by regulating metabolic gene transcription by influencing cAMP pools, or via direct physical interactions between PTS components and other types of carbohydrate transporters (Deutscher, Francke et al. 2006). This results in inhibition of the bound transporter, and thus no carbohydrate transport. In-frame *S. meliloti* PTS deletion mutants exhibited altered carbon metabolism phenotypes, suggesting similar interactions may occur in *Rhizobium* (Pinedo, Bringhurst et al. 2008).

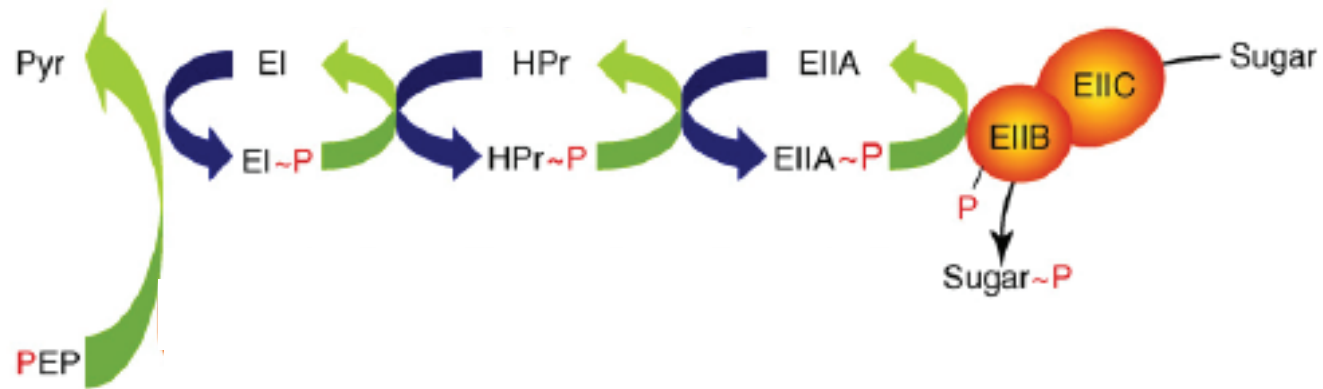


Figure 1.7 A Diagram of the basic phosphotransferase system used by bacteria for the uptake of carbohydrates. Enzyme II A is cytosolic, II B may be cytosolic or membrane bound, and IIC is membrane bound. Enzyme II A and II B are part of the phospho-relay, whereas enzyme II C is the membrane spanning protein. Modified from: Pflüger-Grau, K. and B. Görke (2010). "Regulatory roles of the bacterial nitrogen-related phosphotransferase system." *Trends in microbiology* 18(5): 205-214.

1.4.3 ABC transporters

1.4.3.1 Introduction

One of the most common and largest families of transporters are the ABC transporters. ABC transporters utilize energy from ATP hydrolysis to transport substrates across cellular membrane against the concentration gradient. ABC transporters are responsible for the transport of a diverse range of substrates, both in and out of the cellular compartments. They are found widely distributed in nature, from bacteria (Eitinger , Rodionov et al. 2011) (Higgins and Litton 1992) and archaea (Higgins and Litton 1992 , Horlacher, Xavier et al. 1998) to higher eukaryotes like humans (Eitinger , Rodionov et al. 2011). However, in eukaryotic systems ABC transporters are limited to exporters.

ABC transporters have been demonstrated to transport a wide variety of substrates including sugars, amino acids, metal ions, and peptides. Members of the ABC super family can be defined by the ATPase or ABC protein. There are 48 human ABC genes characterized to date (Dean, Hamon et al. 2001). Defects in some human ABC genes have been implicated in human illnesses such as cystic fibrosis (Dean, Hamon et al. 2001, Gadsby, Vergani et al. 2006), neurological diseases, and retinal degeneration (Dean, Hamon et al. 2001). In bacteria, ABC transporters have been implicated in antibiotic resistance (Dawson and Locher 2006), pathogenesis and cell division (Jones and George 2004).

1.4.3.2 The core structure of an ABC transporter

ABC transporters require the function of multiple protein domains organized in a specific fashion. They are typically defined by the ABC domain or ATPase often called the nucleotide binding domain (NBD). The ABC protein contains the Walker_A and Walker_B

motifs common to all ATP requiring proteins (Walker, Saraste et al. 1982) and the LSGGQ conserved consensus sequence (Nikaido 2002).

The Walker_A motif or P-loop is defined by the amino acid sequence GXXXXGK with X-denoting any amino acid, and is associated with phosphate binding (Walker, Saraste et al. 1982). The Walker_B motif downstream of the P-loop was originally identified as (R/K)XXXXGXXXXLhhhhD, with h representing any hydrophobic amino acid (Walker, Saraste et al. 1982). However, more recently the Walker_B motif was revised to hhhhDE, with the glutamate being essential for ATP-hydrolysis (Hanson and Whiteheart 2005). The LSGGQ motif essential for ATP binding in ABC transporters is also sometimes called the “C-motif” or “Signature motif” as it is unique to ABC type proteins (Nikaido 2002). The ATP binding domain is the most conserved domain in ABC transporters, particularly with the Walker_A, Walker_B motif and the LSGGQ conserved consensus sequence (Nikaido 2002).

ABC transporters generally consist of two membrane spanning, or transmembrane domains (TMD), with 6 membrane spanning regions forming the pathway the substrate passes through. This is often referred to as the permease. This domain plays a role in substrate specificity, likely explaining the diversity in amino acid sequence found in transmembrane domains (Higgins and Litton 1992). The structural constraints required for the function of the trans membrane domain can be satisfied by many amino acid combinations. There is relatively little sequence conservation that exists in this domain between different ABC transporters (Higgins and Litton 1992).

The ATP binding domains are associated with the membrane spanning components. Each ABC transporter has 2 ATP binding domains, where both are required for functional transport. In some cases these may be 2 identical ATP binding proteins (Dassa and Bouige

2001, Higgins 2001). In others systems it is coded for by 2 different genes (Biemans-Oldehinkel, Doeven et al. 2006). The permeases are typically composed of two separate genes coding membrane spanning domains, or two copies of the same gene product (Biemans-Oldehinkel, Doeven et al. 2006). However they are occasionally found as a single larger gene product comprising the entire permease (Biemans-Oldehinkel, Doeven et al. 2006). In export systems the nucleotide binding domain can be found fused to the transmembrane domains (Biemans-Oldehinkel, Doeven et al. 2006).

1.4.3.3 Evolution and spread of ABC transporters

ABC transporters are members of an ancient protein family that can be divided into two main groups, uptake systems and export systems. The divergence between importers and exporters likely happened early in their evolution, as ABC transporters segregate phylogenetically based on this function rather than by the prokaryotic versus eukaryotic origin of the transporter (Saurin, Hofnung et al. 1999). This suggests that the divergence between import and export systems happened once in evolution, prior to the differentiation between prokaryotes and eukaryotes (Saurin, Hofnung et al. 1999, Dassa and Bouige 2001). Despite differences between import and export systems, the ABC protein is still one of the more highly conserved protein families, and shows significant conservation in key domains from bacteria to humans (Higgins and Litton 1992 , Saurin, Hofnung et al. 1999, Dawson and Locher 2006). The basic “tweezer” transport mechanism is also conserved in both importers and exporters (Dassa and Bouige 2001, Biemans-Oldehinkel, Doeven et al. 2006). ABC transporters as a group are responsible for the transport of a large and diverse number of substrates. This leads to more divergence in the membrane spanning permease proteins

(Higgins and Litton 1992) and any accessory proteins or domains (Biemans-Oldehinkel, Doeven et al. 2006).

ABC transporters, and specifically carbohydrate importers are evolutionarily old systems found in both gram positive, gram negative and archeal organisms with strikingly strong homology between archeal and bacterial transport components (Horlacher, Xavier et al. 1998).

1.4.3.4 Discovery of ABC-type importers in bacteria

While studying energy coupling mechanisms for active transport in *E. coli*, glutamine uptake was shown to be different in that its transport was shown to require ATP for energy, and an ATPase activity (Berger 1973). While studying osmotic shock in *E. coli* models, the transport of some molecules was shown to be osmotic shock sensitive, with the ability to transport the molecules permanently lost after osmotic shock (Berger and Heppel 1974) .

Osmotic shock was known to remove the outer membrane and release the components of the periplasmic space. It was subsequently shown that if the released protein content of the periplasm could be supplied to osmotically shocked bacteria in a high enough concentration, transport could be restored (Berger and Heppel 1974). It was demonstrated that the osmotic shock sensitive transporters were the ones that required ATP and ATPases. It was concluded that there were two classes of transporters. 1) A shock sensitive class that required a then unknown periplasmic protein (the SBP), and was dependant on ATP and an ATPase; and 2) another class not dependant on a periplasmic protein, that utilized a different energy coupling mechanism to power transport (Berger and Heppel 1974).

1.4.3.5 Prokaryotic importers

Prokaryotic importers alone are responsible for the transport of amino acids, nucleosides, carbohydrates, sources of nitrogen, phosphorus, iron and sulphur, vitamins, and trace elements such as Mn, Zn, Ni ions (Eitinger , Rodionov et al. 2011).

Prokaryotic ABC importers are composed of the “core components” common to all ABC-type transport systems. They are powered by the Walker_A/Walker_B, LSGGQ containing ABC-type ATPase (Chen, Lu et al. 2003) with membrane spanning permeases making up the channel the substrates are translocated through (Biemans-Oldehinkel, Doeven et al. 2006).

Gram negative importers however, also require a periplasmic substrate binding protein for the function of the transport system (Berger and Heppel 1974, Higgins and Litton 1992 , Higgins 2001). The periplasmic substrate binding protein plays a central role in substrate specificity. They are structurally conserved, generally having 2 globular domains with a cleft that forms the substrate binding site (Higgins and Litton 1992). However, there is little sequence similarity between binding proteins for different substrates, likely due to different requirements for specific substrate binding (Schneider 2001).

1.4.3.6 Gram positive and Archaea substrate binding protein dependant importers.

Initially binding protein dependant transport systems were thought to be limited to Gram negative organisms as the substrate binding protein is localized to the periplasmic space. However, subsequently binding protein dependant transport systems in gram positive organisms were identified. The evidence for these systems in Gram positive bacteria were first reported in *Streptococcus pneumonia*. Similar systems are also found in the mycoplasma

species *Mycoplasma hyorhinis* (Gilson, Alloing et al. 1988). They maintain a high affinity substrate binding protein near the membrane by means of an NH₂ terminal lipo-amino acid anchor (Gilson, Alloing et al. 1988). Subsequently binding protein dependant systems were discovered in the hyperthermophilic archaeon *Thermococcus litoralis* (Xavier, Martins et al. 1996). Despite the presence of binding protein dependant ABC transport systems having been identified in Gram-positive, gram-negative and archaeal systems, most of the work on bacterial ABC importer systems has been limited to gram-negative models. Of this work, most has been done using the *E. coli* maltose transporter system as a model (Schneider 2001).

1.4.3.7 Carbohydrate Uptake Transporters 1 and 2 (CUT 1 and 2)

Bacterial ABC transport systems utilized for the uptake of carbohydrates can be broken into two main classes: Carbohydrate uptake transporters 1 and 2 (CUT 1 and 2) (Schneider 2001). While both require a substrate binding protein, in the CUT 1 family transmembrane domains are more typically expressed as two separate polypeptides. These transmembrane domains interact with two copies of the ABC protein. CUT 2 transporters typically have a single hydrophobic transmembrane gene product that forms a homodimer (Schneider 2001); however these distinctions are not universal (Quentin and Fichant 2000).

CUT 1 and CUT 2 transporters are defined by differences in their ABC protein (Schneider 2001). CUT 1 ABC proteins contain the standard Walker_{A/B} motifs common to all ATPase proteins, as well as the LSGGQ signature motif that defines the ABC protein (Schneider 2001). CUT 2 transporters are defined by an ABC protein that contains a fusion of two nucleotide binding domains, and thus is almost twice the size of their CUT 1

counterpart (Schneider 2001). However, the second nucleotide binding domain of the CUT 2 transporter is non functional due to a lysine being substituted for an arginine in the C-terminal Walker_A motif. This variant has been shown to result in a loss of ATP hydrolysis function (Schneider, 2001).

Transport kinetics studies (Szmelcman, Schwartz et al. 1976), spheroplast and proteoliposome reconstitution of transport experiments (Davidson and Nikaido 1991) and structural studies (E. and Hunke 1998, Jones and George 2004) have given rise to a clearer understanding of the mechanism and kinetics of sugar uptake systems in bacteria. Most of this work has been done in the CUT 1 *E. coli* maltose transporter (Bordignon, Grote et al. 2010).

Most of the work reported for CUT 2 systems has been carried out using the *E. coli* ribose transporter as a model (Schneider 2001). Currently the arrangement of the CUT 2 transporter components and its mechanism is assumed to be similar to the CUT 1 (Schneider 2001). Little work has been done using CUT 2 models to confirm these assumptions. Artificial constructs fusing RbsA (the ABC protein) to RbsC (the permease) have been used to restore growth on ribose *in vivo* (Park, Cho et al. 1999). This suggests a 1:1 ration of the components, and supports the assumption that the arrangement is the same as the CUT 1 transporter. However, functional reconstitution of ribose transport using proteoliposomes has not been successful, and despite the larger ABC protein, models for the transport mechanism are assumed the same.

1.4.3.8 E. coli Maltose transporter as a model for the mechanism of transport

The *E. coli* maltose transporter is one of the best studied models of ABC transport (See Figure 1.8). It consists of the SBP (MalE), the TMDs of the permease (MalF and MalG), and two ABC protein subunits (MalK₂). Crystal structures of four different conformations of the MalK dimer have been isolated, (Chen, Lu et al. 2003 , G., J.M. et al. 2005) as well as two of the full transporter (Oldham, Khare et al. 2007) (Khare, Oldham et al. 2009). This is the only ABC importer with two X-ray models of the full transporter, as structures are available for both the open and the closed state. These data have been used to confirm the basic “alternating access” model for transport.

Conserved sequences in MalF and MalG known as the “EAA-motifs” each interact with a MalK monomer via a cleft on top of MalK (Dassa and Hofnung 1985). The interaction is mediated by a salt bridge between Arg-47 of each MalK monomer and the glutamates of the “EAA-motifs” (MalF, Glu-401; MalG, Glu-190) (Oldham, Khare et al. 2007).

The two ABC proteins are bound at the C-terminal region known as the regulatory domain (RD) (Bordignon, Grote et al. 2010). The two ABC proteins share two ATP molecules. Each ATP is bound between the Walker_A motif of one subunit and the LSGGQ motif of the other. The full closure of the ABC protein occurs only when two ATP molecules are bound to the dimer. The Walker_A motif from one subunit and the LSGGQ motif from the other are involved in electrostatic interactions with the phosphates from the ATP, allowing the ATP to bind the proteins (Bordignon, Grote et al. 2010).

ATP hydrolysis results in a change in conformation, as illustrated by the ATP unbound dimer structure. Although conformational changes from the closed ATP-bound

state, to the semi-open structures and finally the open structure occur, they all show the same interaction between the C-terminal regions of MalK (Bordignon, Grote et al. 2010) . In this model the MalK subunits function as “tweezers”, with the C-terminal RD region acting as a pivot. This “tweezer” motion forces changes to the conformation of the permease, from periplasmic side open (ATP bound), to cytoplasmic side open (ATP unbound) (Bordignon, Grote et al. 2010).

How the availability of the substrate in the periplasm is communicated to the cytosolic transporter components is not well understood. What is known is that MalE (SBP) interacts with the periplasmic MalF-P2 loop. The MalF-P2 loop undergoes conformational changes throughout the catalytic cycle (Grote, Polyhach et al. 2009) . Rearrangements of MalF-P2 loop were found to be dependent on ATP and ligand bound MalE. The MalF-P2 loop is found in three conformations. These correlate with the open, closed and semi-open conformations of MalK₂. Full closure of the ATP-bound MalK₂ occurs when the ligand bound MalE binds to the TMDs, of the permeases MalF and MalG (Grote, Polyhach et al. 2009). The ATPase activity of MalK₂ is significantly increased by adding ligand bound MalE (Landmesser, Stein et al. 2002) . This leads to the hydrolysis of ATP, and the opening of the “tweezer”. The current model for the maltose transport cycle in *E. coli* is summarized in Figure 1.8 (Bordignon, Grote et al. 2010).

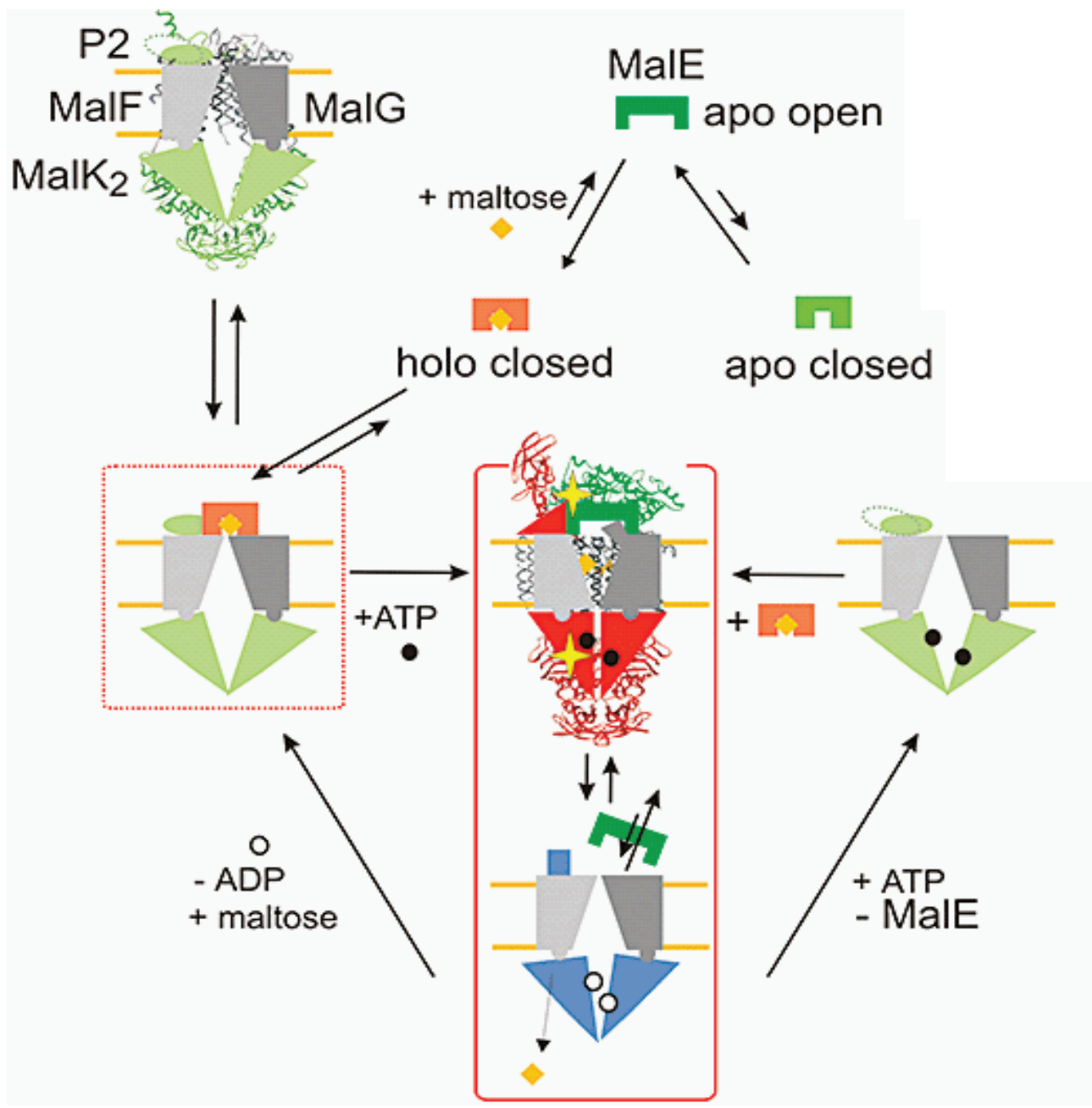


Figure 1.8. The current model for the ATP dependant ABC transport mechanism. The equilibrium between an open apo-, a closed apo-and a closed holo-form of MalE is presented. The TMDs are shown to interchange between a cytoplasmic open (ADP-bound) and a periplasmic-open(ATP-bound) state (alternating mechanism). The correspondence of three different states for MalK2 and MalF-P2 is indicated by the same colour coding, the flexibility of MalF-P2 in the absence of MalE is indicated by a dotted green ellipse. Transitions between the two conformers highlighted in the red box represent ATP-dependent import of maltose in the cell, with the stars denoting the instable nature of the ATP-bound intermediate. The dotted box highlights a short-lived intermediate (the apo-state) in the cell. Taken from: Bordignon, E., et al. (2010). "The maltose ATP_binding cassette transporter in the 21st century—towards a structural dynamic perspective on its mode of action." Molecular Microbiology 77(6): 1354-1366.

1.4.3.9 Variations of the ABC importer architecture

The vast majority of ABC type importers are organized with the standard architecture of two TMD forming the permeases (one or two gene products), two ABC proteins (again one or two gene products), with a single SBP. However variations to this organization have been identified.

Systems where the TMDs are each fused to a substrate binding domain exist. This results in two substrate-binding domains (SBPs) per complete transporter. An example of a transporter with this architecture is the glycine betaine transporter from *Lactococcus lactis* (Biemans-Oldehinkel and Poolman 2003)

In other transporters, two SBPs are fused to each TMD. Two of these units make up the complete permease, meaning these systems have four substrate binding domains. An example of this arrangement is the GlnPQ transporter from *Lactococcus lactis* (van der Heide and Poolman 2002, Schuurman-Wolters and Poolman 2005). Transporters with SBPs fused to the TMDs have been identified in Gram-negative bacteria but are not as common as in gram-positive models (van der Heide and Poolman 2002).

There are multiple examples of a single “core transporter” (TMD and NBD) recognizing more than one SBP. Each SBP is specific for different substrates, but use the “same core transporter”. The HisQMP “core transporter” from *S. typhimurium* has two SBPs. ArgT binds lysine, ornithine, and arginine, and HisJ binds arginine and histidine (Thomas 2010). The CbcWV “core transporter” has three SBPs, CaiX, BetX, CbcX, which recognize carnitine, choline, and betaine respectively (Thomas 2010).

Genome sequencing reveals that Rhizobia contain a disproportionate number genes encoding ABC type-proteins relative to most genomes (Kaneko, Nakamura et al. 2000 ,

Galibert, Finan et al. 2001 , Young, Crossman et al. 2006). Because of this abundance, Rhizobium species present a good model to search for unique or novel ABC-transporter systems.

1.5 Protein-protein interactions

1.5.1 Basic factors responsible for protein-protein interactions.

Protein-protein interactions are mediated by many complex factors that can be difficult to predict; however a few major factors are typically significant determinants in protein-protein surface complementarity. Protein-protein interaction sites are formed at protein surfaces with good size, shape and electrostatic complementarity. The flexibility of the molecules involved also play a major role (Moreira, Fernandes et al. 2007).

The contact size of the interaction is a key determinant in protein-protein interactions. For example, heterodimer interfaces typically need to be larger than 600 Å. The energy of protein-protein binding is directly related to the amount of buried hydrophobic surface area. As such, most protein-protein interfaces are frequently hydrophobic, and bury large regions of nonpolar surface area (Jones and Thornton 1996). These types of interactions are driven by the gains in free energy that result from these regions moving from a polar (aqueous) to a nonpolar environment, increasing entropy and favoring complex formation. They are mediated by van der Waals contacts between nonpolar amino acid residues. Hydrophobic interactions result in tightly packed residues that are organized as patches that tend to protrude from the surface (Jones and Thornton 1996).

Electrostatic complementarity is also a major factor involved in protein-protein interactions. Electrostatic complementarity of the interacting protein surfaces promotes complex formation, and determines the lifetime of a complex. Amino acid side chains can form hydrogen bonds at the protein interfaces, or between the protein contact surfaces and the surrounding water molecules (Jones and Thornton 1996).

1.5.2 Protein binding hot spots

Alanine-scanning mutagenesis has demonstrated that most of the protein–protein binding energy is due only to a few amino acid residues located at the protein-protein interface. These residues are termed hot spots. Hot spots have now been defined as a site where alanine mutations cause an increase in binding free energy of at least 2.0 kcal/mol. (Moreira, Fernandes et al. 2007). It has been shown that on average only 9.5% of the residues located at an interface are hot spots. Hot spots of one interacting face are typically packed against the hot spot of its interacting partner (Moreira, Fernandes et al. 2007)

ASEdb is a database consisting of alanine substitutions and their effects on the free energy of binding in protein interactions. An analysis of ASEdb has demonstrated that hot spots are not random, but are biased in both location, and amino acid composition (Moreira, Fernandes et al. 2007). Although no geometric reason has been proposed, hot spots tend to be compact and centralized, with very few hot spot residues found at the edges of an interface. Tryptophan, arginine and tyrosine are all particularly common in hot spots. Tryptophan appears in hot spots at a particularly high rate as it is a hydrogen bonding donor, has a large hydrophobic surface and can protect fragile hydrogen bonds from water. Additionally, due to a significant difference in size, a tryptophan to alanine substitution can generate a large cavity, which creates a highly complex destabilization. Conversely, leucine, serine, threonine, and valine residues are almost absent in known hot spots (Moreira, Fernandes et al. 2007).

1.5.3 Protein-Protein interactions involving transport systems and sugar kinases.

An ABC transporter requiring a sugar kinase for functional transport appears to be novel. There are no current examples of sugar kinases directly or indirectly interacting with ABC transporters. However, there is a precedent of other protein-protein interactions affecting the function of some transport systems. There are also examples of sugar kinases interacting with non-transport systems proteins.

1.5.3.1 Glucose PTS component EIIA^{glc} interacts with LacY and MelB

LacY is a H⁺-symporter that exploits electrochemical potential for the uptake of several galactosides. The structure of LacY has been determined, revealing a protein that contains 12 hydrophobic transmembrane domains connected by hydrophobic loops (Abramson, Smirnova et al. 2003). Substrate binding appears to be located at the interfaces of helices IV, V and VII (Deutscher, Francke et al. 2006). The current model for lactose transport has LacY in a conformation where the sugar binding site is facing the periplasm, termed “outward.” The binding of the sugar appears to result in a change in the structural conformation to one where the substrate binding site is exposed to the cytoplasm (termed “inward”). LacY binds to the PTS protein EIIA^{Glc} when in the “outward” conformation, preventing the structural change to “inward” facing conformation, thus blocking the transport cycle (Hoischen, Levin et al. 1996, Seok, Sun et al. 1997, Sondej, Sun et al. 1999, Sondej, Weinglass et al. 2002, Sondej, Vazquez-Ibar et al. 2003). The residues involved in this interaction have been identified by mutagenesis, as changing Thr-7 or Met-11 to Ile prevents EIIA^{Glc} from binding LacY (Sondej, Sun et al. 1999).

1.5.3.2 MalK physically interacts with MalT

MalK has been shown to sequester MalT, the positive regulator of its own transcript, when it is in the resting state (Richet, Davidson et al. 2012). Departure from the resting state abolishes this interaction (Panagiotidis, Boos et al. 1998, Richet, Davidson et al. 2012) Two regions of MalK responsible for the interaction with MalT have been identified by mutagenesis (Richet, Davidson et al. 2012). Patch W267, made up of residues A248, A250, I251, Q253, Q255, W267 (Richet, Davidson et al. 2012) and Patch D297 defined by residues L291, D297, E350 (Richet, Davidson et al. 2012).

1.5.3.3 EIIA^{Glc} interacts with MalK

Two EIIA^{Glc} molecules bind to the MalK₂ dimer. One EIIA^{Glc} binds the nucleotide binding domain of one subunit, and the regulatory domain of the other subunit. The other EIIA^{Glc} binds the inverse (Chen, Oldham et al. 2013). This prevents the “tweezers-like motion” of the transporter, and renders it non-functional. Interestingly, MalK appears to have different binding sites for EIIA^{Glc} and MalT. Mutations to MalK affecting the binding of EIIA^{Glc} do not affect the binding of MalT, and *vice versa* (Deutscher, Francke et al. 2006).

1.6 Thesis objectives

Rhizobium leguminosarum bv. *trifoli* Rlt100 has been shown to be less competitive for nodule occupancy when unable to utilize rhamnose as a sole carbon source (Oresnik, Pacarynuk et al. 1998). This led to the subsequent investigation of the genes and proteins involved in *Rhizobium leguminosarum* rhamnose catabolism.

This investigation revealed a rhamnose catabolic sugar kinase that appeared to exhibit unique features. Most rhamnose catabolic kinases have activity phosphorylating rhamnulose (Rodionova, Li et al. 2013); however the Rlt100 rhamnose catabolic kinase RhaK was shown to exhibit activity directly phosphorylating rhamnose (Richardson, Hynes et al. 2004 , Richardson and Oresnik 2007).

In addition to this, the associated CUT 2 type rhamnose ABC transporter has been shown to require RhaK for the uptake of rhamnose (Richardson and Oresnik 2007). Although RhaK has been shown to have activity directly phosphorylating rhamnose, it is believed that unlike PTS transport, RhaK is not involved in a single transport and phosphorylation event. Since RhaU is a mutarotase that acts directly on rhamnose it must act between the transport and phosphorylation of rhamnose at least some of the time (Richardson and Oresnik 2007). The general mechanisms of ABC-type transporters have been well characterized (Chen, Lu et al. 2003 , Jones and George 2004). However, no mechanism has been confirmed for any CUT 2 class transporter, and no other known ABC-type transporters are known to require a sugar kinase for activity. The role of RhaK in this association is unknown.

To date, the inability to catabolize various carbon compounds have been associated with the inability to compete for nodule occupancy in Rhizobia. However, only one, myo-inositol catabolism appears to affect this competition for nodule occupancy in more than one *Rhizobium* species (Gordon, Ryder et al. 1996, Oresnik, Pacarynuk et al. 1998, Fry, Wood et al. 2001, Jiang, Krishnan et al. 2001, Yost, Rath et al. 2006, Poysti, Loewen et al. 2007, Geddes, Pickering et al. 2010, Kohler, Zheng et al. 2010, Vanderlinde, Hynes et al. 2013).

At the onset of my research, it was hypothesized that the RhaK protein's ability to phosphorylate rhamnose, and its role in transport were separate activities. It was reasoned that its role in transport was likely through a direct or indirect physical interaction with a component of the transporter. It was predicted that these two functions of the protein could be uncoupled, and that the ability to transport the sugar would be dependent on this physical interaction. In addition, it was predicted that in other related organism similar interactions would be indentified in orthologous proteins, and that mutations to their RhaK or other rhamnose catabolic proteins may have an effect on competition for nodule occupancy.

With this in mind, my thesis objectives were to:

- 1) Determine if the kinase function and transport role of RhaK are independent functions that can be uncoupled.
- 2) Determine if the role RhaK plays in transport is dependent on either a direct or indirect physical interaction with a transporter component. If so, to determine the nature of this interaction.
- 3) Determine if RhaK dependant rhamnose transport is an anomaly unique to *R. leguminosarum*, or if it is also seen in other related organisms.

- 4) Determine the biological role of RhaK phosphorylating rhamnose directly. How is the catabolic pathway of *R. leguminosarum* different than the canonical phosphorylated rhamnose pathway?
- 5) Investigate if the inability to catabolize rhamnose is a more general determinant in competition for nodule occupancy, or if it is restricted to *R. leguminosarum*.

Chapter 2:

Carbohydrate Kinase (RhaK) Dependant ABC-Transport of Rhamnose in *Rhizobium*

leguminosarum:

Genetic Separation of Kinase and Transport Activities

2.1 Introduction

Cells require a specific and regulated way to transport substrates across membranes. One of the largest families of transporters is the ATP binding cassette (ABC) transporters. ABC transporters utilize free energy from ATP hydrolysis to transport substrates across a membrane. They are widely distributed in all domains of life and are involved in transport that affects diverse biological functions (Dassa and Bouige 2001, Davidson and Chen 2004, Rea 2007).

Members of the ABC super family are defined by the ATPase protein that contains the Walker_A and Walker_B motifs, along with a LSGGQ conserved consensus sequence (Higgins 2001, Higgins and Litton 2004). Functional ABC importers generally consist of two proteins that have transmembrane domains, consisting of 6 membrane spanning regions (permeases), two ABC proteins that are cytoplasmically localized and contain the ATP binding domains, and a periplasmically localized substrate binding protein for the function of the transport system (Bordignon, Grote et al. 2010). Relatively few sequence similarities are found between binding proteins for different substrates as the periplasmic substrate binding protein plays a central role in substrate specificity (Higgins 1992).

Gram negative bacterial ABC transport systems responsible for the import of carbohydrates can be broken into two main classes; carbohydrate uptake transporters 1 and 2 (CUT 1 and 2) (Schneider 2001). Most work on bacterial carbohydrate ABC importer systems has been done using the CUT1 *E. coli* maltose transport system as a model and many aspects of its function are understood at the atomic level giving rise to a clear understanding of the mechanism and kinetics of sugar uptake systems (Chen, Lu et al. 2003, Lu, Westbrook et al. 2005, Oldham, Khare et al. 2007, Khare, Oldham et al. 2009, Richet,

Davidson et al. 2012). In contrast, less work has been carried out using CUT2 transporters as models.

CUT2 transporters are classified primarily on the basis of the ABC protein (Schneider 2001). The ABC protein tends to be about 500 amino acids long, which contrasts with about 300 amino acids generally found in members of the CUT1 family (Schneider 2001). This protein appears to have arisen by a fusion of 2 ABC domains (Eitinger, Rodionov et al. 2011). Although the ABC protein contains two ATP binding sites, it is not clear that both nucleotide binding motifs are functional in ATP hydrolysis since a key lysine residue in the C-terminal Walker_A is substituted with an arginine (Schneider 2001 , Bordignon, Grote et al. 2010). This type of mutation has been shown to adversely affect activity (Schneider and Hunke 1998). Due to its size, it is not clear whether CUT2 transporters function with one or two copies of this larger ABC protein (Schneider 2001 , Eitinger, Rodionov et al. 2011).

The genomes of Rhizobium species tend to have high numbers of genes encoding ABC transport systems. *Sinorhizobium meliloti*, *R. leguminosarum*, and *Mesorhizobium loti* contain 200, 269, and 216 annotated ABC genes respectively (Kaneko, Nakamura et al. 2000, Finan, Weidner et al. 2001, Galibert, Finan et al. 2001, Young, Crossman et al. 2006). In contrast, *Pseudomonas aeruginosa*, which has a comparable genome size, and can also be found as a soil organism, contains about 124 (Stover, Pham et al. 2000). Since Rhizobia do have a high number ABC transport systems, and a number of these strains are being used as model systems, studying them in these organisms may lead to findings that give insight into nuances associated with ABC transport.

In *R. leguminosarum*, strain Rlt100, the inability to transport or catabolise rhamnose leads to a decreased ability to compete for nodule occupancy relative to the wild-type

(Oresnik, Pacarynuk et al. 1998). Transport of rhamnose was shown to be encoded by a CUT2 type ABC transporter encoded by *rhaSTPQ* (Richardson, Hynes et al. 2004). Strains carrying a mutation in the sugar kinase (*rhaK*) that is necessary for the catabolism of rhamnose were unable to transport labelled rhamnose into the cell, despite the components of the transporter being transcribed, translated and localized to the membrane (Richardson and Oresnik 2007), suggesting that RhaK affects ABC transport in addition to having kinase activity. In this chapter we address the hypothesis that the kinase activity of RhaK is distinct from its ability to affect the transport of rhamnose.

2.2 Materials and Methods

2.2.1 Bacterial strains and culture conditions

E. coli strains used in this work were grown on LB as a complex medium at 37°C (Davis, Botstein et al. 1980). *R. leguminosarum* strains were grown at 30°C on TY as a complex medium (Beringer 1974), or on VMM as a defined medium (Vincent 1970).

Carbon sources were filter sterilized and added to a final concentration of 15 mM.

Antibiotics were added at the following concentrations for solid media: Tetracycline (Tc) 5 µg/ml, Streptomycin (Sm) 200 µg/ml, Neomycin (Nm) 200 µg/ml, kanamycin (Kan) 50 µg/ml, and ampicillin (Amp) 100 µg/ml, and halved for growth in liquid media. (Table 2.1).

Table 2.1 Strains and Plasmids

Strain or Plasmid	Relevant Genotype	Reference or Source
<i>R. leguminosarum</i>		
Rlt100	W14-2 Sm ^r , wild-type	(Baldani <i>et al.</i> 1992)
Rlt144	Rlt100, <i>rhaK50::Tn5-B20</i>	(Richardson <i>et al.</i> 2004)
Rlt106	Rlt100, <i>rhaT2::Tn5-B20</i>	(Oresnik <i>et al.</i> 1998)
<i>E. coli</i>		
DH5α	<i>endA hsdR17 supE44 thi-1 recA1 gyrA96 relA1 (argF-lacZYA) U169 80dlacZ M15</i>	(Hanahan 1983)
MT616	MT607 (pRK600)	(Finan <i>et al.</i> 1986)
Plasmids		
pRK7813	Broad host range vector, Tc ^r	(Jones <i>et al.</i> 1987)
pRK600	pRK2013 <i>npt::Tn9</i> , Cm ^r	(Finan <i>et al.</i> 1986)
pMR110	pRK7813 with <i>rhaK</i> ⁺ expressed from p _{lac} promoter	(Richardson <i>et al.</i> 2007)
pMR178	<i>rhaK</i> with N-terminal His ₆ tag, in pRK7813	This work
pDR170	pRK7813 with <i>rhaK72</i> with a N-terminal His-Tag expressed from p _{lac} promoter	this work
pLit28i	high copy number <i>colEI</i> cloning vector Ap ^r	New England Biolabs
pDR1	<i>rhaK</i> in pLit28i	this work
pDR22	pRK7813 with <i>rhaK66</i> expressed from p _{lac} promoter	this work
pDR70	pRK7813 with <i>rhaK65</i> expressed from p _{lac} promoter	this work
pDR72	pRK7813 with <i>rhaK70</i> expressed from p _{lac} promoter	this work
pDR73	pRK7813 with <i>rhaK72</i> expressed from p _{lac} promoter	this work
pDR74	pRK7813 with <i>rhaK73</i> expressed from p _{lac} promoter	this work
pDR77	pRK7813 with <i>rhaK88</i> expressed from p _{lac} promoter	this work
pDR78	pRK7813 with <i>rhaK74</i> expressed from p _{lac} promoter	this work
pDR79	pRK7813 with <i>rhaK91</i> expressed from p _{lac} promoter	this work
pDR80	pRK7813 with <i>rhaK84</i> expressed from p _{lac} promoter	this work
pDR81	pRK7813 with <i>rhaK76</i> expressed from p _{lac} promoter	this work
pDR82	pRK7813 with <i>rhaK78</i> expressed from p _{lac} promoter	this work
pDR144	pRK7813 with <i>rhaK69</i> expressed from p _{lac} promoter	this work

pDR146	pRK7813 with <i>rhaK79</i> expressed from p _{lac} promoter	this work
pDR147	pRK7813 with <i>rhaK85</i> expressed from p _{lac} promoter	this work
pDR155	pRK7813 with <i>rhaK92</i> expressed from p _{lac} promoter	this work
pDR157	pRK7813 with <i>rhaK95</i> expressed from p _{lac} promoter	this work
pDR158	pRK7813 with <i>rhaK97</i> expressed from p _{lac} promoter	this work
pDR159	pRK7813 with <i>rhaK115</i> expressed from p _{lac} promoter	this work
pDR160	pRK7813 with <i>rhaK114</i> expressed from p _{lac} promoter	this work
pDR161	pRK7813 with <i>rhaK96</i> expressed from p _{lac} promoter	this work
pDR162	pRK7813 with <i>rhaK104</i> expressed from p _{lac} promoter	this work
pDR163	pRK7813 with <i>rhaK103</i> expressed from p _{lac} promoter	this work
pDR164	pRK7813 with <i>rhaK109</i> expressed from p _{lac} promoter	this work
pDR165	pRK7813 with <i>rhaK116</i> expressed from p _{lac} promoter	this work
pDR166	pRK7813 with <i>rhaK111</i> expressed from p _{lac} promoter	this work
pDR167	pRK7813 with <i>rhaK117</i> expressed from p _{lac} promoter	this work
pDR168	pRK7813 with <i>rhaK119</i> expressed from p _{lac} promoter	this work
pDR169	pRK7813 with <i>rhaK121</i> expressed from p _{lac} promoter	this work
pDR181	pRK7813 with <i>rhaK122</i> expressed from p _{lac} promoter	this work
pDR182	pRK7813 with <i>rhaK64</i> expressed from p _{lac} promoter	this work
pDR183	pRK7813 with <i>rhaK90</i> expressed from p _{lac} promoter	this work
pDR184	pRK7813 with <i>rhaK94</i> expressed from p _{lac} promoter	this work
pDR185	pRK7813 with <i>rhaK98</i> expressed from p _{lac} promoter	this work
pDR186	pRK7813 with <i>rhaK99</i> expressed from p _{lac} promoter	this work
pDR187	pRK7813 with <i>rhaK100</i> expressed from p _{lac} promoter	this work
pDR188	pRK7813 with <i>rhaK112</i> expressed from p _{lac} promoter	this work
pDR189	pRK7813 with <i>rhaK118</i> expressed from p _{lac} promoter	this work
pDR191	pRK7813 with <i>rhaK105</i> expressed from p _{lac} promoter	this work
pDR192	pRK7813 with <i>rhaK106</i> expressed from p _{lac} promoter	this work
pDR193	pRK7813 with <i>rhaK80</i> expressed from p _{lac} promoter	this work

2.2.2 Genetic manipulations

Conjugations between *E. coli* strains and *R. leguminosarum* were performed as described previously using the mobilizing strain MT616 (Finan, Hirsch et al. 1985).

2.2.3 DNA manipulations, constructions

Standard techniques were used for gel electrophoresis, restriction enzyme digests, and isolation of plasmid DNA (Sambrook and Russell 2001). Nucleotide sequencing was carried out by cycle sequencing using a Big-Dye version 3.1 kit. Sequencing reactions were carried out as recommended by the manufacturer and resolved using an ABI 3130 sequencer.

2.2.4 Construction of pDR1

The *rhaK* coding region was removed from pMR110 using *Bam*HI and *Hind*III, and ligated into pLit28i. The resultant construct was confirmed by sequencing and named pDR1 (Table 1). To over express a RGS-His6 tagged variant of RhaK72 in *R. leguminosarum*, a C-terminal tagged version was constructed. pDR73 was used as a template. The primers 5' HindRBS, -5' ATATAAGCTTGGAGAACTGCAGATGACCGCCAGTTCCTATC-3' (containing a *Hind*III restriction site, RBS, and the 5' annealing portion of *rhaK*), and 3His 6B, 5'-ATGGATCCCTAATGATGATGATGATGATGCGATCCTCTTGCCATCGCCGC GTACC-3' (containing a *Bam*HI restriction site, a C-terminal RGS-His6 tag, and the 3' annealing portion of *rhaK*) and *rhaK72* as a template were used to create a C-terminal tagged RhaK72. The PCR product was subsequently cut with *Bam*HI and *Hind*III and cloned into pRK7813 such that it was expressed from the *plac* promoter. The construct was sequenced and named pDR170.

2.2.5 Generation of Tn7 in frame insertional rhaK alleles

In-frame insertion mutations were generated using New England Biolabs GPS-LS Mutagenesis Kit. Briefly, an *in vitro* mutagenesis was carried out using pDR1, pGPS5 (transposon donor), and TnsABC (transposase). The reaction was transformed into library grade competent cells. Transformants carrying insertions within the coding sequence of *rhaK* were identified by restriction digestion using *Bam*HI and *Hind*III. The point of insertion of each allele was determined by sequencing using the suggested primers that were designed to the ends of the Tn7. Inserts that were within the coding sequence were subsequently recloned into pRK7813. The Tn7 construct had *Pme*I sites placed such that the Tn7 DNA could be deleted by cutting with *Pme*I and religating. This would leave behind a 15 base pair insert (10 base pairs from the transposon (TGTTTAAACA) and a 5 bp duplication at the site of insertion) that could either generate 5 amino acid insert or a stop codon in the final protein. Each allele was confirmed by sequencing and conceptually translated to determine the nature of the insertion.

2.2.6 Rhamnose transport assay

Transport assays were carried out as previously described with slight modifications (Richardson, Hynes et al. 2004, Richardson and Oresnik 2007, Richardson, Carpena et al. 2008). Radioactive [³H] rhamnose (5 Ci/mmol) was purchased from American Radiolabeled Chemicals Ltd. (St. Louis, Mo.). Strains were grown overnight in 5-ml TY cultures and then subcultured into 5 ml of VMM–glycerol-rhamnose. Cells were harvested by centrifugation (5,000 x g for 10 min) at mid-log phase (OD 600 0.5 to 0.7), washed twice in VMM-

glycerol, then resuspended in defined salts medium to a final OD 600 of 0.3 in a total volume of 2 ml. Transport assays were initiated by the addition of [³H] rhamnose to a final concentration of 2 μM, and aliquots of 0.5 ml were withdrawn at appropriate time points and rapidly filtered through a Millipore 0.45-μm Hv filter on a Millipore sampling manifold. Filtered cells were immediately washed with 5 ml of defined salts medium to wash away unincorporated label. Label incorporation was quantified using a liquid scintillation spectrophotometer (Beckman LS6500). Samples were taken every 20 s and continued for up to 2 min. Transport rates were generally linear over the first minute of the assay. Longer term accumulation assays were carried out in a similar fashion except incubation times with label were extended.

2.2.7 Biochemical assays

An assay for rhamnose kinase activity was developed by adapting a fructose kinase assay that has been previously described (Anderson and Sapico 1975, Geddes and Oresnik 2012). Briefly *R. leguminosarum* cells were grown and cell-free lysates were prepared as previously described (Oresnik and Layzell 1994, Poysti, Loewen et al. 2007, Pickering and Oresnik 2008). The assay was initiated with the addition of rhamnose. Rates were determined from linear portions of NADH oxidation data that were corrected for background oxidase activity and were proportional to the amount of extract assayed. Malate dehydrogenase assays were carried out as previously described (Finan, Oresnik et al. 1988).

2.2.8 Isolation of cell lysate membrane fraction

Strains were grown on defined medium to mid-log phase, harvested and cell free extracts were prepared. Membranes were isolated by subsequent ultracentrifugation

(Beckman Coulter MaxE) of the cell free extracts using a TLA 100.3 rotor (45,000 x g for 30 min at 4°C). Membrane pellets were subsequently resuspended in extraction buffer. Isolated membrane fractions were routinely monitored for cytoplasmic cross contamination by assaying for malate dehydrogenase activity.

2.2.9. Western blots

Cell fractions were analyzed by SDS-PAGE (Laemmli 1970) and transferred to a nitrocellulose membrane for Western blot analysis using a His6 monoclonal antibody as a primary antibody and a goat-anti-mouse secondary antibody linked to HRP. HRP was detected colourmetrically with an opti-4CN substrate detection kit (Bio-Rad Laboratories) as previously described (Laemmli 1970, Richardson, Carpena et al. 2008).

2.2.10 Bioinformatic analysis and molecular modelling RhaK

Basic protein alignments were carried out using CLUSTAL-X (Thompson, Gibson et al. 1997) and Praline (Simossis, Kleinjung et al. 2005). Structural models of RhaK were constructed by using PHYRE² (Kelley and Sternberg 2009). The final scaffold that was used to model RhaK was gluconate kinase from *Lactobacillus acidophilus* (PDB file 3GBT). Final figures were generated in PyMol (Schrödinger 2012).

2.3 Results

2.3.1 RhaK can associate with the cytoplasmic membrane

Strains lacking the kinase RhaK were shown to be unable to transport rhamnose (Richardson, Hynes et al. 2004, Richardson and Oresnik 2007). The inability to transport

rhamnose was not due to the lack of expression or translation of the ABC transporter RhaSTPQ (Richardson and Oresnik 2007). Conversely, insertional mutants lacking the ABC transporter components RhaTPQ, the mutarotase RhaU, and RhaK did not affect rhamnose kinsase activity when *rhaK* was expressed from a plasmid. It was hypothesized that if RhaK affected rhamnose transport, it may interact with either the cytoplasmic membrane or the components of the ABC transporter.

To address this hypothesis two RhaK variants with an RGS-His6x tag on either the C or N terminal were generated such the localization of RhaK could be followed by Western blot analysis. Each variant was capable of complementing *rhaK* mutants for growth using rhamnose as a sole carbon source and thus transport as well. Western blot analysis using a monoclonal antibody to the RGS-His6X epitope resulted in a single band of approximately 55 kdal (Figure 2.1). Cell lysates were further separated by ultracentrifugation into cytoplasmic and membrane fractions. Assaying these fractions for malate dehydrogenase activity showed that the membrane fraction did not contain detectable activity, suggesting that the membrane fractions were devoid of cytoplasmic contamination.

The data showed that RhaK could be found in the membrane fraction (Figure 2.1). Since rhamnose kinase activity is directly related to the amount of RhaK detected by Western Blot analysis, RhaK activity was used to determine the proportion of RhaK that was associated with each of the fractions (Figure 2.1, panel A). When the proportion of the total RhaK activity of associated with each fraction was calculated it was found that about 7% of

the total activity could be found associated with the membrane fraction. Similar results were obtained in both a wild-type, as well as a *rhaK* background (Figure 2.1, panel A).

To determine how tightly RhaK was associated with the membrane, isolated membrane fractions were subjected to salt washes. Whereas a 50 mM NaCl wash did not appear to give results that were qualitatively different from an unwashed membrane, the majority of the RhaK signal was lost with a 150 mM wash. This suggests that a proportion of RhaK can interact with the cytoplasmic membrane in a manner that is consistent with it being a peripheral membrane protein (Figure 2.1, panel B).

The ABC rhamnose transporter in *R. leguminosarum* consists of the periplasmically localized solute binding protein RhaS, the integral membrane proteins RhaP and RhaQ that are the permease components, and the cytoplasmic protein RhaT, which is the ABC component. Since an absence of RhaK results in a lack of transport activity, we hypothesized that RhaK may be directly interacting with components of the ABC transporter. To address this hypothesis the plasmid pMR178, which contains the tagged version of RhaK, was mated into Rlt106. Rlt106 contains a Tn5 insertion within *rhaT*, thus eliminating the *rhaTPQ* as well as the mutarotase *rhaU* (Richardson, Carpena et al. 2008) and the chromosomal copy of *rhaK*. Western blot analysis of the membrane and cytoplasmic fractions show that the distribution of the His-tagged RhaK in this background is comparable to that seen in both the wild-type background as and the *rhaK* mutant background. This suggests that the RhaK membrane interaction is not dependent on the presence of the rhamnose ABC transporter components, since RhaK localizes to the membrane fraction even in the absence of the transporter components (Figure 2.1, panel C).

Figure 1.

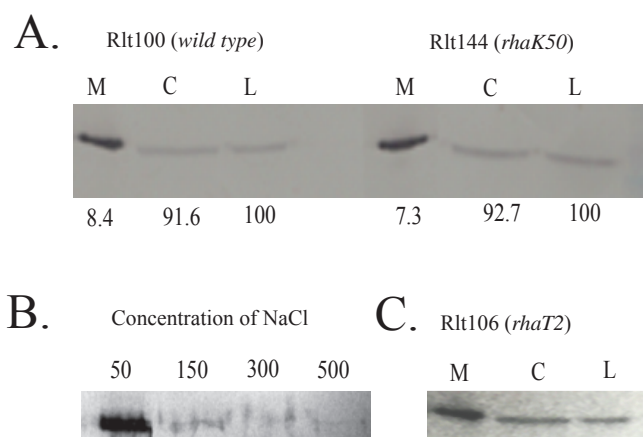


Figure 2.1. Western blot analysis of RhaK in cellular fractions. (A) His-tagged wild-type RhaK expressed in wild-type (Rlt100) and *rhaK50* (Rlt144). Cleared cell lysates (L) were fractionated into membrane (M) and cytoplasmic (C) fractions. Equal volumes each fraction were separated by SDS-PAGE, blotted and detected. The numbers below the each lane indicate proportion of total rhamnose kinase activity from the lysate that was found in each fraction. (B) Salt washes of membrane bound RhaK. Membrane samples were run on SDS-PAGE, blotted and detected. NaCl wash concentrations are indicated above the panel. (C) Western Blot analysis of His-tagged RhaK expressed in Rlt106 (*rhaT2*). Designations are as in panel A.

2.3.2 Generation of *rhaK* mutant alleles

The characterization of RhaU demonstrated that the structure of this protein was consistent with its function as a cytoplasmic rhamnose mutarotase (Richardson and Oresnik 2007). This means RhaU must temporally separate RhaK phosphorylating rhamnose from the role RhaK plays in transport. These data strongly suggests that the ability of RhaK to affect transport and its ability to phosphorylate its substrate can be uncoupled physiologically.

To address this hypothesis an in-frame insertional mutagenesis of RhaK was carried out. The *rhaK* open reading frame was cloned into pLit28i as a *Bam*HI/*Kpn*I fragment yielding pDR1, which was subsequently mutagenized. Putative transposon insertional mutants were screened by restriction to identify those that generated an insert band that was larger than *rhaK* indicating the presence of the transposon. Thirty-nine unique insertional mutants within *rhaK* were generated in this manner. The exact location of each insertion was determined by sequencing. Since there are two *Pme*I sites located in the Tn7, the internal portion of the transposon was removed by cutting with *Pme*I and religating. The resultant construct had a 15 bp insertion within *rhaK*. These insertional mutants were subsequently cut with *Bam*HI and *Hind*III then recloned into pRK7813 (Table 1).

Not including inserts that were found in identical positions, the average spacing between inserts was 12 amino acids (high of 50, low of 1). Ten of the insertions generated stop codons, leaving 29 inserts that generated 5 amino acid inserts. In addition to the 39 inserts presented, 15 insertions in same locations as one of the 39 inserts presented were also isolated (generating identical insertions). We note that the largest gap in our coverage appears to occur between allele *rhaK72* and *rhaK109* near the C-terminus of the protein (Figure 2.2).

2.3.3 Screening of linker scanning *rhaK* alleles

All the alleles that were generated were screened for their ability to complement the *rhaK* strain Rlt144 for growth on rhamnose. In addition, since the build up of a phosphorylated intermediate has also been correlated with the inability to complement (Geddes, Pickering et al. 2010), the alleles were also screened on medium containing rhamnose and glycerol (Table 2.2). The results show that 25 of the insertions were not affected in their ability to complement Rlt144 (Table 2.2) and thus were not further examined. Thirteen of the remaining insertions were unable to complement Rlt144 for growth using rhamnose as a sole carbon source. Eleven of these insertions were predicted to generate stop codons that would yield a truncated protein. The two remaining alleles, *rhaK73* and *rhaK111* were of particular interest for further analysis.

A single insert, *rhaK72*, was initially scored as unable to complement Rlt144 following 3 days of incubation. It was found that colonies would appear after extended periods of growth (6 days). To determine whether this growth was due to second site mutations arising in the culture, single colonies were isolated, purified and tested for re-growth. It was found that upon re-streaking these colonies exhibited the same slow growth phenotype suggesting the growth that was exhibited was due to the *rhaK72* allele and not a second site mutation.

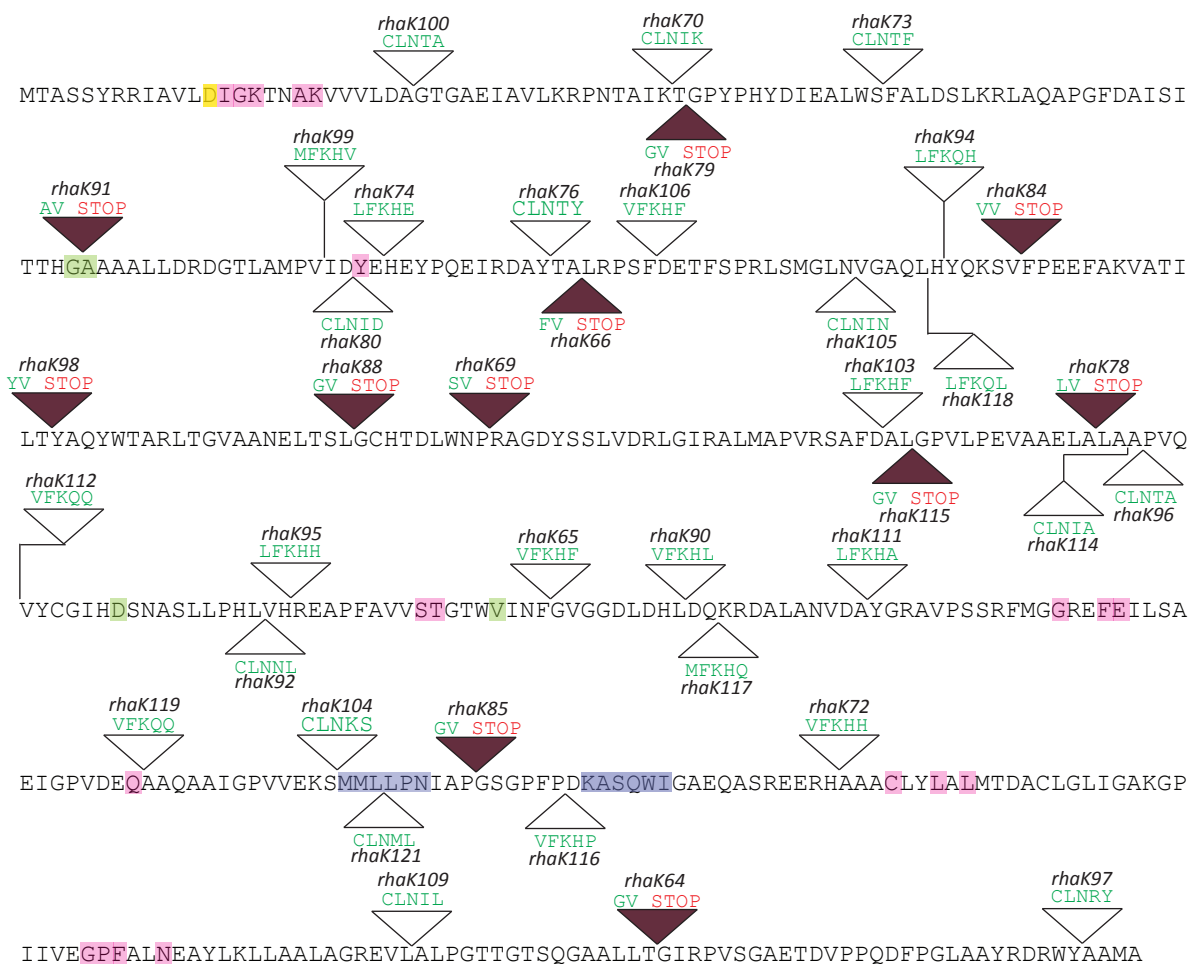


Figure 2.2. Linker scanning mutagenesis. Sites and sequences of mutants generated. Insertion sites are denoted by triangles. Filled triangles denote inserts that generate truncations. The conceptual translation generated by each insert is shown on the top of the triangle. Highlighted amino acids denote residues deduced to play a role in either ATP binding (Red), Mg²⁺ binding (Yellow), substrate binding (green), or are predicted to be a dimerization domain (blue) based on other related carbohydrate kinases.

Table 2.2. Complementation analysis of *rhaK50* using *rhaK* linker scanning alleles

Strain	Relevant genotype ^a	Rhamnose	Rhamnose/Glycerol	Glycerol
Rlt100	Wild-type	+	+	+
Rlt144	<i>rhaK50</i>	-	+	+
Rlt144 (pDR22)	<i>rhaK66</i>	-	+	+
Rlt144 (pDR70)	<i>rhaK65</i>	+	+	+
Rlt144 (pDR72)	<i>rhaK70</i>	+	+	+
Rlt144 (pDR73)	<i>rhaK72</i>	+/-	+	+
Rlt144 (pDR74)	<i>rhaK73</i>	-	+	+
Rlt144 (pDR77)	<i>rhaK88</i>	-	+	+
Rlt144 (pDR78)	<i>rhaK74</i>	+	+	+
Rlt144 (pDR79)	<i>rhaK91</i>	-	+	+
Rlt144 (pDR80)	<i>rhaK84</i>	-	+	+
Rlt144 (pDR81)	<i>rhaK76</i>	+	+	+
Rlt144 (pDR82)	<i>rhaK78</i>	-	+	+
Rlt144 (pDR144)	<i>rhaK69</i>	-	+	+
Rlt144 (pDR146)	<i>rhaK79</i>	-	+	+
Rlt144 (pDR147)	<i>rhaK85</i>	-	+	+
Rlt144 (pDR155)	<i>rhaK92</i>	+	+	+
Rlt144 (pDR157)	<i>rhaK95</i>	+	+	+
Rlt144 (pDR158)	<i>rhaK97</i>	+	+	+
Rlt144 (pDR159)	<i>rhaK114</i>	+	+	+
Rlt144 (pDR160)	<i>rhaK115</i>	-	+	+
Rlt144 (pDR161)	<i>rhaK96</i>	+	+	+
Rlt144 (pDR162)	<i>rhaK104</i>	+	+	+
Rlt144 (pDR163)	<i>rhaK103</i>	+	+	+
Rlt144 (pDR164)	<i>rhaK109</i>	+	+	+
Rlt144 (pDR165)	<i>rhaK116</i>	+	+	+
Rlt144 (pDR166)	<i>rhaK111</i>	-	+	+
Rlt144 (pDR167)	<i>rhaK117</i>	+	+	+
Rlt144 (pDR168)	<i>rhaK119</i>	+	+	+
Rlt144 (pDR169)	<i>rhaK121</i>	+	+	+
Rlt144 (pDR182)	<i>rhaK64</i>	-	+	+
Rlt144 (pDR183)	<i>rhaK90</i>	+	+	+

Rlt144 (pDR184)	<i>rhaK94</i>	+	+	+
Rlt144 (pDR185)	<i>rhaK98</i>	-	+	+
Rlt144 (pDR186)	<i>rhaK99</i>	+	+	+
Rlt144 (pDR187)	<i>rhaK100</i>	+	+	+
Rlt144 (pDR188)	<i>rhaK112</i>	+	+	+
Rlt144 (pDR189)	<i>rhaK118</i>	+	+	+
Rlt144 (pDR191)	<i>rhaK105</i>	+	+	+
Rlt144 (pDR192)	<i>rhaK106</i>	+	+	+
Rlt144 (pDR193)	<i>rhaK80</i>	+	+	+

Complementation was scored by judging growth on defined medium of Rlt144 carrying each of the alleles. Growth is as follows: +, wild-type; +/-, slow growth; -, no growth.

^a Rlt144 carries the *rhaK50::Tn5-B20*. For clarity, only the allele being tested is listed.

2.3.4 Biochemical properties of linker scanning alleles

We hypothesized that with respect to affecting rhamnose transport and kinase activity, 4 types of mutant alleles should be possible. These are, presence of both activities, absence of both activities, presence of transport and the lack of kinase activity, and finally presence of kinase and lack of transport activity. Each of the remaining alleles was introduced into Rlt144 and screened for the ability to confer transport and kinase activity. The majority of these alleles did not show any appreciable differences from Rlt144, or were obviously non functional because the insertion generated a stop codon (Figure 2.2). Thus, representative alleles from these subsets were chosen for further analysis.

The alleles chosen were as follows; two insertions (*rhaK64* and *rhaK85*) that represent the longest truncated versions of RhaK, and three insertions (*rhaK70*, *rhaK92*, and *rhaK97*) from the N- terminal, middle, and C- terminal region of RhaK respectively that were indistinguishable from the wild-type based on complementation analysis. The remaining alleles (*rhaK72*, *rhaK73*, and *rhaK111*), had complementation phenotypes that were consistent with affecting kinase and/or transport activity.

The results show that alleles that encode premature stop codons did not confer either kinase or transport activity (Figure 2.3). We note that although *rhaK64* does not confer either transport or kinase activity, insertions that do not generate stop codons in this area (*rhaK109* and *rhaK97*) do not impair RhaK function (Table 2.2, Figure 2.3). Not surprisingly, alleles that retained the ability to complement a *rhaK* mutation retained kinase activity and conferred the ability to transport rhamnose near wild-type levels (Figure 2.3).

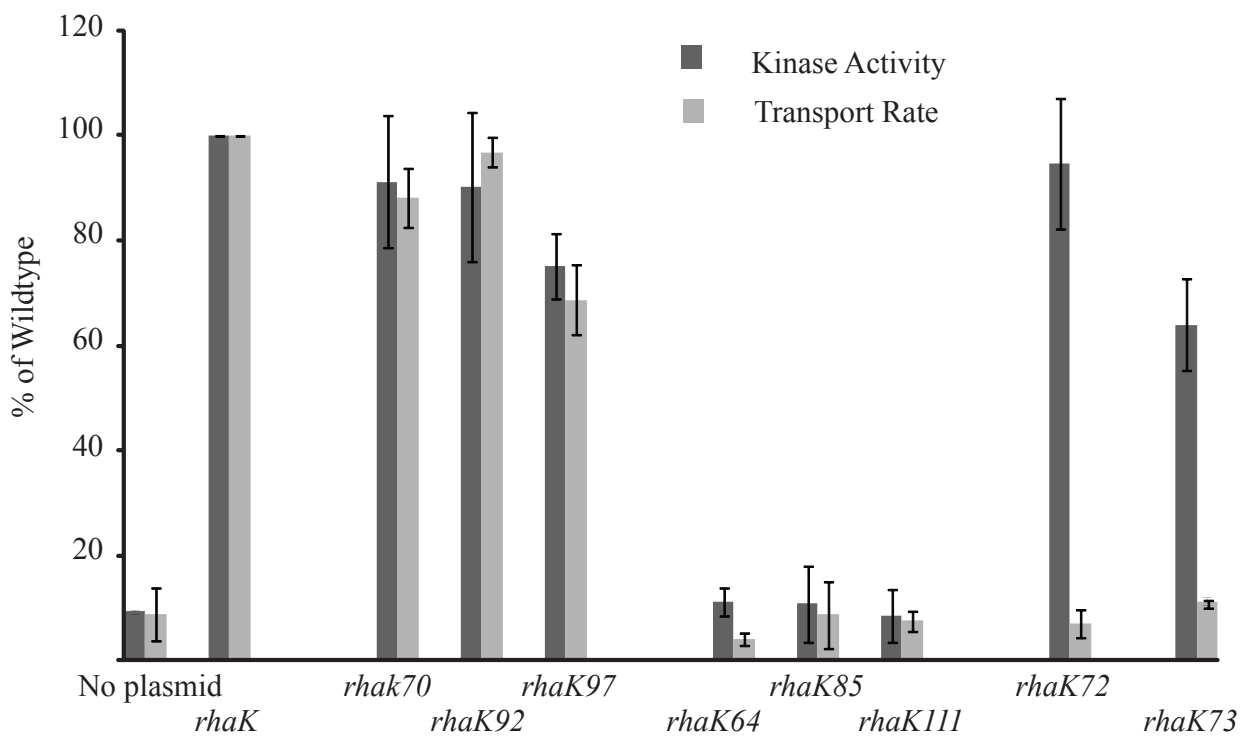


Figure 2.3. Transport and kinase rates of representative *rhaK* alleles. RhaK variants expressed from plasmids in Rlt144 were assayed for transport and kinase activities. Whole cell assays were used to measure uptake of radio-labelled rhamnose whereas cleared lysates were used to measure kinase activity. Transport and kinase rates are presented as a percent of the wild-type. Error bars represent standard deviation on the mean of at least 3 biological replicates. Typical wild type kinase rates were approximately 600 μ moles/min/mg. Typical wild type transport rates were about 10 nmoles/min/mg.

Whereas the *rhaK111* allele was unable to confer the ability to transport rhamnose or kinase activity, *rhaK73* had appreciable rhamnose kinase activity (Figure 2.3). Assays of Rlt144 carrying the *rhaK72* showed that this allele encoded a variant that had kinase activity that was not significantly different from the wild-type, but it could not confer the ability to transport rhamnose over background levels (Figure 2.3). Western blot analysis using His-tagged versions of *rhaK72* showed that it localized identically to the wild-type.

2.3.5 Alleles that uncouple transport and kinase activity have different complementation phenotypes

rhaK72 and *rhaK73* both uncouple the ability to confer transport and kinase activity (Figure 3). In an attempt to gain new insight in to how these alleles might be affecting the ability of the cell to utilize rhamnose, a growth experiment was carried out to quantitate growth. Rlt100 growing on defined medium using our growth conditions typically has a doubling time of between 8 to 10 hours (Table 2.3, (Richardson, Carpena et al. 2008)). If Rlt144 carrying the *rhaK50* allele is inoculated into defined medium with rhamnose as a sole carbon source no measurable doubling occurs. Rlt144 complemented with the wild-type *rhaK* on a plasmid, clearly grows; however it is not unusual to note a slightly longer doubling time (Table 2.3).

Introduction of either *rhaK72* or *rhaK73* on a multicopy plasmid into Rlt144 shows significantly different growth when compared with the strain carrying the wild-type *rhaK* allele or each other. Whereas the doubling time of Rlt144 carrying the *rhaK72* allele was

Table 2.3. *rhaK72* and *rhaK73* alleles confer different growth rates when expressed in Rlt144 (*rhaK50*::Tn5)

Strain	Relevant Genotype (Chromosomal/plasmid)	Doubling time ^a (h) ± SD
Rlt100	Wild-type	9.5 ± 0.9
Rlt144	<i>rhaK50</i> / -	NG ^b
Rlt144 (pMR110)	<i>rhaK50</i> / <i>rhaK</i> ⁺	12.7 ± 0.1
Rlt144 (pDR73)	<i>rhaK50</i> / <i>rhaK72</i>	21.1 ± 1.1
Rlt144 (pDR74)	<i>rhaK50</i> / <i>rhaK73</i>	NG

Doubling times (hrs) were calculated from cultures growing in defined media over an 16 hour time interval. All cultures were in mid-log phase over the entire time interval. Numbers represent the mean ± S. D. of 3 independent cultures.

^a Doubling time = $\ln(2)/((\ln(N_2/N_1))/T)$ Where N_1 = initial OD₆₀₀, N_2 = final OD₆₀₀, and T= time (hrs)

^b No detectable growth.

calculated to be about 21 hours, the doubling time could not be calculated from Rlt144 carrying the *rhaK73* allele because there was very little growth (Table 2.3).

Initial rhamnose transport assays that were carried out with *rhaK72* or *rhaK73* in Rlt144 did not show transport activity that was over the level of background in a 2 minute assay (Figure 2.3). To determine if residual rhamnose uptake could be detected, extended incubations of Rlt144 carrying *rhaK72* and *rhaK73* were carried out. Consistent with the growth data, Rlt144 carrying *rhaK72* was able to accumulate about 59 nmol/mg protein whereas Rlt144 carrying *rhaK73* accumulated approximately 6 nmol/mg protein (Figure 2.4). Both of these values is significantly greater than the value that was determined for Rlt144 (Figure 2.4).

2.3.6 Molecular architecture of a RhaK

Bioinformatic, genetic, and biochemical data support that RhaK is a sugar kinase. Our initial analysis was capable of identifying a limited number of domains and motifs associated with RhaK (Figure 2.2). Briefly, it was shown that RhaK belonged to a family of FGGY-type kinases, that a conserved P-loop motif used for the binding of ATP was found close to the N-terminus, and that based on CLUSTAL and PRALINE alignments to the *E. coli* GlpK (Hurley, Faber et al. 1993, Hurley 1996), the active site residue necessary for the catalysis was D248 (Richardson and Oresnik 2007).

The generation of linker scanning insertions that uncoupled the proteins kinase and transport activities necessitated a more in-depth analysis of the protein to develop hypotheses regarding how RhaK might have such disparate activities. Initial experiments focused on over-expressing His-tagged variants in *E. coli* with the ultimate goal of generating a

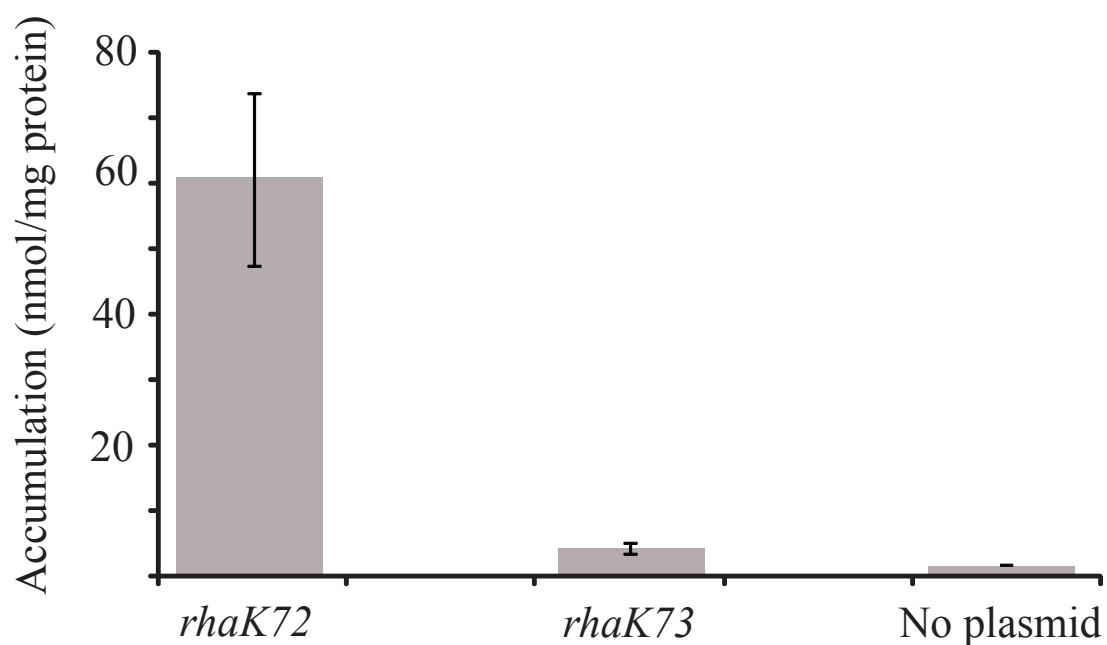


Figure 2.4. Accumulation of rhamnose in Rlt144 carrying either *rhaK72* or *rhaK73*. Rlt144, or Rlt144 carrying either *rhaK72* or *rhaK73*, were incubated with labelled rhamnose for sixty minutes. Whole cell assays were used to measure uptake of radio-labelled rhamnose. Data are presented as nmol rhamnose accumulated/mg protein. The corresponding wild-type transport rate for this experiment was 10.2 ± 0.8 nmoles/mg/min. Error bars represent standard deviation on the mean of 3 biological replicates.

experimentally determined structural model. These experiments were complicated by the number of rare arginine codons that are found in RhaK. Expressing *rhaK* in either *R. leguminosarum* or *S. meliloti* circumvented some of the low yield issues that were encountered; however the isolated protein did not show *in vitro* activity although the constructs were capable of complementing strains carrying a *rhaK* mutation. To gain insight into how the insertion alleles may be changing RhaK we instead opted to generate a model of RhaK using the PHYRE² server.

The PHYRE² server predicts the secondary structure of the query protein, and matches this pattern to that of proteins that have been crystallized (Kelley and Sternberg 2009). The output data modelled RhaK to 20 different crystallized proteins. The top 19 ranked hits were all carbohydrate kinases. The percent sequence identity varied from 14 to 39%. The typical alignment coverage was 97%, and the confidence level for the predicted model was 100% in each case. The top hit was a gluconate kinase from *Lactobacillus acidophilus* (pdb c3gbtA), with 98% alignment coverage, 15% amino acid identity, and 100% confidence level in its prediction. Visual analysis of the 19 models that were generated to the carbohydrate kinases, showed a high degree of similarity in the overall structure. A common anomaly that we note is a “floating helix” at the C- terminus of the protein. This type of anomaly appears to be a result of the difference in lengths of the protein being modelled and the protein to which it is being scaffolded. Since there was general agreement with all the models, we chose to use the model that was generated by scaffolding RhaK onto the top hit.

The predicted model shows a two lobed protein (Figure 2.5, panel A). The active site residue is found at the base of a long cleft, the ATP binding site and substrate binding sites

are also found along the N- terminal lobe that makes up the active site cleft. Other notable features include a pair of beta sheets that act as a likely dimerization domain (Di Luccio, Petschacher et al. 2007). Moreover the entire shape and structure of the model predicts a carbohydrate kinase that has a great number of conserved structural features. Most notably there is a high degree of agreement on which amino acids appear to be surface exposed (Figure 2.5, panel A).

The RhaK72 insertion occurs immediately following A363 in RhaK, and generates a five amino acid insertion (VFKHH, Figure 2.2). The region from R357 to I380 in RhaK is predicted with a high degree of confidence (9 on scale of 9) to form an alpha helical region. The inclusion of the 5 amino acids in the RhaK72 variant within this helix is not predicted to disrupt the helix since the secondary structure prediction across this area still maintains a high confidence in the probability of an alpha helix forming. The net result is a helix of 23 amino acids in the RhaK72 variant whereas the wild-type contains a 24 amino acid alpha helix. Modelling this into a three dimensional space predicts that this is a helix that leads to a surface exposed loop. Overlaying the predicted RhaK72 structure with the wild-type suggests that this insertion leads to a change in the surface loop region of the variant (Figure 2.5, panel B).

The RhaK73 insertion occurs immediately following W56 in RhaK, and generates the five amino acid insertion CLNTF (Figure 2.2). The region I52 to A69 of RhaK are predicted with high confidence levels to form a 17 amino acid alpha helix. With the inclusion of the 5 amino acids generated from the insertion, the secondary structure in this area is still predicted with high confidence levels to be an alpha helix. Placement of this helix in the RhaK model

predicts this to be a surface exposed helix that is found in the N- terminal lobe of RhaK (Figure 2.5, panel C).

Taken together, mapping of both inserts that have uncoupled kinase activities from the transport of rhamnose into the cell map to two distinct regions of the protein. It is noteworthy that the predicted changes for both of the insertions results in changes on the same protein face (Figure 2.5, panel D), suggesting perhaps this protein face plays a role in the protein-protein interaction of RhaK with other protein(s).

Based on the generated model and the positioning of the Rhak72 insertion it was noted that closely related RhaK orthologs contained a conserved amino acid sequence (R358 E369 E360 R361). To test if this region is directly involved in a protein-protein interaction, amino acids 358 to 361 were replaced with alanines yielding allele *rhaK122*. Testing of this construct showed that the growth phenotype of Rlt144 carrying *rhaK122* was indistinguishable from that of the wild type. In addition, kinase and transport rates were also comparable to the wild type, indicating no significant effect was generated by this substitution .

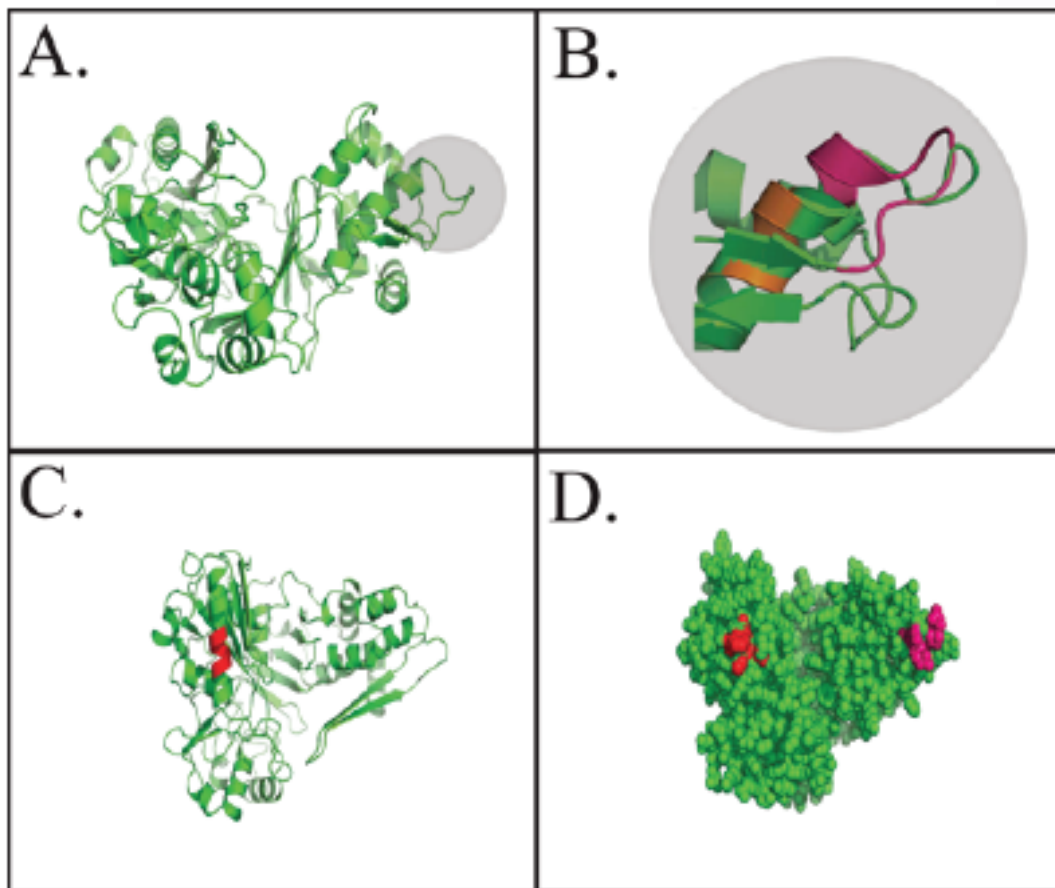


Figure 2.5. Predicted structure of wild type RhaK and variants. Structural models were generated using PHYRE². The RhaK amino acid sequence was scaffolded to gluconate kinase from *Lactobacillus acidophilus* (PDB c3gbtA). (A) Overall predicted structure of RhaK. (B) RhaK72 variant model superimposed onto RhaK. The Region highlighted in A, corresponds to the region depicted in B. Orange ribbon depicts the amino acid insertion, whereas the pink depicts regions where the structure differs. (C). RhaK73 variant model superimposed onto RhaK (top view relative to panel A). Red ribbon depicts the amino acid insertion. (D) Composite model (top view relative to panel A) depicting surface residues that are predicted to be changed in RhaK73 (red) and RhaK72 (pink).

2.4 Discussion

In this paper we have initiated a characterization of how RhaK affects transport of rhamnose into the cell. Two hypotheses regarding how RhaK may affect transport were previously put forward (Richardson and Oresnik 2007). The first, is that rhamnose itself may be a negative regulator that affects its own ABC transporter. The second, is that RhaK, in addition to the known catabolic activity, also affected the activity of the ABC transporter. The isolation of two alleles of *rhaK* that retain kinase activity and have lost the ability to confer *in vivo* transport demonstrates that these activities can be uncoupled (Figure 2.3). Moreover these data provide direct evidence that the mechanism by which RhaK affects rhamnose transport in *R. leguminosarum* is not due to a negative regulatory effect of the rhamnose on its own ABC transporter.

In an attempt to demonstrate direct interaction of RhaK with its transporter we were able to demonstrate that a small proportion of RhaK that had rhamnose kinase activity could be found loosely associated with the membrane fraction (Figure 2.1). We note that this association was not dependent on the presence of the ABC transporter components, and that this was only tested with *rhaK* expressed from a low copy number plasmid. The significance of this result is not clear at present and may be resolved with further study.

Prior to this work, the only RhaK variant that was constructed was in the P-loop region which is used for the binding of ATP (Richardson and Oresnik 2007). These variants eliminated both transport and kinase activity (Richardson and Oresnik 2007). Assuming that ability to affect kinase and transport were independent, we had anticipated isolating four mutant classes from our mutagenesis of *rhaK*; these were variants that had neither activity affected, variants that had both activities affected, and variants that had one or the

other affected. The results of our analysis show that we did not isolate alleles that encoded variants that retained the ability to transport but lost kinase activity. Although we achieved our anticipated coverage of the open reading frame following our mutagenesis, we note that a number of our inserts had generated stop codons leaving the possibility that this class of variant may still exist. Alternately, it could be possible that the ability to affect transport may be linked to the ability of RhaK to bind ATP or carry out kinase activity. It is not unprecedented for protein-protein interactions to be dependent on phosphorylation state (Hurley, Faber et al. 1993, Hurley 1996).

The alleles *rhaK72* and *rhaK73* confer distinct phenotypes when they are used to complement a *rhaK* mutant (Table 2.2, Table 2.3, Figure 2.3, Figure 2.4). It could be that these two mutations affect two separate interactions or that both alleles affect a single protein-protein interaction. Although the growth differences are clear (Table 2.3), it could be that the *rhaK72* allele allows some residual transport activity that is not above the level of detection in our assays. We do not have data that would allow us to favour either of these scenarios at this time.

Two alleles, *rhak72* and *rhak73* encoded variants that retained kinase activity but were unable to complement transport activity *in vivo* (Figure 2.3). Mapping of these inserts showed that the 5 amino acid insertion (CLNTF) insertion that defined *rhaK73* was following W56, whereas the *rhaK72* insertion (VFKHH) was closer to the C-terminal quarter of the protein after H362. Aligning of RhaK to closely related kinases using CLUSTALX (Thompson, Gibson et al. 1997) and Praline (Simossis, Kleinjung et al. 2005) was useful in identifying residues that were involved in substrate binding and catalysis when compared to the well studied glycerol kinase (GlpK) (Pettigrew, Smith et al. 1998). Both *rhaK72* and

rhaK73 insert 7 and 8 amino acid residues prior to conserved leucine residues. Structural modelling of RhaK using the PHYRE² server produced a model that had 97% coverage and 100% confidence levels when compared to GlpK. Mapping of these two insertional variants suggest that both appear to be in regions that are strongly predicted to be helices and that the resulting changes appear as changes on the same surface of the protein. Further construction and testing of the “REER” motif that is conserved in closely related RhaK orthologs however did not change RhaK’s ability to affect rhamnose transport. Whereas the insertional variant *rhaK72* moved residues further apart, the *rhaK122* allele targeted a group of charged amino acids and changed these to uncharged neutral amino acids. It could be that amino acid spacing rather than charge may be more important for the function of this area. Alternately, this particular region is not important to transport activity. These hypotheses are currently being addressed.

This work was carried out to test the hypothesis that kinase and transport activities are affected by RhaK. Isolation of *rhaK72* and *rhaK73* alleles of *R. leguminosarum rhaK* clearly show that these activities can be separated. Our working hypothesis is that RhaK interacts via protein-protein interaction to affect the rhamnose ABC transporter. It is not clear whether this will occur as a direct or indirect interaction. Precedent for interacting proteins affecting transport exist. In particular MalK has been shown to interact both with EIIA^{glc} (Samanta, Ayvaz et al. 2003), and the maltose regulator MalT has been shown to interact with, and be sequestered at the membrane by MalK when it is complexed with its permease components (Boos and Böhm 2000, Richet, Davidson et al. 2012).

Our work is now focusing on identifying proteins that may interact with RhaK to affect rhamnose transport. In addition, we have embarked on generating a crystal structure

for RhaK to be used for more detailed structure function studies to define the motif(s) on RhaK that are used to affect rhamnose transport by the CUT2 ABC transporter RhaSTPQ.

Chapter 3

The sugar kinase (RhaK) that is necessary for the catabolism and transport of rhamnose in *Rhizobium leguminosarum* directly interacts with the ABC transporter component (RhaT).

3.1 Introduction

Cells are separated from the external environment by a selectively permeable membrane. Because of this, a means to transport physiologically relevant substrates across the membrane is required. (Higgins 1992, Dassa and Bouige 2001, Davidson and Chen 2004, Rea 2007). The largest and most widely distributed family of transporters are the ATP binding cassette (ABC) transporters (Higgins 1992). ABC transporters are responsible for the transport of a very diverse set of substrates, and can be divided into two main classes; importers and exporters (Higgins 1992, Dassa 2003, Davidson and Chen 2004, Rea 2007). Much of our understanding of their structure, mechanism, and kinetics comes from the study of a few model transporters.

Gram negative ABC importers typically consist of two membrane spanning permeases (encoded by either one or two genes), a periplasmic substrate binding protein, and two copies of the ABC-type ATPase protein (Boos and Shuman 1998, Boos and Böhm 2000, Schneider 2001, Eitinger, Rodionov et al. 2011). The CUT1 class maltose transporter from *E. coli* has served as a model for Gram negative ABC-type importers (Boos and Shuman 1998, Schneider and Hunke 1998, Schneider 2001, Chen, Lu et al. 2003, Oldham, Khare et al. 2007, Khare, Oldham et al. 2009). Carbohydrate uptake transporters, however, are divided into two main groups: carbohydrate uptake transporter 1 and 2 (CUT 1 and CUT 2). This division is based on the ATPase protein. The CUT 2 ABC protein is typically about 200 amino acids larger than the CUT1 family counterpart and appears to have arisen from a fusion of two ABC domains. These ABC proteins contain two ATP binding sites (Schneider 2001); however, a functional substitution of a key lysine residue in the C-terminal Walker A motif likely leads to the loss of ATP hydrolysis function in this second domain (Schneider

2001). Unlike the CUT1 class transporters, relatively little evidence is available to help elucidate the mechanism and arrangement of the CUT2 class transporters.

In *Rhizobium leguminosarum* genes encoding the determinants necessary for the transport and catabolism of rhamnose are arranged in two transcripts that are regulated by a negative regulator (RhaR) (Richardson, Hynes et al. 2004). The catabolism is carried out by a mutarotase (RhaU), an isomerase (RhaI), a kinase (RhaK), and a dehydrogenase/aldolase (RhaD) (Richardson and Oresnik 2007, Richardson, Carpena et al. 2008). The transport of rhamnose is dependent on a CUT 2 type ABC transporter system encoded by RhaSTPQ ((Richardson, Hynes et al. 2004, Richardson and Oresnik 2007, Rivers and Oresnik 2013) Chapter 2). Functional characterization of this locus revealed that the transport of rhamnose as carried out by its associated ABC transporter was affected by mutations in the gene *rhaK*. The absence of *rhaK* did not affect either the transcription, translation, or membrane localization of the components of the ABC transporter (Richardson and Oresnik 2007). Using an in-frame insertional mutagenesis strategy it was shown that the function of RhaK as a sugar kinase and its role in transport could be genetically separated (Rivers and Oresnik 2013).

Homology models of RhaK indicate it has a typical sugar kinase structure. It contains N-terminal and C-terminal lobes separated by a central active site cleft. The two insertional alleles capable of uncoupling transport function from kinase function mapped to the N and C terminal lobes respectively. Based on these models the inserts appear localized to the same protein face (Figure 2.5)(Chapter 2). Taken together these observations suggests the possibility that RhaK is capable of interacting with another protein along this face, and that this interaction affects the ability of RhaSTPQ to transport rhamnose (Chapter 2). In an effort

to determine if RhaK is interacting with components of the rhamnose transporter, Western blots of RhaK were carried out in a wild-type background, and in a genetic background where the components of the transporter are not expressed (Figure 2.1) . A small proportion (7%) of RhaK was found associated with the membrane fraction in both backgrounds. This indicates that the membrane association was not dependent upon the presence of the transporter (figure 2.1). The way these experiments were carried out required the over-expression of *rhaK*. Therefore it was not clear if the result was physiologically relevant, and we could not eliminate the hypothesis that RhaK directly interacts with components of the rhamnose transporter. Since the basic architecture of an ABC transporter is well characterized (Higgins and Litton 1992 , Quentin and Fichant 2000, Biemans-Oldehinkel, Doeven et al. 2006) it is likely that RhaT is the only component physically available to interact directly with RhaK. In this chapter we address the hypothesis that RhaK directly interacts with the ABC protein RhaT.

3.2 Materials and Methods

3.2.1 Bacterial strains, plasmids, and media

The bacterial strains and plasmids used and generated in the work are listed in Table 3.1. Bacterial strains used were routinely grown at 30°C on complex (either LB or TY) or defined (VMM) medium (Vincent 1970, Beringer 1974, Sambrook and Russell 2001). Defined media were modified as previously described (Oresnik, Pacarynuk et al. 1998). *E. coli* strains were routinely grown at 30°C when in broth cultures and at 37°C on agar medium. When required antibiotics were used at the following concentrations: ampicillin

Table 3.1 Strains and plasmids.

Strains and Plasmids	Relevant genotype	Reference or source
Strains		
DH5α	<i>endA hsdR17 supE44 thi-1 recA1 gyrA96 relA1 Δ(argF lacZYA) U169 φ80dlacZΔM15</i>	B.R.L. Inc
BTH101	<i>cya⁻</i>	Karimova et al. 1998
Plasmids		
pW3C1	Rhamnose locus contained in pRK7813	Oresnik et al. 1998
pMR110	<i>rhaK</i> in pRK7813	Richardson and Oresnik, 2007
pDR73	<i>pRK7813</i> with <i>rhaK72</i> expressed from <i>plac</i> promoter	(Rivers & Oresnik, 2013)
pDR74	<i>pRK7813</i> with <i>rhaK73</i> expressed from <i>plac</i> promoter	(Rivers & Oresnik, 2013)
pUT18	T18 fragment of <i>B.pertussis</i> <i>Cya</i>	Karimova et al. 1998
pKT25	T25 fragment of <i>B.pertussis</i> <i>Cya</i>	Karimova et al. 1998
pTM10	<i>rhaK</i> F1 in pKT25	This work
pTM13	<i>rhaK</i> F2 in pKT25	This work
pTM15	<i>rhaK</i> F3 in pKT25	This work
pTM17	<i>rhaK</i> F4 in pKT25	This work
pDR201	<i>RhaT F1</i> in pUT18	This work
pDR202	<i>RhaT F2</i> in pUT18	This work
pDR204	<i>RhaT F3</i> in pUT18	This work
PDR206	<i>RhaT F4</i> in pUT18	This work
pDR211	<i>RhaT F5</i> in pUT18	This work
pDR209	<i>rhaK72</i> F1 in pKT25	This work
pDR210	<i>rhaK73</i> F4in pKT25	This work
Primers		
	Sequence 5'-3'	
<i>rhaK</i> F1-S	GTACCTGCAGGAATGACCGCCAGTTCCTATCG	This work
<i>rhaK</i> F1-aS	GTACAGGGATCCGTCGAAGGAGGGGCGC	This work
<i>rhaK</i> F2-S	GTACCTGCAGGAGCCTATACGGCCTTGCGC	This work
<i>rhaK</i> F2-aS	GTACAGGGATCCCACGGGGGCTGCGAGC	This work
<i>rhaK</i> F3-S	GTACCTGCAGGAGCCGCGGAGCTTGCGC	This work

<i>rhaK</i> F3-aS	GTACAGGGATCCGGAACCGGGGGCGATATTC	This work
<i>rhaK</i> F4-S	GTACCTGCAGGAATGCTGCTGCCGAATATCGC	This work
<i>rhaK</i> F4-aS	GTACAGGGATCCCTATGCCATCGCCGCGTAC	This work
HindRhaTF	ATATAAGCTTGATGAACGCCGCTTTCAACA	This work
RhaT R2	ATATGGATCCTCGTTCTCGGCCACCGTCAG	This work
RhaT F2	ATATAAGCTTCTGACGGTGGCCGAGACC	This work
RhaT R3	ATATGAATCCTCATCGGCGATTTTCGTAAAGCTC	This work
RhaT F3	ATATAAGCTTGGAGCTTTACGAAATCGCCG	This work
RhaT R4	ATATGGATCCTCGCCGAGAATCTCGCCCTT	This work
RhaT F4	ATATAAGCTTGAAGGGCGAGATTCTCGGC	This work
RhaT R5	ATATGGATCCTCATCCAGCCGCTCGGCATA	This work
RhaT F5	ATATAAGCTTGTATGCCGAGCGGCTGGAT	This work
BaMRhaTR	ATATGGATCCTCCGCATTGCCGGTGGCGG	This work

(amp) $100 \mu\text{g ml}^{-1}$; kanamycin (Kan), $50 \mu\text{g ml}^{-1}$. Growth was routinely monitored spectrophotometrically at 600 nm. When X-gal was used, X-gal concentration was $20 \mu\text{g ml}^{-1}$. 1 mM IPTG was used to induce expression of the vectors as necessary.

3.2.2 Clonings, and genetic manipulations

Standard techniques were used for, amplification of DNA by PCR reactions, DNA isolation, restriction enzyme digests, ligations, transformations, and agarose gel electrophoresis (Sambrook and Russell 2001). Nucleotide sequencing was carried out by cycle sequencing using Big-Dye version 3.1, kit. Sequencing reactions were carried out as recommended by the manufacturer and read using an ABI 3130 sequencer as previously described (Chapter 2).

3.2.3 Bacterial two-hybrid analysis

Bacterial two-hybrid analysis was carried out essentially as described (Karimova, Ullmann et al. 2002). Fragments from *rhaT* and *rhaK* were amplified using the primers listed in Table 3.1, and cloned into pUT18 and pTK25 vectors respectively. Final constructs were sequenced to ensure that they were in-frame with the T18 and T25 fragments of adenylate cyclase in pUT18 and pTK25.

3.2.4 Generation of motif variants

Variants of pTM10 (*rhaK* F1) were designed and purchased from GeneScript . The variants were supplied in pUC57 (Table 1). The inserts were each subsequently re-cloned into pKT25 to carry out 2 hybrid assays.

3.2.5 *Beta-galactosidase assays*

Bacterial strains carrying carrying plasmids to be assayed for beta-galactosidase activity were grown to late log phase (generally over night) in the appropriate medium and sub-cultured. Assays were carried out on log phase cells essentially as described previously (Oresnik, Pacarynuk et al. 1998).

3.2.6 *Phylogenetic and bioinformatic analysis*

Basic Protein alignments of RhaK were carried out using Praline or Clustal (Thompson, Gibson et al. 1997, Simossis, Kleinjung et al. 2005). RhaK homology models were constructed using the PHYRE² server (Kelley and Sternberg 2009), using gluconate kinase from *Lactobacillus acidophilus* (Protein Data Bank [PDB] file 3GBT) as previously described (Rivers and Oresnik 2013). Phylogenetic analysis was carried out using the Phylogeny.fr server using default settings (http://www.phylogeny.fr/version2_cgi/index.cgi). Briefly, protein sequences were aligned using MUSCLE V3.7 (Edgar 2004), and subsequently curated using Gblocks to eliminate poorly aligned sequences (Talavera and Castresana 2007). Phylogenies were constructed using PhlML v 3.0. Trees were rendered using TreeDyn (Chevenet, Brun et al. 2006). Consensus sequences of putative protein motifs were generated using WebLogo (Crooks, Hon et al. 2004). Final figures were constructed using Adobe Illustrator as previously described (Geddes, Hauser et al. 2013).

3.3 Results

3.3.1 Two hybrid analysis of *RhaK* and *RhaT*

Briefly, the bacterial two hybrid system is based on “interaction-mediated reconstitution” of adenylate cyclase activity in *E. coli*. The *Bordetella pertussis* adenylate cyclase consists of two domains (T18 and T25) that do not have enzymatic activity when physically separated. These two domains are expressed on separate plasmids and can be fused to proteins of interest. The T18 and T25 fragments may be brought into close proximity if a protein-protein interaction occurs between the proteins of interest. This results in the functional reconstitution of adenylate cyclase activity, and thus cAMP synthesis. The cAMP produced by the reconstituted chimeric enzyme regulates expression of the *lac* operon in a non-reverting adenylate cyclase deficient (*cya⁻*) *E. coli* strain (BTH101). This allows the direct screening for interactions using LB-X-gal media, and quantification of the relative strength of an interaction using beta-galactosidase assays.

Initially, full length *RhaK* and *RhaT* were translationally fused to the T18, and T25 fragments. No evidence of any interaction between *RhaK* and *RhaT* was detected. Since it was likely that the size of the translational fusion was too large to facilitate the reconstitution of the *B. pertussis* adenylate cyclase, a strategy was devised to subdivide both *RhaK* and *RhaT* into smaller functional domains. These fragments were sub-cloned into pUT18 and pKT25 vectors and assayed for interactions.

RhaK was divided into 4 segments based on the secondary structure as predicted by the PHYRE² server, and a homology model previously generated (Figure 2.5)(Rivers and Oresnik 2013). Each predicted alpha helix and beta sheet was contained fully intact in at least one of these fragments. Primers were generated to amplify each of the corresponding

DNA. Each of these RhaK fragments were PCR amplified, and cloned into pKT25 in frame with the T25 fragment. This yielded plasmids pTM10 (RhaKF1), pTM13 (RhaK2), pTM15(RhaKF3), and pTM17 (RhaKF4) (Table 3.1, Figure 3.1), with the largest coding for a 120 amino acid fragment of RhaK.

Similarly, based on the predicted secondary structure and known functional domains associated with ABC proteins, five fragments encompassing the entire length of RhaT were generated and cloned into pUT18 in frame with the T18 fragment. This yielded plasmids pDR201 (RhaTF1), pDR202 (RhaTF5), pDR204 (RhaTF3), pDR206 (RhaTF4), and pDR211 (RhaTF2) (Table 3.1, Figure 3.1). To test for an interaction between fragments of RhaK and RhaT, all possible combinations were co-transformed into BTH101. These were plated onto LB plus X-gal, and incubated at both 37°C and 30°C as suggested (Karimova, Ullmann et al. 2000).

BTH101 grown on LB + X-gal media yields blue colonies when pTM10 or pTM15 is co-transformed with pDR201, indicative of an interaction. The blue colour was more intense when incubated at 37 °C, and thus all further replicates were incubated at 37 °C. The plasmid pTM10 contains the N-terminal region of RhaK fused to the T25 fragment of adenylate cyclase (RhaKF1). The plasmid pTM15 contains the C-terminal region of RhaK fused to the T25 fragment of adenylate cyclase (RhaKF4). The plasmid pDR201 contains the N-terminal portion of RhaT translationally fused to the T18 fragment of adenylate cyclase (RhaTF1). To ensure that these were bona fide positive interactions, the plasmids were subsequently isolated and retransformed into new BTH101 cells. This again resulted in blue colonies on medium containing X-gal. The isolated constructs were re-sequenced to confirm they indeed contained the original fragments. Three representative colonies from each co-transformation

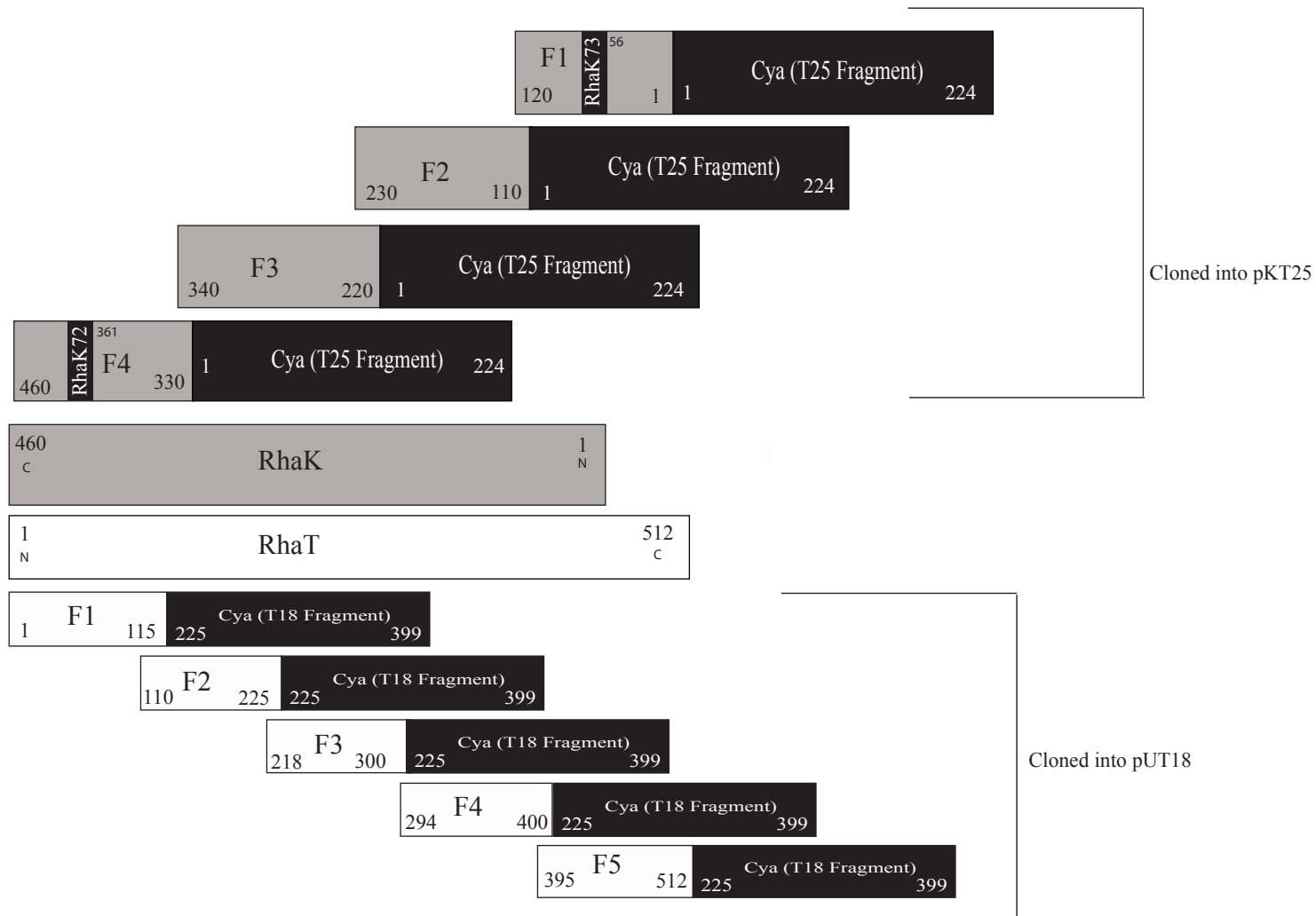


Figure 3.1. A diagram indicating which residues make up each RhaK and RhaT fragment translationally fused to Cya fragments T25 and T18. The amino acid residues included in each fragment are indicated at the bottom, with the N- and C-terminus of the wild-type proteins labelled. The RhaK-CyaT25 fusion proteins have been named RhaKF1-F4 as indicated. The RhaT-CyaT18 fusion proteins have been named RhaTF1-F5 as indicated. The sites of the insertions that define the *rhaK72* and *rhaK73* alleles are indicated

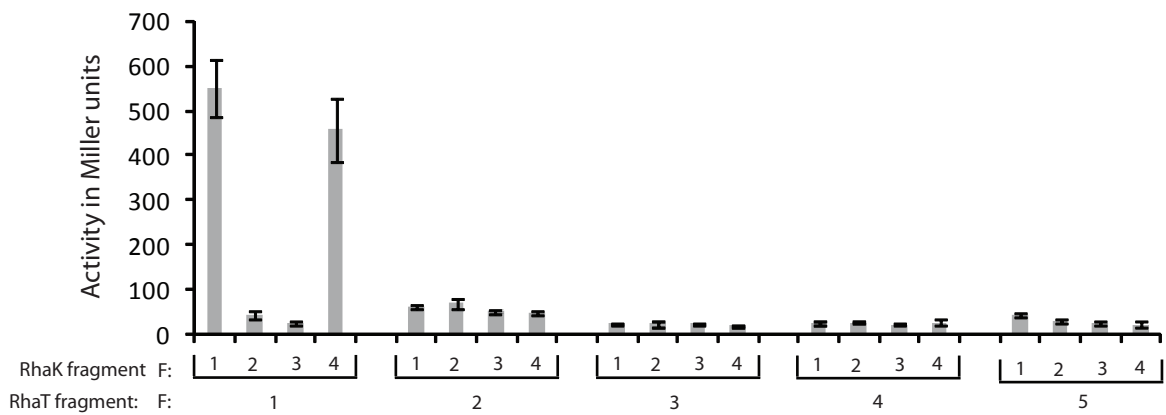


Figure 3.2. Beta-galactosidase activity of BTH101 cells co-transformed with the RhaK and the RhaT fragment indicated. RhaKF1-F4, and RhaTF1-F5 are as defined in Figure 3.1. Growth conditions and assays were carried out as indicated in the material and methods section, with activity reported in miller units. Data represents three biological replicates each assayed in triplicate, with the error bars representing the standard deviation of all replicates.

were selected to quantify this interaction using betagalactosidase assays (Figure 3.2). The results show that each putative positive interaction yielded values of approximately 600 Miller units, whereas the typical background value for the negative control was approximately 30 Miller units (Figure 3.2).

3.3.2 The *rhaK72* and *rhaK73* alleles disrupt two hybrid interactions

Two *rhaK* alleles (*rhaK72* and *rhaK73*) shown to affect rhamnose transport but not kinase activity were previously isolated (Figure 2.3). Interestingly, the in-frame insertions that define *rhaK73* and *rhaK72* map within the segments that correspond to the wild-type N- and C-terminal fragments found in pTM10 and pTM17 respectively. Since it was previously hypothesized that these alleles defined either a direct or indirect point of interaction with RhaT, it was decided these alleles would be tested directly for the ability to interact with RhaT using the two-hybrid system. The primers that were used to generate the RhaKF1 and RhaKF4 fragments flank the *rhaK73* and *rhaK72* inserts. Utilizing the same primers, these regions were PCR amplified with pDR74 (*rhaK73*) and pDR73 (*rhaK74*) as the template. The resultant constructs, pDR209 (RhaK73F1) and pDR210 (RhaK72F1), were transformed into BTH101 strains containing either RhaTF1, RhaTF2, RhaTF3, or RhaTF4.

The results show that when the plasmids carrying RhaKF1 (pTM10) or RhaKF4 (pTM17) were transformed into BTH101 strains carrying RhaTF1 (pDR201), the resultant colonies were blue on medium that contained X-gal. However, when either plasmid pDR209 (RhaK73F1), or pDR210 (RhaK72F4), was transformed into BTH101 carrying RhaTF1 the colonies were white on medium containing X-gal. Similar to what was seen previously, RhaKF1 (pTM10) or RhaKF4 (pTM17) co-transformed into BTH101 with RhaTF1

(pDR201) yield 400-600 Miller units of activity in beta-galactosidase assays (Figure 3.3). However, RhaK73F1 (pDR209) or RhaK72F4 introduced into BTH101 containing RhaTF1(pDR201) yields 132 ± 23 and 63 ± 7 Miller units respectively (Figure 3.3). RhaK73F1 with RhaTF1, yielded beta-galactosidase assay values similar to background values attained in the other negative interactions (between 45-78 units). However, RhaK72F4 with RhaTF1 yielded values greatly diminished but significantly greater than background values (Figure 3.3). It is noteworthy that the *rhaK72* allele was also capable of conferring a very slow growth phenotype in strains carrying *rhaK* mutations, and can restore marginal rhamnose transport. The *rhaK73* allele confers no growth, and no detectable transport in a *rhaK* mutant background (Rivers and Oresnik 2013).

3.3.3. Identification of putative motifs in the C- and N- terminal fragments of RhaK

Initial characterization of RhaK demonstrated that it had rhamnose kinase activity and that a functional protein was necessary for the transport of rhamnose (Richardson and Oresnik 2007). The initial bioinformatics analysis that was carried out did not identify any motifs that may play a role in interacting with the rhamnose ABC transporter (Rivers and Oresnik 2013). Phylogenetic analysis did however identify other putative orthologs, suggesting this was not a unique protein (Richardson and Oresnik 2007). Since the number of sequenced bacterial genomes has increased since the initial characterization, it was reasoned that a re-analysis coupled with the functional data we had generated might be useful in identifying regions that potentially participate in an interaction with RhaT.

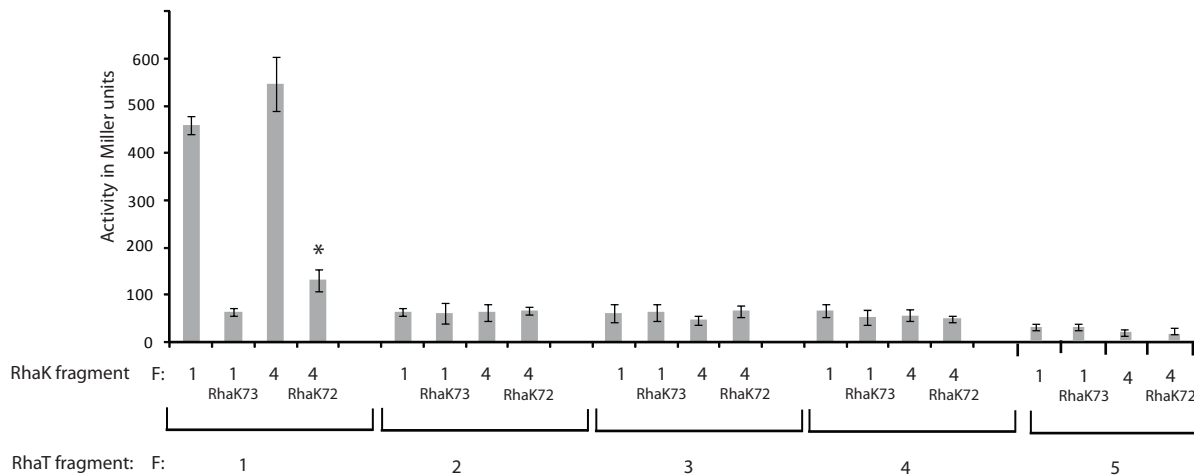


Figure 3.3. Beta-galactosidase activity of BTH101 cells co-transformed with the RhaK and the RhaT fragment indicated. RhaKF1, RhaKF4, and RhaTF1-F5 fragments are as defined in Figure 3.1. RhaK72 or RhaK73 listed below the fragment number indicates the insertion that defines RhaK72 or RhaK73 is present in the corresponding fragment. Growth conditions and assays were carried out as indicated in the material and methods section, with activity reported in miller units. Data represents three biological replicates each assayed in triplicate, with the error bars representing the standard deviation of all replicates. The * indicates this value is found to be statistically different from all other negative values using a Student's t-test ($p = 0.05$).

To define a RhaK dataset the amino acid sequence of RhaK was used as a BLAST query sequence against the IMG database of complete, draft, and permanent draft genomes. To refine the data set, it was assumed that proteins with similar biochemical properties would be found with similar operon structures. Carbohydrate kinases that were found in operons that were not syntenous, or where sequence coverage did not cover the entire operon, were omitted. One hundred and twenty-nine BLAST hits with an e value lower than e^{-150} were retained. Additionally, hits that came from “environmental samples” that were not identified to the species level were also discarded. Finally each hit was manually curated to eliminate multiple identical representatives from the same species. If within a species the RhaK orthologs were identical at the amino acid level only one representative was included in the analysis. The final data set consisted of 52 RhaK-like sequences.

Phylogenetic analysis of the sequences was carried out using the *E. coli* rhamnulose kinase (RhuK) as an out group. The data support the hypothesis that the RhaK-like sequences separated into 2 and possibly 3 clades (Figure 3.4). Whereas the clade that contained the *Mesorhizobium* species was not well supported, a clade that contained the *Sinorhizobium* (*Ensifer*) species and another that predominantly contained the *Rhizobium*, *Agrobacterium*, *Ochrabactrum*, *Brucella* and a few others were strongly supported (Figure 3.4).

RhaK from Rlt100 clustered with other *R. leguminosarum*, *R. eltli*, *R. tropici* as well as *A. rhizogenes* and *A. radiobacter*. *A. tumefaciens* and *R. lupini* were separated from the *R. leguminosarum*; however all of these were within a single clade (Figure 3.4). It is of note that whereas the original published analysis had grouped the *S. meliloti* and *A. tumefaciens*

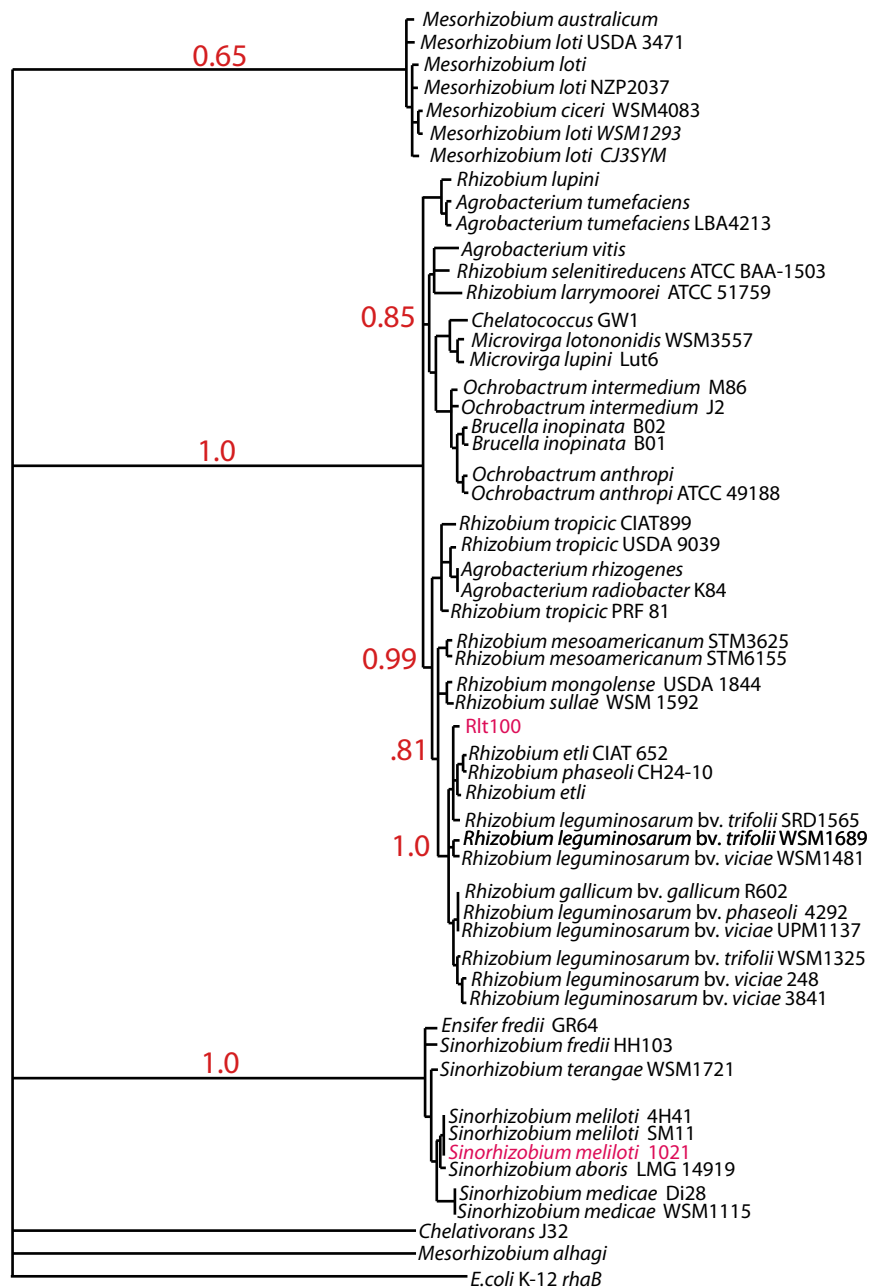


Figure 3.4. A phylogenetic tree of RhaK-like proteins from α -proteobacteria. Phylogenetic analysis was carried out using the Phylogeny.fr server as indicated in the materials and methods section. The support for the major nodes are indicated in red, whereas nodes with less than 50% support have been collapsed. *Rlt100* and *S. meliloti* are highlighted in pink solely to bring attention to their location.

strains together (Richardson and Oresnik 2007), this analysis clearly shows that there are differences among all of these strains.

Since there was divergence between the *S. meliloti* and *R. leguminosarum* RhaK proteins, we wished to determine if the *S. meliloti* RhaK could functionally replace the *R. leguminosarum* RhaK; reasoning that if it could, an alignment of the C-terminal and N-terminal RhaK fragments, coupled with previous functional genetic characterization might give some insight into regions that interact with RhaT. To carry this out the *S. meliloti rhaK* was isolated by using the ORFeome platform (Schroeder, House et al. 2005) . The *rhaK* open reading frame was recombined into the destination vector pCO37, yielding pDR35. Introduction of pDR35 into Rlt144 (*rhaK50::Tn5*) resulted in functional complementation, suggesting that *S. meliloti* RhaK can carry out the same sugar kinase function, and the same role in rhamnose transport as the Rlt100 RhaK .

3.3.4. Identification of a putative interacting domain

Since the *S. meliloti* Rm1021 RhaK homolog was able to complement Rlt100 carrying *rhaK* mutations, it was assumed that that both of these proteins must be able to interact with the *R. leguminosarum* rhamnose transporter component RhaT. In an effort to identify putative interacting motifs an analysis that included only the 23 RhaK protein sequences from groupings that contained *S. meliloti* and *R. leguminosarum* was carried out. The analysis consisted of first aligning the proteins on the basis of local regions of identity using Praline. This alignment was then supplemented with a secondary structure analysis using the PHYRE² server, information from previously characterized *rhaK* variants (Rivers and Oresnik 2013), as well as functional residue predictions based on glycerol and xylulose

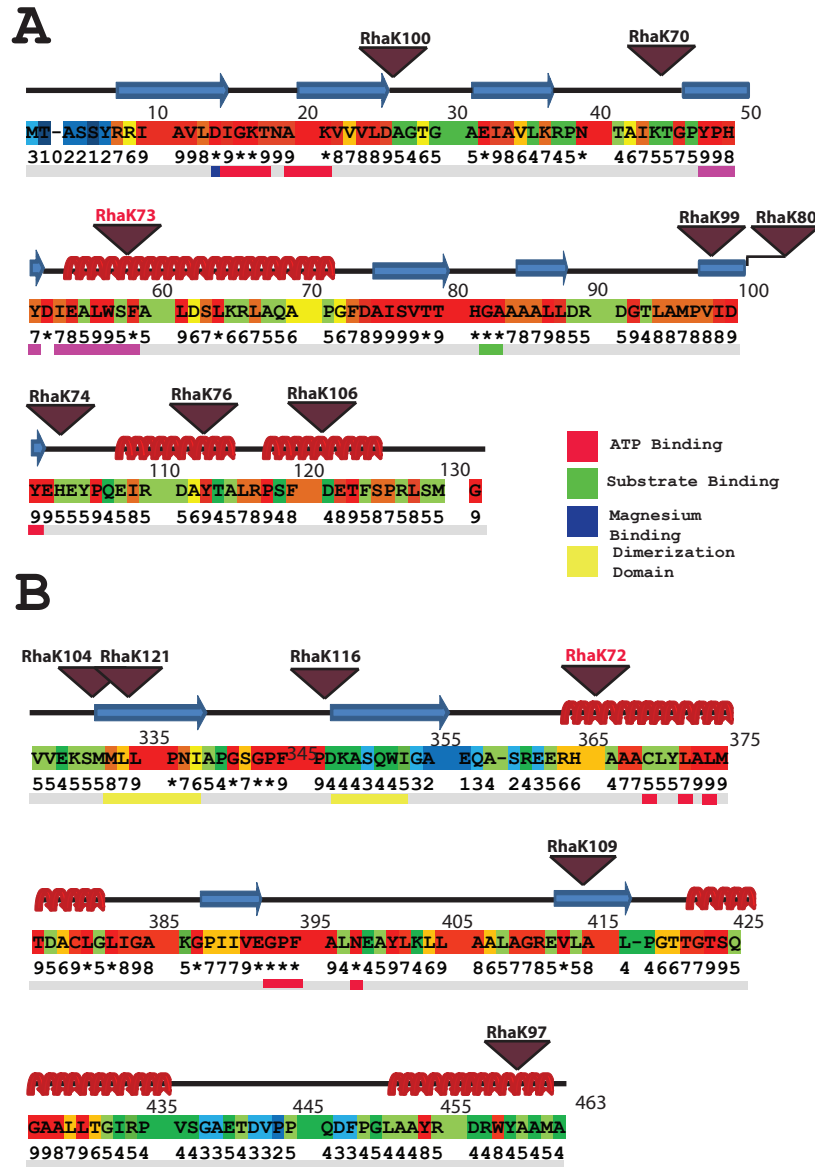


Figure 3.5. The amino acid conservation of 23 RhaK like proteins. The sequences analyzed were chosen as they were found in the phylogenetic groupings that contained the *S. meliloti* and *R. leguminosarum* RhaK proteins. The regions corresponding to RhaKF1 (panel A) and RhaKF4 (panel B) are displayed, with only the amino acid sequence from *R. leguminosarum* shown. A conservation score was calculated using PRALINE, with the scoring ranging from 0 (least conserved) to 10 (invariant, denoted by an *). The conservation score at each site is listed underneath each amino acid. Each amino acid is highlighted by a colour assigned by PRALINE, which corresponds to the conservation score. Colours range from dark blue (least conserved) to bright red (Most conserved). The secondary structure of these regions, as predicted by the Phyre² server is indicated above the amino acid sequence. Regions predicted to be β -sheets are indicated by blue arrows, while regions predicted to form a helix are indicated by red cartoon helices. The sites of the previously generated insertional mutations are indicated by triangles, their allele numbers are listed above. RhaK72 and RhaK73 are highlighted in pink. The colour bar below the conservation number depicts the predicted function of a region. This is as determined by a comparison to the homologous regions of the *E. coli* *GlpK* and the *E. coli* xylulose kinase. Regions of this bar that are grey indicate no predicted function. The predicted functions of the coloured regions are as indicated in the legend. One region is highly conserved with no known function. It is highlighted in purple to identify it as a putative binding domain.

kinase (Anderson, DeLaBarre et al. 2007, Di Luccio, Petschacher et al. 2007). An overview of the C- and N- terminal fragments that were positive in the two-hybrid screens was generated (Figure 3.5). The results show that many regions of high conservation were identified (Figure 3.5). These include regions identified as being involved in kinase dimerization, ATP binding, substrate binding, and magnesium binding. Two areas of conservation were shown to be present in the regions surrounding the *rhaK72* and *rhaK73* insertions sites. Although the region surrounding the *rhaK72* insertion site within the C-terminal fragment showed conservation, this region may also be associated with ATP binding (Figure 3.5b). However in the N-terminal fragment, proximal to the insertion site that defined *rhaK73*, a 12 amino acid region was identified with no predicted function (Figure 3.5a). This region was predicted to be a surface exposed alpha helix that was disrupted in the RhaK73 variant (Rivers and Oresnik 2013).

3.3.5 Characterization of the N-terminal motif

To visually represent the region that makes up the putative motif, a sequence logo of the aligned regions was generated (Figure 3.6a). The motif consists of 12 amino acids, 7 of which are invariant. If this sequence is used as a query for a BLASTP search of the NCBI database; the top 100 hits that were returned are sequences of carbohydrate kinases from the α -proteobacteria. The highest scoring hits are either predicted to be involved in rhamnose catabolism or are in operons that show synteny to the *R. leguminosarum* rhamnose catabolic operon. Within the Rhizobia this region does not appear to be conserved in other kinases outside those predicted to be involved in rhamnose catabolism. However, this region does not seem to be conserved in all kinases involved in rhamnose catabolism. For example, the

E. coli RhuK does not share conservation in this region. Taken together, the data support the hypothesis that this region may constitute a motif involved in a physical interaction with RhaT.

To systematically determine the relative importance of each residue in the proposed binding motif an alanine scanning mutagenesis of the region was carried out. Each of the amino acids was converted to an alanine except for amino acid 74, which was an alanine. To generate a control model unable to interact, a variant in which each of the conserved residues was converted to an alanine was also constructed. These variants were cloned into the appropriate vector and assayed for an interaction with RhaTF1.

The results show all of the substitutions made to the predicted binding domain appeared to significantly affect the ability of the two fragments to interact (Figure 3.6). The RhaKF1 variant containing multiple substitutions, RhaKF1* (containing a substitution of all of the conserved residues) resulted in beta-galactosidase activity reduced to background levels. This variant however contained multiple substitutions that by themselves nearly abolished beta galactosidase activity. The substitution that affected this interaction the least was the E73A. This substitution, however, still resulted in an approximate 50 % reduction in beta-galactosidase activity (Figure 3.6). The region in the middle of this sequence appears more tolerant to alanine substitution, whereas the four N-terminal most substitutions (Y67A, P68A, H69A, Y70A) and the 2 C-terminal most substitutions (S77A, F78A) resulted in the most severe reduction in beta galactosidase activity. The four N-terminal most substitutions are predicted to make up a beta sheet that leads into a surface exposed helix. Part of this beta sheet itself also appears to be surface exposed (Figure 3.6)(Rivers and Oresnik 2013). The

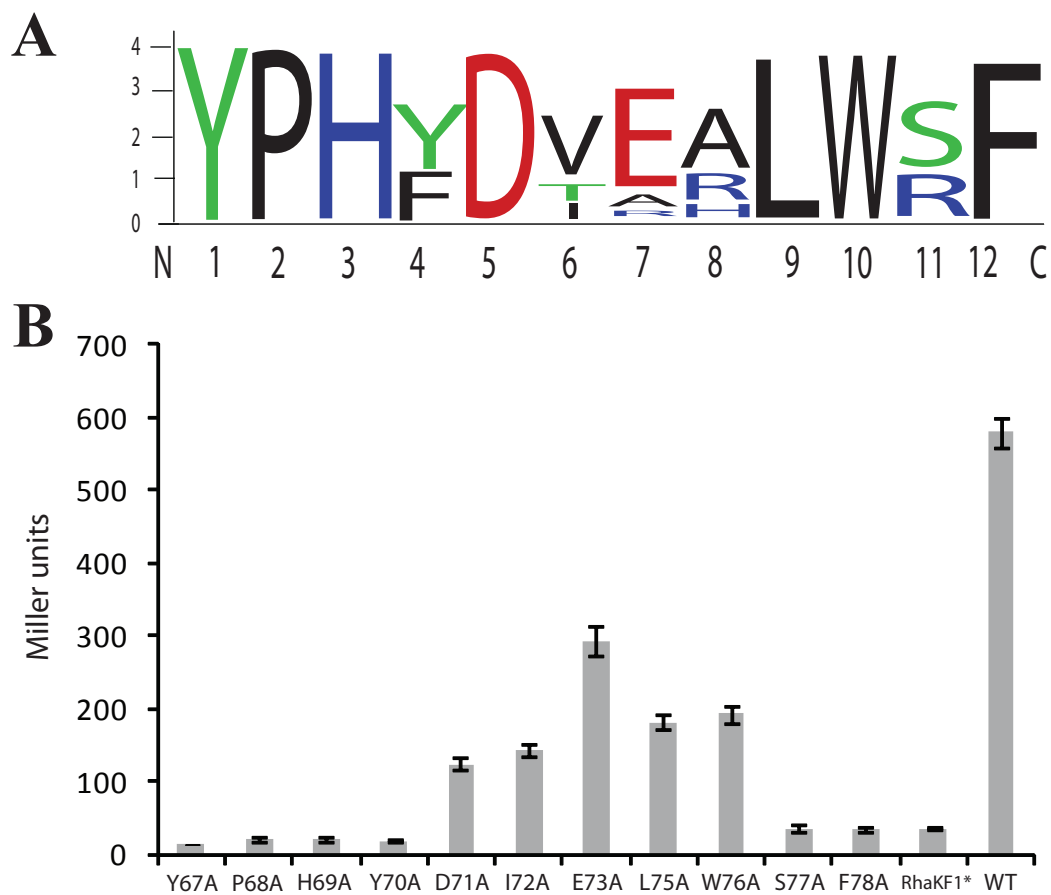


Figure 3.6. (A) The consensus sequence of the putative binding motif as generated using WebLogo. (B) Beta-galactosidase activity of BTH101 cells co-transformed with a RhaKF1 variant fragment, and RhaTF1. Each variant is identified on the x-axis by the amino acid substitution it contains. RhaKF1* contains a substitution of all of the invariant amino acids from panel (A). Growth conditions and assays were carried out as indicated in the material and methods section, with activity reported in miller units. Data represents three biological replicates each assayed in triplicate, with the error bars representing the standard deviation of all replicates.

five least severe substitutions appear to make up the N-terminal most residues of the surface exposed alpha helix. The 2 C-terminal most substitutions (S77A, F78A) localize to a predicted surface exposed alpha helix. Changing either of these 2 residues abolishes beta-galactosidase activity. Of note, the site of the linker scanning insertion sequence in the *rhaK73* allele is localized between these two amino acids (Figure 3.6).

3.4 Discussion

The data show that the N-terminal and C-terminal fragments of RhaK, RhaKF1 and RhaKF4 respectively, can both interact with an N-terminal fragment of the ABC protein RhaT using a bacterial two-hybrid system (Figure 3.2). Linker scanning mutagenesis of RhaK had previously been utilized to uncouple the *in-vivo* ability of RhaK to affect whole cell rhamnose transport from its ability to function as a kinase (Rivers and Oresnik 2013). Since the two alleles (*rhaK72* and *rhaK73*) previously isolated as being able to uncouple these activities were localized within the N- and C-terminal regions defined by RhaKF1 and RhaKF4, these alleles were directly tested and found to disrupt the ability of these fragments of interacting with the RhaT N-terminal fragment (Figure 3.3). Taken together these data strongly suggest that the interaction found in the two-hybrid assays are correlated with physiologically relevant phenotypes and are unlikely to be due to spurious interactions due to the nature of the assay (Bram 2012).

It is also worth noting that when the amino acid inset found in the RhaK72 variant is inserted into F4 fragment, it has a statistically higher beta-galactosidase activity when compared to the RhaK73 variant inserted into in the F1 fragment (132 units versus 65 units respectively, Figure 3.3). This correlates well with previously published data that show that when the *rhaK72* allele is used to complement a chromosomal *rhaK* mutation *in trans* it was capable of conferring a very slow growth with some residual labelled rhamnose accumulation. This was not observed with the *rhaK73* allele (Rivers and Oresnik 2013).

Although this may imply that the strength of the interaction is correlated with the ability to confer transport, more work will be needed to support this hypothesis.

Based on predictive modelling it had been previously proposed that the *rhaK72* and *rhaK73* alleles encoded proteins with surface variations (Rivers and Oresnik 2013).

Coupling this modelling with the experimental evidence generated in this and previous linker-scanning work, allowed us to delineate a conserved region within the N-terminal fragment of RhaK that had no predicted function. Changing any of the residues in this region resulted in at least a 50% loss in its ability to interact with the N-terminal fragment of RhaT in the bacterial 2-hybrid assay. Together the data support the hypothesis that RhaK directly interacts with RhaT. At present our data do not allow us to predict how this interaction might occur, or what role the region proximal to the *rhaK72* insertion in the C-terminal fragment of RhaK may be playing.

Phylogenetic analysis had been previously used in an attempt to identify other proteins that may have the ability to interact with ABC transporters (Richardson and Oresnik 2007). Although the number of sequences was limited, putative orthologues were identified (Richardson and Oresnik 2007). Since a greater number of sequences is now available, repeating the analysis showed that RhaK-like sequences can be divided into separate clades with *Mesorhizobium*, *Sinorhizobium*, and *Rhizobium* RhaK sequences all grouping separately. Although the *rhaK* genes that encode these proteins are found in syntenous operons predicted to be used for rhamnose catabolism, it is of note that the RhaK protein sequences show the greatest degree of divergence. For example, RhaK from *R. leguminosarum* Rlt100 and *S. meliloti* share only 55 % identity. The physiological/biochemical ramifications of these differences are unclear at this time since

rhaK from *S. meliloti* expressed from a plasmid can complement a *rhaK* mutation in *R. leguminosarum* for both transport as well as biochemical activity.

In light of this work it is clear that the nature of the interaction of RhaK with RhaT should be further characterized. Whereas the physical characterization of the kinase may allow a more robust modelling of the interacting domains, it may be possible to use an analogous genetic approach to find sites on RhaT that allow the transport of rhamnose to be RhaK independent. Moreover since RhaK appears to be the most divergent component involved in the transport and catabolism of rhamnose, the characterization of orthologs may provide insight into some of the nuances of this interaction.

Chapter 4

The *Sinorhizobium meliloti* Rm1021 Rhamnose Kinase (RhaK_{Sm}) is required for the function of its associated ABC-type transporter, and only has activity phosphorylating rhamnulose.

4.1 Introduction

The use of symbiotic rhizobia for biological nitrogen fixation is very important for agriculture. In addition to using inoculum strains that are efficient at fixing nitrogen, they must also form nodules on the host plant while competing with indigenous rhizobia (Triplett and Sadowsky 1992). Whereas our understanding of the events involved during nodule development are understood in detail (Oldroyd and Downie 2008, Oldroyd, Marray et al. 2011, Udvardi and Poole 2013), those that affect rhizosphere growth, root colonization or other events that can affect competition for nodule occupancy are not.

Metabolic capacity has been correlated with the ability to compete for nodule occupancy in strains of *R. leguminosarum* (Wiebo, Marek-Kozaczuk et al. 2007, Geddes and Oresnik 2014). This finding correlates well with the predominance of ABC type sugar transporters that are encoded on the megaplasmids found in both *R. leguminosarum* as well as *S. meliloti* (Mauchline, Fowler et al. 2006, Ramachandran, East et al. 2011). A number *R. leguminosarum* or *S. meliloti* mutants unable to catabolize specific carbon compounds have been found to be unable to compete for nodule occupancy. This phenotype has been associated with the inability to catabolize myo-inositol, glycerol, homoserine, erythritol, arabinose, and rhamnase (Gordon, Ryder et al. 1996, Oresnik, Pacarynuk et al. 1998, Fry, Wood et al. 2001, Jiang, Krishnan et al. 2001, Yost, Rath et al. 2006, Poysti, Loewen et al. 2007, Geddes, Pickering et al. 2010, Kohler, Zheng et al. 2010, Vanderlinde, Hynes et al. 2013). Of these, only myo-inositol catabolism appears to affect competition for nodule occupancy in more than one species.

The locus necessary for the utilization of rhamnose in *R. leguminosarum* was found to be plasmid encoded (Baldani, Weaver et al. 1992). The region necessary was subsequently isolated and it was shown that a transcript at this locus was inducible by seed exudate and that mutants that were specifically unable to utilize rhamnose were uncompetitive for nodule occupancy (Oresnik, Pacarynuk et al. 1998). Subsequent genetic characterization showed that the locus consisted of 2 transcripts; one encoding an ABC-type transporter (RhaRSTPQ) as well as a mutarotase (RhaU) (Richardson, Carpena et al. 2008), and sugar kinase (RhaK) (Richardson, Hynes et al. 2004). On the basis of physiological, biochemical and genetic evidence the rhamnose catabolic pathway in *R. leguminosarum* was hypothesized to be initiated by RhaK (Richardson 2004). Further characterization of *rhaK* revealed that the absence of RhaK also affected the transport of rhamnose into the cell (Richardson and Oresnik 2007). The absence of kinase and transport activity did not affect either the transcription or translation of *rhaSTPQ*, suggesting that RhaK, in addition to having activity as a kinase, also played a role in affecting the transport of rhamnose (Richardson and Oresnik 2007). A linker scanning mutagenesis of *rhaK* from *R. leguminosarum* and the subsequent characterization of these variants clearly showed that the ability to affect transport could be genetically separated from the ability of RhaK to act as a sugar kinase. (Rivers and Oresnik 2013). RhaK from *S. meliloti* 1021 has been shown to complement a *rhaK* mutation in *R. leguminosarum* Rlt100, showing they must carry out the same primary biological function. Implicitly, this includes an involvement in the transport of the sugar. However, the respective RhaK proteins are the most divergent proteins in their respective operons, and recent phylogenetic analysis (Figure 3.4) shows greater divergence than the initial analysis revealed (Richardson and Oresnik 2007).

This work was initiated to address two hypotheses; 1) that the ability to catabolize rhamnose may be a determinant in competition for nodule occupancy and thus also plays a role in allowing *S. meliloti* to be competitive for nodule occupancy on alfalfa, and 2) to determine whether *rhaK* from *S. meliloti* has the same activity and role in transport as the RhaK from *R. leguminosarum*.

Here we report the isolation of *S. meliloti* mutants unable to utilize rhamnose as a sole carbon source, their subsequent characterization, and a direct test of the hypothesis that the inability to utilize this sugar affects the ability to compete for nodule occupancy.

4.2 Materials and Methods

4.2.1 Bacterial strains, plasmids, and media

The bacterial strains and plasmids used and generated in the work are listed in Table 4.1. *S. meliloti* strains were routinely grown at 30°C using TY or LB as a complex medium (Beringer, Beynon et al. 1978, Sambrook and Russell 2001). *S. meliloti* strains were grown on either LB or TY as a complex medium. Both *R. leguminosarum* and *S. meliloti* strains were grown on VMM as a defined medium (Vincent 1970), that was modified as previously described (Oresnik, Pacarynuk et al. 1998). When required antibiotics were used at the following concentrations: tetracycline (Tc) either 5 or 10 $\mu\text{g ml}^{-1}$; neomycin (Nm), 200 $\mu\text{g ml}^{-1}$; kanamycin (Kan), 50 $\mu\text{g ml}^{-1}$; streptomycin (Sm), 200 $\mu\text{g ml}^{-1}$; rifampicin (Rf), 100 $\mu\text{g ml}^{-1}$; chloramphenicol (Cm), 20 $\mu\text{g ml}^{-1}$; gentamycin (Gm) 20 or 50 $\mu\text{g ml}^{-1}$. Growth was routinely monitored spectrophotometrically at 600 nm.

Table 4.1. Strains and Plasmids

Stain or Plasmid	Genotype or phenotype	Reference or source
Strains		
<i>S. meliloti</i>		
Rm1021	SU47 <i>str-21</i> ; Sm ^r	(Meade et al., 1982)
SRmA137	Rm1021 <i>rhaR</i> ::Tn5, (Nm ^r)	(This work)
SRmA211	Rm1021 <i>rhaK</i> :: Tn5, (Nm ^r)	(This work)
SRmA191	Rm1021 <i>rhaI</i> :: Tn5, (Nm ^r)	(This work)
SRmA145	Rm1021 <i>rhaT</i> :: Tn5, (Nm ^r)	(This work)
SRmA146	Rm1021 <i>rhaD</i> :: Tn5, (Nm ^r)	(This work)
SRmA943	Rm1021 Δ <i>rhaP</i>	(This work)
<i>R. leguminosarum</i>		
Rlt100	W14-2 Sm ^r , wild-type	(Oresnik 1998)
Rlt144	Rlt100 <i>rhaK</i>	(Richardson 2004)
Rlt106	Rlt100 <i>rhaT2</i> ::Tn5-B20, (Nm ^r)	(Oresnik et al., 1998)
Rlt105	Rlt100 <i>rhaD1</i> ::Tn5-B20, (Nm ^r)	(Oresnik et al., 1998)
Rlt130	Rlt100 <i>rhaI39</i> ::Tn5-B20, (Nm ^r)	(This Work)
Rlt117	Rlt100 <i>rhaR25</i> ::Tn5-B20, (Nm ^r)	(Richardson 2004)
Rlt128	Rlt100 <i>rhaP36</i> ::Tn5-B20, (Nm ^r)	(Richardson 2004)
Rlt151	Rlt100 <i>rhaQ38</i> ::Tn5-B20, (Nm ^r)	(This work)
<i>E. coli</i>		
DH5 α	<i>endA hsdR17 supE44 thi-1 recA1 gyrA96 relA1 (argF-lacZYA) U169 80dlacZ M15</i>	(Hanahan., 1983)
MT616	MT607 (pRK600)	(Finan et al., 1986)
Plasmids		
pRK7813	Broad host range vector, Tc ^r	(Jones, 1987)
pCO37	pRK7813 containing <i>attB</i> sites; Gateway-compatible destination vector	(Jacob et.al, 2008)
pRK600	pRK2013 <i>npt</i> ::Tn9, Cm ^r	(Finan et al., 1986)
pW3C1	pRK7813 cosmid containing entire rhamnose locus	(Oresnik et al., 1998)

pW3R1	pWRC1, <i>rhaD1</i> ::Tn5-B20	(Oresnik et al., 1998)
pW3R2	pWRC1, <i>rhaT2</i> ::Tn5-B20	(Oresnik et al., 1998)
pMR84	pWRC1, <i>rhaP36</i> ::Tn5-B20	(This Work)
pMR110	<i>R.leguminosarum rhaK</i> ⁺ in pRK7813	(Richardson & Oresnik, 2007)
pMR53	<i>R.leguminosarum rhaR</i> ⁺ in pRK7813	(Richardson & Oresnik, 2007)
pDR32	<i>S.meliloti rhaI</i> ⁺ in pCO37	(This Work)
pDR35	<i>S.meliloti rhaK</i> ⁺ in pCO37	(This Work)
pDR190	<i>S.meliloti rhaR</i> ⁺ in pCO37	(This Work)

4.2.2 Genetic techniques

Mutagenesis of Rm1021 with Tn5 was carried out using pRK602 as previously described (Poysti and Oresnik 2007). Putative mutants were routinely single colony purified three times and retested for phenotype. Strains that appeared to carry the desired phenotype were routinely transduced into Rm1021 to ensure that the Tn5 and the marker were 100% linked by transduction (Finan, Hartweg et al. 1984). Conjugations were carried out essentially as previously described (Finan, Oresnik et al. 1988).

Plasmids pDR32, pDR35, and pDR190 were constructed by using the GatewayTM compatible vector pCO37 (Jacob, Adhamn et al. 2008). Briefly, *S. meliloti* ORFeome clones (Schroeder, House et al. 2005), were recombined into pCO37 as previously described (House, Mortimer et al. 2004, Geddes, Pickering et al. 2010). The ORFeome clones for *rhaI*, *rhaK*, and *rhaR* were used to construct pDR32, pDR35, and pDR190 respectively. The resultant constructs were completely sequenced.

4.2.3 DNA manipulations

Standard techniques were used for gel electrophoresis, restrictions, ligations and PCR reactions (Sambrook and Russell 2001). Arbitrary PCR reactions to determine the location of Tn5 inserts in the genome were performed as previously described (Poysti, Loewen et al. 2007). Nucleotide sequencing was accomplished by cycle sequencing using a Big Dye, version 3.1 kit as recommended by the manufacturer and resolved using an ABI3130 sequencer.

4.2.4 Rhamnose transport assay

Transport assays were carried out as previously described (Rivers and Oresnik 2013). Radioactive [³H] rhamnose (5 Ci/mmol) was purchased from American Radiolabeled Chemicals Ltd. (St. Louis, Mo.). Transport assays were initiated by the addition of tritiated rhamnose to a final concentration of 2 μM, and aliquots of 0.5 ml were withdrawn at appropriate time points and rapidly filtered through a Millipore 0.45μm Hv filter on a Millipore sampling manifold. Filtered cells were immediately washed with 5 ml of defined salts medium, and the radioactivity that remained on the filter was determined using a liquid scintillation spectrophotometer (Beckman LS6500). Typically transport rates were linear over the first minute of the assay.

4.2.5 Enzyme assays

R. leguminosarum and *S. meliloti* cells were grown and cell-free lysates were prepared as previously described (Oresnik and Layzell 1994). L-rhamnose isomerase activities were determined by measuring to formation of rhamnulose (ketose formation) utilizing the cysteine-carbazole method as previously described (Dische and Borenfreund 1951, Takagi and Sawada 1964). Sugar kinase assays were carried out as described (Anderson and Sapico 1975), and modified to use rhamnose as a substrate (Rivers and Oresnik 2013).

4.2.6 β-Galactosidase assays

S. meliloti cultures were containing transcriptional fusions were first grown overnight in either TY or LB broth and subsequently sub-cultured in defined broth medium containing

the desired carbon sources. Cultures used for assays were typically at an OD₆₀₀ of approximately 0.5. Assays were carried out essentially as described (Miller 1972), and as previously modified (Richardson, Hynes et al. 2004).

4.2.7 Plant assays

Plant symbiotic assays were carried out using alfalfa (*Medicago sativa* cv. Rangelander) essentially as previously described (Oresnik, Charles et al. 1994). Competition for nodule occupancy experiments were carried out essentially as previously described. Essentially *S. meliloti* cultures were grown overnight in LB and diluting the cultures 1/100 in sterile distilled water and inoculating each Leonard Jar assembly with 10 ml (approximately 10⁶ cfu/seedling). The ratio of the strains in the inoculum was determined by plating an appropriate dilution of the inoculum and screening to determine the proportion of each strain in the inoculum. This was compared to the proportion of nodules occupied by each strain as previously described (Geddes and Oresnik 2012). Statistical significance was determined by using a paired Student's t test.

4.3 Results

4.3.1 Identification of a rhamnose catabolic operon in *S. meliloti*

To identify mutations that could affect the ability of *S. meliloti* to utilize rhamnose as a sole carbon source approximately 10,000 Tn5 generated mutants, from approximately 10 independent mutagenesis experiments were generated. Mutants were screened for their inability to grow on rhamnose as a sole carbon source while retaining the ability to grow on

glucose as a sole carbon source. Mutants with this phenotype were single colony purified three times, retested for phenotype, and checked by transduction to show that the transposon and the phenotype were 100% transducible (typically 50-100 colonies screened). Ten mutants were isolated from these screens (Table 4.1). In addition to these, a strain carrying a mutation in *tpiA* was also isolated on the basis of a slow growth phenotype and has been previously reported (Poysti and Oresnik 2007). The site of the Tn5 insertion in each of the mutants was determined by sequencing the product of an arbitrary PCR reaction that used the genomic DNA from each of the mutants as template. Each of the inserts were localized to the genome of Rm1021 by using the generated sequence as a BLASTn query against the Rm1021 genome. The 10 mutations were mapped to the locus delineated by *SMc02321* and *SMc03003* (Figure 4.1). We note that the genes at this locus are currently annotated as *rha* based on identity and synteny to that of *R. leguminosarum* (Richardson, Hynes et al. 2004). Other studies have also identified genes in this region as being involved with the ability to utilize rhamnose as a sole carbon source (Mauchline, Fowler et al. 2006, Poysti and Oresnik 2007, Jacob, Adhamn et al. 2008, Rodionova, Li et al. 2013). The annotation of the genes at this locus are in agreement with what has been previously determined for *R. leguminosarum* (Richardson, Hynes et al. 2004).

4.3.2 Complementation of *S. meliloti* rhamnose mutants with *R. leguminosarum* rhamnose genes.

To aid in the characterization of this locus, complementation cloning experiments were carried out. Attempts to isolate a complementing cosmid from two independent cosmid banks were repeatedly unsuccessful. Subsequent screening of these pooled banks using PCR

primers that were designed to a region in *rhaP* were also unable to generate a PCR product. This suggesting that the rhamnose region of *S. meliloti* is either poorly represented, or completely absent from these cosmid banks. In light of this, heterologous complementation experiments were designed to corroborate that the *S. meliloti* locus had the same operon structure as that of *R. leguminosarum*.

Cosmids containing either the *R. leguminosarum* wild-type *rha* region (pW3A1), or carrying a transposon in either of the two transcripts (*rhaD* or *rhaT*; pW3AR1 and pW3AR2 respectively), were mobilized into *S. meliloti* mutants as well as the appropriate *R. leguminosarum* mutants, Rlt105 (*rhaDI*) and Rlt106 (*rhaT2*). Transconjugants were tested for the ability to use rhamnose as a sole carbon source. As previously shown pW3A1 complemented both *R. leguminosarum* mutants Rlt105 and Rlt106 (Oresnik, Pacarynuk et al. 1998). As expected, strain Rlt105 was complemented by pW3AR2 but not pW3AR1 while Rlt106 was complemented by pW3AR1, but not by pW3AR2. The *S. meliloti* mutants were complemented by the cosmid pW3C1 that contains the entire wild-type locus but not by either pW3AR1 or pW3AR2 .

4.3.3 *S. meliloti rhaDI* mutants are not sensitive on rhamnose/glycerol media

In *R. leguminosarum* it was found that strains carrying either a *rhaDI* or a *rhaI* mutation could not grow on defined medium containing rhamnose and glycerol (Richardson, Hynes et al. 2004). This phenotype was shown to be dependent on *rhaK*, and is consistent with the build-up of a phosphorylated rhamnose compound (Richardson, Hynes et al. 2004, Richardson and Oresnik 2007).

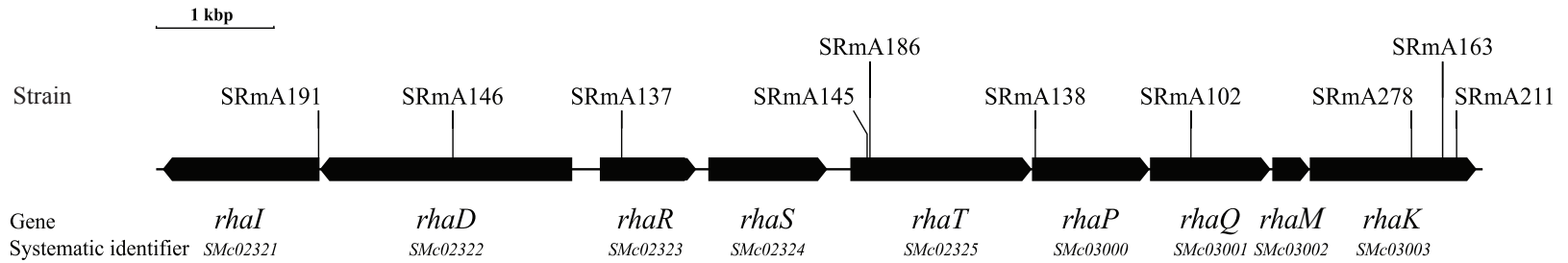
To determine if *S. meliloti* mutants unable to utilize rhamnose behaved similarly, mutants were tested for their ability to grow on defined media containing rhamnose and/or glycerol (Table 4.2). It was found that whereas the *R. leguminosarum* strains carrying a *rhaDI* mutation could not grow on defined medium containing rhamnose and glycerol, the *S. meliloti* mutants could grow equally well on either medium (Table 4.2). However, the over-expression of *rhaI* *in trans* in a *rhaDI* mutant background, resulted in less robust growth on minimal media containing rhamnose and glycerol. This is the same phenotype observed in *E. coli* strains that catabolize rhamnose via the canonical phosphorylated rhamnose pathway (Power 1967).

4.3.4 *S. meliloti* *rhaK* (*rhaK_{Sm}*) and *rhaI* (*rhaI_{Sm}*) can complement *R. leguminosarum* mutants

S. meliloti strains carrying *rhaDI* mutations can grow on defined medium containing glycerol and rhamnose. *R. leguminosarum* cosmids cannot complement *S. meliloti* growth on rhamnose as a sole carbon source when the cosmids carry an insert in any of the rhamnose utilization genes. This suggested that either there is a fundamental difference in how these two organisms catabolised the sugar, or how the transcripts at the locus are regulated. To resolve if the difference could be due to a difference in how the sugar was catabolized, we wished to determine if the terminal genes in both operons (*rhaI* and *rhaK*) were capable of heterologously complementing corresponding mutations in each organism.

To carry this out, each of the genes was cloned into a broadhost range expression vector. *rhaI* and *rhaK* from *S. meliloti* (referred to as *rhaI_{Sm}* *rhaK_{Sm}*) were recombined from

A.



B.

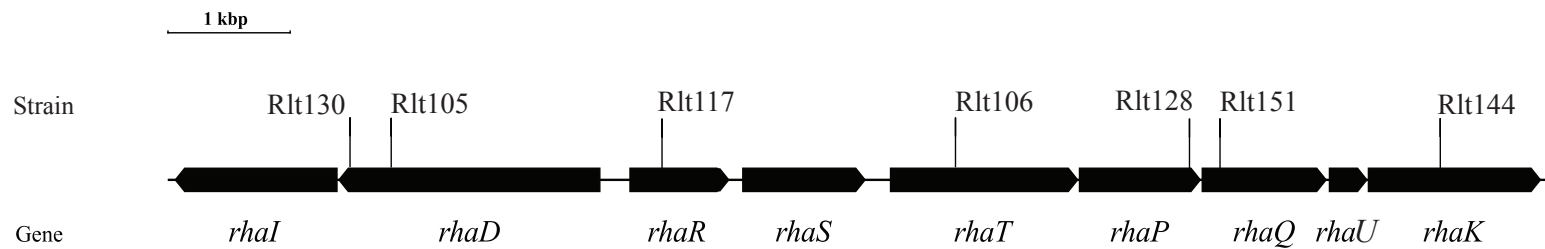


Figure 4.1. The rhamnose catabolic regions in *S. meliloti* strain Rm1021 (A), and *R. leguminosarum* bv. *trifolii* strain Rlt100 (B). Solid arrows represent the genes and the direction of transcription of at each locus. Vertical lines represent sites of Tn5 or Tn5-B20 insertion. Strain names corresponding to specific alleles are given above the genes, gene names are given below the genes. In the case of *S. meliloti*, the systematic identifier numbers are also shown.

Table 4.2. *S. meliloti* displays different conditional growth phenotypes than *R. leguminosarum*

Strain	Relevant characteristics Chromosomal (plasmid)	Glyc	Rham	Rham/Glyc
Rlt100	wild-type	+	+	+
Rlt144	<i>rhaK</i>	+	-	+
Rlt105	<i>rhaDI</i>	+	-	-
Rm1021	wild-type	+	+	+
SRmA211	<i>rhaK</i>	+	-	+
SRmA146	<i>rhaDI</i>	+	-	+
SRmA191	<i>rhaI</i>	+	-	+
SRmA191 (pDR32)	<i>rhaI</i> (<i>rhaI</i> ⁺)	+	+	+
SRmA146 (pDR32)	<i>rhaDI</i> (<i>rhaI</i> ⁺)	+	-	+/-

+ indicates ability to grow on VMM minimal media plus indicated carbon source
 - indicates inability to grow on VMM minimal media plus indicated carbon source
 +/- indicates weak growth on VMM minimal media plus indicated carbon source

the *S. meliloti* ORFeome into the pCO37; yielding pDR32 and pDR35 respectively. We had previously constructed pMR110 which contained *rhaK* from *R. leguminosarum* (*rhaK_{RI}*) (Richardson and Oresnik 2007). All of these constructs were capable of heterologous complementation (Table 4.3). Although we were able to isolate multiple independent *rhaI* clones from *R. leguminosarum* (*rhaI_{RI}*) and verify these by nucleotide sequencing, the constructs were unable to complement *R. leguminosarum* *rhaI* mutations. The reason for their inability to complement was not pursued any further.

4.3.5 *RhaK_{Sm}* does not have measurable rhamnose kinase activity

In *E. coli*, the rhamnose catabolic pathway consists of an isomerase that converts rhamnose to a keto-sugar that is subsequently phosphorylated and finally undergoes an aldolase reaction to yield two three carbon sugars; lactaldehyde, and di-hydroxy-acetone phosphate (Sawada and Takagi 1964, Takagi and Sawada 1964, Takagi and Sawada 1964, Power 1967). In *R. leguminosarum*, it has been shown that RhaK has rhamnose kinase activity (Richardson and Oresnik 2007, Richardson, Carpena et al. 2008, Rivers and Oresnik 2013).

Heterologous complementation experiments with the *rhaI* and *rhaK* suggested that these genes were orthologous. However, *S. meliloti* *rhaDI* mutants do not display a conditional growth inability on medium containing rhamnose and glycerol, and phylogenetic analysis shows the kinases are more divergent than initially expected (Figure 3.4). Because of this, and heterologous cosmid complementation anomalies, we wished to determine if RhaK_{Sm} and RhaI_{Sm} had the same biochemical activity as RhaK_{RI} and RhaI_{RI}.

Table 4.3. Plasmid encoded *rhaK* and *rhaI* can heterologously complement *rhaK* and *rhaI* mutations respectively.

Strain	Relevant Characteristics	pDR32 (<i>rhaI</i> _{Sm}) ^a	pDR35 (<i>rhaK</i> _{Sm})	pMR110 (<i>rhaK</i> _{Rl})
Rlt100	<i>R. leguminosarum</i> , wild-type	+	+	+
Rlt144	<i>R. leguminosarum</i> , <i>rhaK</i> ⁻	-	+	+
Rlt130	<i>R. leguminosarum</i> , <i>rhaI</i>	+	-	-
Rm1021	<i>S. meliloiti</i> , wild-type	+	+	+
SRmA211	<i>S. meliloiti</i> , <i>rhaK</i> ⁻	-	+	+
SRmA191	<i>S. meliloiti</i> , <i>rhaI</i>	+	-	-

Growth was scored as indicated; +, growth comparable to wild-type; -, no growth.

^a Lower case Sm or lower case Rl following gene designation refers to the origin of the gene as *S. meliloti* or *R. leguminosarum* respectively.

The results show that cleared extracts from both *S. meliloti* and *R. leguminosarum* had inducible rhamnose isomerase activity (Table 4.4). These activities were absent in both strains carrying *rhaI* mutations, and were restored with the introduction of *rhaI_{Sm}* on a plasmid (Table 4.4). When the extracts were assayed for rhamnose kinase activity, it was found that Rlt100 had an inducible rhamnose kinase activity that was dependent upon the presence of *rhaK_{Rl}*. Surprisingly, we could not detect this activity in Rm1021 (Table 4.4). This activity could be assayed in *S. meliloti* transconjugants that carried *rhaK_{Rl}*, but not by transconjugants that carried *rhaK_{Sm}* (Table 4.4). Moreover, the *R. leguminosarum* strain Rlt144 was complemented by *rhaK_{Sm}* for growth on rhamnose as a sole carbon source. However, when assayed for rhamnose kinase activity, no measurable rhamnose kinase activity was detected.

4.3.6 The rhamnose transporter *RhaSTPQ* is required for growth on rhamnose

To provide direct evidence that the ABC-type transporter defined by *RhaSTPQ* was responsible for rhamnose uptake, transport assays were carried out using uniformly labelled [³H] rhamnose. *S. meliloti* grown on defined medium with glucose had negligible rates of rhamnose uptake whereas rhamnose grown cells showed significant rates of rhamnose uptake (Figure 4.2). The typical transport rates measured in Rm1021 were approximately 3.5 nmol/min/mg, whereas the typical rate measured for *R. leguminosarum* is more than twice this value (10 nmol/min/mg) (Rivers and Oresnik 2013). When a representative transporter mutant (SRmA145 - containing an insertion in *rhaT*) was assayed, it was found to still have transport rates significantly greater than wild-type glucose grown cells. The

Table 4.4: Kinase and isomerase activity of *R. leguminosarum* and *S. meliloti* strains

Strain	Relevant characteristics	Kinase Activity ^a		Isomerase Activity ^b	
		Glc	Rham ^c	Glc	Rham ^c
Rlt100	<i>R. leguminosarum</i> , wild-type	23±3	211±3	48±4	180 ± 6
Rlt144	Rlt100, <i>rhaK50</i>	44± 6	50 ± 27		
Rlt144 (pMR110)	Rlt100, <i>rhaK50</i> (<i>rhaK</i> ⁺ _{Rl})		640 ± 40 ^d		
Rlt144 (pDR35)	Rlt100, <i>rhaK50</i> (<i>rhaK</i> ⁺ _{Sm})		50 ± 15 ^d		
Rlt130	Rlt100, <i>rhaI</i>			ND ^e	43 ± 9
Rlt130 (pDR32)	Rlt100, <i>rhaI</i> (<i>rhaI</i> ⁺ _{Sm})				269 ± 8 ^d
Rm1021	<i>S. meliloti</i> , wild-type	18 ± 9	19 ± 6	65 ± 7	227 ± 32
SRmA211	Rm1021, <i>rhaK</i>	13±3	19 ± 3		
SRmA211 (pMR110)	Rm1021, <i>rhaK</i> (<i>rhaK</i> ⁺ _{Rl})		383 ± 18 ^d		
SRmA211 (pDR35)	Rm1021, <i>rhaK</i> (<i>rhaK</i> ⁺ _{Sm})		33 ± 8 ^d		
SRmA191	Rm1021, <i>rhaI12</i>			ND	12 ± 1
SRmA191(pDR32)	Rm1021, <i>rhaI12</i> (<i>rhaI</i> ⁺ _{Sm})			ND	152 ± 31

^a Indicates data presented in μmoles/min/mg protein

^b Indicates results reported in milliunits per mg total protein

^c Grown on rhamnose, or rhamnose plus glycerol if the strain is unable to grow on rhamnose as a sole carbon source

^d Genes encoded on pRK7813 are constitutively expressed from a *plac* promoter. Previous experiments confirm similar results are obtained measuring activities in inducing and non-inducing conditions. Because of this, activities were only measured in one condition.

^e Not detected

rhaT insertion in SRmA145 was also polar on *rhaK*, and RhaK has been shown to affect rhamnose transport and growth in *R. leguminosarum* (Rivers and Oresnik 2013). Because of this it was not clear that the inability to grow was solely due to the lack of *rhaTPQ*. To address this, a deletion of *rhaP* was constructed yielding SRmA943. SRmA943 did not grow on rhamnose as a sole carbon source, and had transport rates that were not statistically different from SRmA145. We note that if SRmA943 was incubated for an extended period of time on defined agar medium (greater than 10 days), weak growth could be seen.

RhaK in *R. leguminosarum* affects rhamnose transport, and its kinase activity can be genetically separated from its affect on transport (Richardson and Oresnik 2007, Rivers and Oresnik 2013). To determine if RhaK_{Sm} affected transport, SRmA211 was assayed for rhamnose uptake. The results show that the kinase mutant (SRmA211) had rhamnose transport rates that were comparable to strains carrying an insertion in *rhaT* (SRmA145) or containing a deletion in *rhaP* (SRmA943). Introduction of either *rhaK_{RI}* or *rhaK_{Sm}* on a plasmid allowed the strain to grow on rhamnose as a sole carbon source, as well as take up rhamnose at rates comparable to Rm1021 (Figure 4.2). These data show that the products of *rhaK_{Sm}* and *rhaK_{RI}* both encode proteins that can affect the *S. meliloti* ABC transporter RhaSTPQ.

4.3.7 Cosmids carrying the R. leguminosarum rhamnose locus show reduced expression in Rm1021

The *R. leguminosarum* rhamnose locus has been shown to be negatively regulated by RhaR (Richardson, Hynes et al. 2004). Although *rhaRSTPQUK* were shown to be a single transcriptional unit by complementation experiments, transcription of *rhaRS* appears to occur

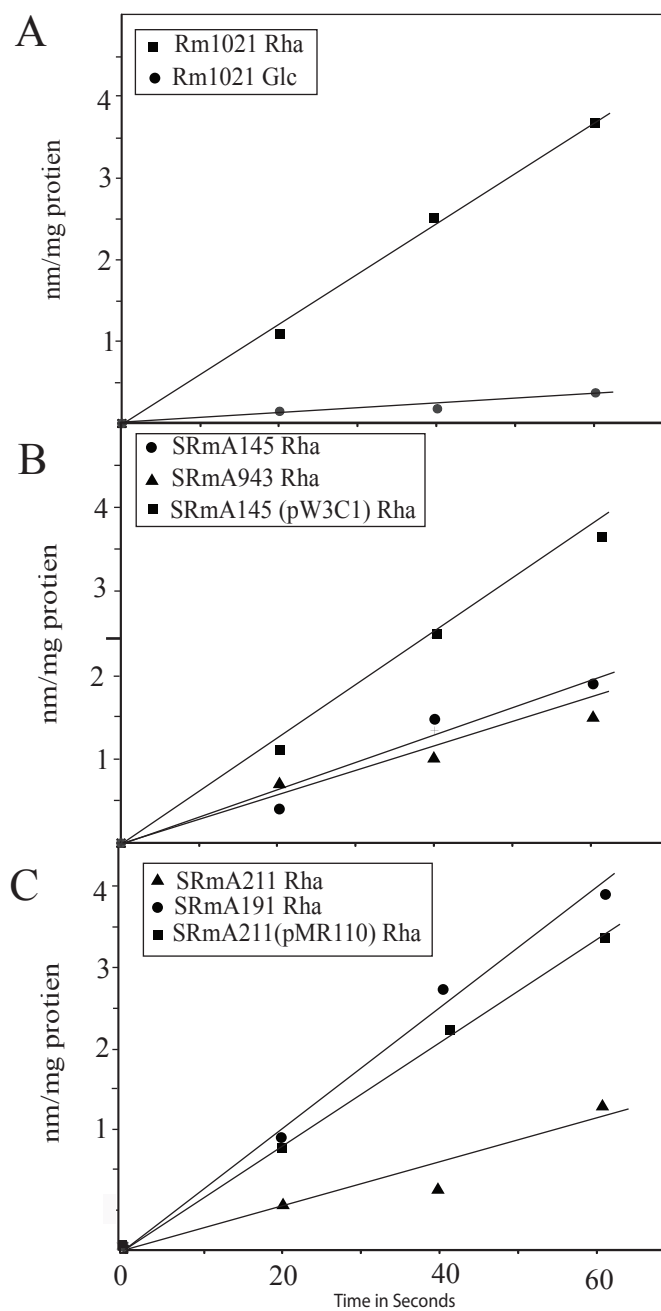


Figure 4.2. Rhamnose transport assays. Strains were grown to mid log phase in defined medium containing either glucose/glycerol or rhamnose/glycerol as indicated. Transport rates were determined by using [³H] rhamnose as described in materials and methods. Symbols for strains are indicated on figure inset. The data shown are from a single experiment that was repeated three times showing consistent results.

under non-inducing conditions. Full operon induction only occurs under inducing (Richardson, Hynes et al. 2004), but was not dependent on the presence of the ABC transporter or *rhaK* (Richardson, Hynes et al. 2004, Richardson and Oresnik 2007).

R. leguminosarum, cosmids carrying inserts in the rhamnase locus cannot complement mutations to any *S. meliloti* rhamnase gene. It seems unlikely this is due to biochemical differences since, *rhaI_{RI}* and *rhaK_{RI}* are capable of complementing the corresponding *S. meliloti* mutants. It is therefore hypothesized that this inability to complement could be due to regulatory differences. In particular, RhaR_{Sm} could act in a dominant fashion to RhaR_{RI}.

To test this, cosmid borne transcriptional fusions were introduced into both Rm1021 and Rlt100. The data shows that the *rhaP36* and the *rhaDI* fusions were induced greater than 6 and 5 fold respectively in a Rlt100 background. These same fusions were either not induced, or very marginally induced in a Rm1021 background (Table 4.5). We note that when the *RhaDI* fusion was introduced into a *S. meliloti rhaR⁻* background this fusion was now capable of being induced by rhamnase. Background expression under non-inducing conditions was also higher in this *rhaR⁻* background (Table 4.5). Introduction of pDR190 (*rhaR_{Sm}* on a broad-host vector) into Rlt100 did not affect the ability of the transconjugants to grow using rhamnase as a sole carbon source. However, consistent with the plasmid fusion data the presence of pDR190 did reduce the expression of chromosomal transcriptional fusion alleles.

Table 4.5. Induction of *R. leguminosarum* cosmid fusions in either *R.leguminosarum* or *S. meliloti* backgrounds.

Strain	Relevant characteristics	Glucose	Rhamnose	Induction ^b
Rlt100	<i>R. leguminosarum</i> , wt	30±8	77±12	-
Rm1021	<i>S. meliloti</i> , wt	7±3	11±4	-
SRmA137	<i>S. meliloti rhaR</i> ⁻	6±2	9±4	-
Rlt100 (pMR84)	Rlt100 (<i>rhaP36</i> cosmid)	54±3	362±20	6.7
Rm1021 (pMR84)	Rm1021 (<i>rhaP36</i> cosmid)	122±20	217±20	1.8
Rlt100 (pW3CR1)	Rlt100 (<i>rhaD1</i> cosmid)	153±10	832±27	5.4
Rm1021 (pW3CR1)	Rm1021 (<i>rhaD1</i> cosmid)	207±18	208±34	1.0
SRmA137 (pW3CR1)	SRmA137 (<i>rhaD1</i> cosmid)	602 ± 31	1934 ± 18	3.2

All trials were grown with glycerol as a secondary carbon source in the media. Data is reported in Miller units.

^b refers to the fold increase in beta-galactosidase activity of a strain grown in rhamnose vs grown in glucose.

4.3.8 *S. meliloti* rhamnase mutants are less competitive for nodule occupancy on alfalfa

R. leguminosarum bv. *trifolii* mutants unable to catabolize rhamnase were found to be severely defective in their ability to compete for nodule occupancy (Oresnik, Pacarynuk et al. 1998). To test the hypothesis that the ability to utilize rhamnase would affect competition for nodule occupancy in *S. meliloti*, a representative mutant unable to catabolize rhamnase was tested against the wild-type. To circumvent the possibility that the build-up of a catabolic intermediate may affect the ability of a mutant to compete with the wild-type, SRmA145 (*rhaT*::Tn5) was chosen because it is unable to transport rhamnase. Preliminary plant growth experiments showed that mutants unable to grow on rhamnase were able to form an effective symbiotic association with alfalfa. Plant dry weights were measured and found to be not significantly different from those inoculated with wild type .

To determine if the inability to utilize rhamnase would affect the ability to compete for nodule occupancy, two ratios of inocula were used. Equal proportions of mutant:wild-type ($45 \pm 7\%$) and a high ratio of mutant:wild-type ($65 \pm 5\%$) were tested. Resulting nodules were sampled after 28-35 days. The isolated bacteria were tested for their ability to grow on rhamnase as a sole carbon source and for their ability to grow in the presence of antibiotics. This was used as a means to determine if they were mutant or wild-type bacteria. In each case, the observed nodule occupancy ratio was lower than the inoculum ratio (Figure 4.3). A test for significance by a paired Student's t-test indicates that the ratio of nodules occupied by SRmA145 is significantly lower ($p < 0.0003$) when the inoculum ratio of SRmA146:Rm1021 was approximately equal. Similar results were found when SRmA145 outnumbered the wild-type two-fold ($p < 0.0007$). Taken together, the ability to use

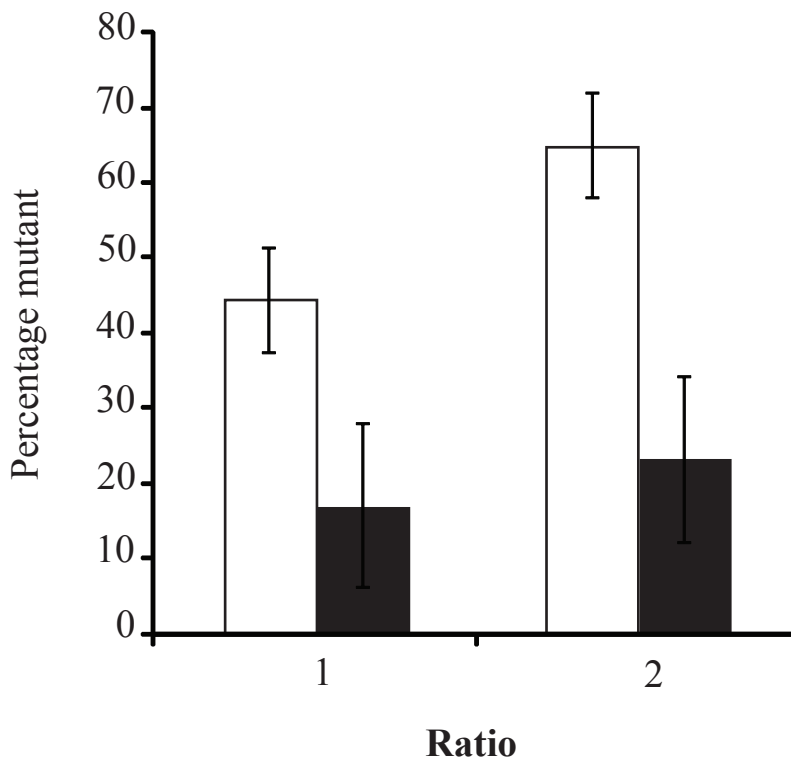


Figure 4.3. Competition for nodule occupancy assays between Rm1021 (wild-type) and the rhamnose catabolic mutant SRmA145 (*rhaT*::Tn5). SRmA145 was inoculated with the wild-type onto alfalfa plants. Open bars represent the percentage of mutant in the inocula; filled bars represent the percentage of nodules occupied by the mutant strain as determined by bacterial isolation from nodules. Ratio 1, n = 6, P < 0.0003; Ratio 2, n = 5, P < 0.0007).

rhamnose as a sole carbon source affects the ability to compete for nodule occupancy in *S. meliloti* in an analogous manner to *R. leguminosarum*.

4.4 Discussion

In this work two putative operons were identified that are necessary for rhamnose catabolism in *S. meliloti*. The operons have the same genetic organization. Each gene product has high similarity to those of *R. leguminosarum* rhamnose catabolic genes products that have been previously described (Richardson, Hynes et al. 2004). It was assumed that the catabolic pathways and the biological significance of rhamnose utilization would be similar to those of *R. leguminosarum* since the region appeared to contain the same genes in the same order with a high similarity at the amino acid level (minimum of 67% similarity). It was initially intended that all future work pertaining to RhaK and its involvement in rhamnose transport would be done in this organism as it is a sequenced, model organism with more available genetic tools. Although the operons and the sequences of the encoded genes are highly similar, there are differences between *R. leguminosarum* and *S. meliloti* rhamnose catabolism.

Using radioactive assays and coupled enzyme assays, RhaK_{RI} has previously been shown to use rhamnose as a substrate (Richardson & Oresnik, 2007) (Rivers and Oresnik 2013). In this work we were able use this coupled enzyme assay to detect this activity. Utilizing this assay however, we were unable to detect rhamnose kinase activity over background for RhaK_{Sm}. This combined with isomerase assay data and heterologous complementation results, suggests a model where rhamnose is first isomerized into rhamnulose in both *R. leguminosarum* and *S. meliloti*. In this model, both strains follow the

isomerisation step with the phosphorylation of rhamnulose. This is then likely followed by an aldolase reaction. This is supported by; 1) RhaI_{RI} and RhaI_{Sm} having the same activities , 2) RhaK_{Sm} not having activity on rhamnose directly, but still being able to complement a RhaK_{RI} mutant for growth on rhamnose and 3) a *tpiA* mutation shows slow grow on rhamnose as a sole carbon source, suggesting dihydroxy acetone phosphate is one of the end products of the rhamnose catabolic genes (Geddes and Oresnik 2014). This suggests that the product of RhaK_{RI} phosphorylating rhamnose directly is not a catabolic intermediate required for growth on rhamnose. This leaves the role of this product unknown. However, this does explain the rhamnose/glycerol growth sensitivity phenotypes found in *R. leguminosarum* (Figure 4.2)(Richardson, Hynes et al. 2004). Although previously published phylogenetic analysis suggests that RhaK_{RI} and RhaK_{Sm} are very closely related (Richardson & Oresnik, 2007), more recent analysis suggests that these fall in phylogenetically distinct groups (Figure 3.4).

R. leguminosarum strains with *rhaK* mutations are unable to transport rhamnose or phosphorylate the sugar (Richardson and Oresnik 2007). Implicit in complementing a *R. leguminosarum* strain carrying a *rhaK* mutation with RhaK_{Sm} is that both the kinase and the transport activities are complemented. This suggests that like *R. leguminosarum*, *S. meliloti* *rhaK* mutants are unable transport rhamnose. This was confirmed by uptake assays. Similar to the equivalent mutation in *R. leguminosarum*, kinase mutants with the kinase delivered *in trans* on a plasmid have transport rates restored to wild type levels. This indicates that this transporter also requires the kinase for transport activity. Transport assays show that *rhaK* and *rhaT* mutants of *S. meliloti* have a significantly lower transport rates for rhamnose. However, a substantial amount of rhamnose still appears to accumulate in these strains. This

is unlike *R. leguminosarum* rhamnase transport and suggests two possibilities. First, *S. meliloti* could possess another transporter capable of transporting rhamnase. Second, the accumulation of rhamnase is due to either diffusion or non-specific uptake through another transporter. A strain carrying a deletion of *rhaP* did not show growth on rhamnase as a sole carbon source after 3 days (the normal time this strain is scored). Though this strain did show some minimal growth after 12 days, this shows the minimal accumulation of rhamnase in the deletion strain is insufficient for normal growth on rhamnase as a sole carbon source. Thus, *rhaSTPQ* encodes a kinase requiring, rhamnase transporter that is required for robust growth on rhamnase.

It has been shown that RhaK_{RI} has activity directly on rhamnase. This activity is likely not necessary for the catabolism of the sugar, as RhaK_{Sm} which lacks this activity is sufficient for complementing *R. leguminosarum rhaK* mutations. It has also been shown that RhaR_{Sm} prevents full induction of the *R. leguminosarum* catabolic genes in the presence of rhamnase. It's known that DeoR type negative regulators typically interact with phosphorylated catabolic intermediates (Mortensen, Dandanell et al. 1989). It is possible that the product of RhaK_{RI} acting directly on rhamnase is involved in this type of regulation. Alternately RhaK_{RI} acting directly on rhamnase may lead to the generation of a product that has no catabolic role. This enzyme function may be an artifact of evolution. Further work on the regulation of the operon will attempt to determine the difference between RhaR_{RI} and RhaR_{Sm} as well as the potential role of this phosphorylated product.

The original basis of this work was to test the hypothesis that rhamnase catabolism plays a role in competition for nodulation in *S. meliloti*. Our data clearly shows that *S. meliloti* rhamnase mutants, like *R. leguminosarum* rhamnase mutants, are compromised in

their ability to compete against the wild type for nodule occupancy (Figure 4.3). Compared to *R. leguminosarum*, the *S. meliloti* wild type does not grow very well on rhamnose as a sole carbon source. However, mutants unable to use rhamnose as a sole carbon source appear to have a similar inability to compete for nodule occupancy against their wild type. It is possible that at some point during the colonization or infection process rhamnose is available and the inability to use this carbon source may affect the ability to directly compete with the wild type. Consistent with this possibility, small amounts of rhamnose have been measured in the root exudates of legumes (Knee, Gong et al. 2001). However, it is also possible this inability to compete for nodule occupancy is not due to the inability to metabolize this sugar, but due to an effect on the regulation of other genes necessary for competition. Characterization of the cause of the competitive phenotype of *S. meliloti* mutants unable to use rhamnose is the subject of ongoing work.

Chapter 5:

General thesis conclusions

5.1 Thesis conclusions and observations

5.1.1 Project history

At the onset of my graduate research, it was known that RhaK was necessary for the transport of rhamnose as carried out by its associated ABC transporter (RhaSTPQ). This in and of itself was unique; however this protein also appeared to catalyse the direct phosphorylation of the sugar rhamnose (Richardson, Hynes et al. 2004). This lead in part to the conclusion rhamnose catabolism in *R. leguminosarum* was likely accomplished using a novel pathway. A pathway where the sugar was first phosphorylated, then isomerized after the phosphorylation step (Richardson, Hynes et al. 2004). The reasons RhaK was necessary for transport, and the catabolic pathway utilizing this novel phosphorylated rhamnose product were unknown. At the onset of my research, the primary hypothesis was that RhaK dependant sugar transport, and phosphorylation of rhamnose could be uncoupled. It was predicted that this role in transport would be due to a novel physical interaction. My work to date has allowed us to gain insight into the role of RhaK in rhamnose transport and to partially characterize its interaction with the transporter. As well, I have proposed a pathway for rhamnose catabolism in *R. leguminosarum*.

5.1.2 The kinase activity and transport role of RhaK can be uncoupled

To address the hypothesis that RhaK dependant phosphorylation of rhamnose and its role in transport could be uncoupled genetically, a bank of in-frame insertional alleles were generated. This approach turned out to be very successful generating two alleles of interest, *rhaK72* and *rhaK73*. Each allele was incapable of complementing RhaK dependant transport

while still maintaining the ability to phosphorylate rhamnose (Rivers and Oresnik 2013). Although this mutagenesis strategy was successful in generating transport negative kinase positive alleles, it was unable to generate alleles with the inverse phenotype (Rivers and Oresnik 2013). This suggests the possibility that though not sufficient for transport, a catabolically active kinase may be required to complement transport. This is supported by the fact that an amino acid substitution in which a P-loop lysine is substituted for an aspartate or methionine (RhaK(K16D), RhaK(K16M)) renders RhaK catabolically inactive and the cell incapable of transporting rhamnose (Richardson and Oresnik 2007). It is possible that we simply did not saturate the protein sufficiently, and a kinase negative, transport positive allele is possible. It is also possible that generating such an allele is not possible due to structural constraints, but the transport mechanism still doesn't require that RhaK be catalytically active.

5.1.3 RhaK physically interacts with RhaT

Initially, it was hypothesised that the RhaK dependant transport phenotype was due to a direct physical interaction with a transporter component. To test this, a western blot was utilized to detect RhaK in the membrane fraction (Figure 2.1). The presence of RhaK could be detected in the membrane fraction and when exposed to a 500 mM salt wash this protein was washed out of the membrane fraction (Figure 2.1). Although this indicates RhaK is peripherally associated with the membrane, unexpectedly the mutants of RhaK that could no longer complement transport were still associated with the membrane in the same manner as the wild-type. This association was also shown to be independent of the presence of RhaT. RhaK72 and RhaK73 displaying an ability to transport rhamnose is therefore not due to

disrupting the membrane association seen in the western blot. This leaves the relevance of this peripheral association with the membrane unknown. The percent of RhaK associated with the membrane does not represent a large fraction of the total RhaK in the cell (Figure 2.1). This interaction may be real and biologically relevant, or it may simply be an artifact of *rhaK* being expressed from a multi-copy (10-20/cell) plasmid.

Using a two-hybrid approach, regions of the N and C terminus of RhaK were found to interact with a region in the N-terminal fragment of RhaT (Figure 3.2). These regions were interestingly also the regions where in-frame insertions generated alleles incapable of transport (Figure 3.2, Figure 3.1, Figure 2.2). When these insert alleles were tested using a two-hybrid system they did not interact with the N-terminal region of RhaT (Figure 3.3). It is clear from this that RhaK interacts with RhaT and that this interaction is required for transport. It is however unclear how this relates to the membrane association.

5.1.4 A putative binding domain has been identified in the N-terminal region of RhaK that is required for an interaction with RhaT

The regions of RhaK interacting with the transporter were analysed using bioinformatics (Figure 3.5), and a putative binding domain was identified in the N-terminal region of RhaK. Using an alanine scanning mutagenesis approach it was shown that this region was not tolerant to substitutions. All of the substitutions significantly affected the interaction between the fragments of RhaK and RhaT (Figure 3.3). This region was defined as a putative binding domain.

Although this sequence is not entirely conserved, it begs the question whether this represents a more widely distributed domain required for interacting with a subsets of ABC

transporters. To date, this sequence seems limited to α -proteobacteria sugar kinases predicted to be involved in rhamnose catabolism (Figure 3.4). Perhaps, similar interactions exist in other systems but are dependent on a different kinase amino acid sequence, one that would have co-evolved with its transporter. Despite being the most divergent protein in the operon (Figure 3.4) all examples of α -proteobacteria RhaK like proteins that have been examined to date appear to be highly conserved in this region (Figure 3.4, Figure 3.5). Within the *Rhizobiaceae*, *S. meliloti* Rm1021 is from one of the most divergent clades examined. It still requires its RhaK (RhaK_{Sm}) for the transport of rhamnose, and RhaK_{Sm} can complement the uptake and catabolism of rhamnose in *R. leguminosarum*. This shows this phenomenon is not an unusual artifact of evolution restricted to *R. leguminosarum*. It is likely at least as widely distributed as the phylogeny included in this thesis (Figure 3.4) and perhaps represents a novel class of transporter.

5.1.5 Some RhaK alleles have intermediate transport phenotypes

It is clear that RhaK interacting with RhaT is not simply an on or off phenomenon, as one allele (*rhaK72*) is less severe in regards to all of its phenotypes ((Rivers and Oresnik 2013) Figure 3.3). It has been shown that this allele has a small but measureable uptake of labelled rhamnose. The RhaK72F1- RhaTF1 two hybrid interaction quantified using beta-galactosidase assays yielded activity double that of the typical background, and double that of the RhaK73F4 allele (Figure 3.3). In a *R. leguminosarum* *rhaK* mutant background the *rhaK72* allele can restore growth on rhamnose as a sole carbon source, though only to an exceedingly slow rate (Rivers and Oresnik 2013). Other alleles confer no practical change in growth or transport phenotypes; still others completely abolish all measureable growth and

uptake. This is consistent with the hypothesis that alterations in amino acid composition of RhaK are leading to changes in the surface shape and/or charge to varying degrees. This in turn is affecting the protein–protein interaction to varying degrees. I believe changes in this interaction are directly resulting in differences in rhamnose uptake rates, and thus growth rates on rhamnose as a sole carbon source. It seems likely that RhaK72 has a weak or transient interaction with RhaT due to less severe alterations in surface charge or shape.

Although all the alanine scanning mutants severely impacted the RhaK fragments' ability to interact with RhaT, the mutagenesis yielded fragments that varied in how significantly they diminished the interaction with the N-terminal fragment of RhaT. It would be interesting to see if these amino acid substitutions also lead to varying intermediate transport and growth phenotypes. I predict that the growth and transport rates would vary in correlation with the two-hybrid beta-galactosidase activity detected. However, we have not made similar mutations to the full length genes to determine the *in vivo* effect.

5.1.6 Developing a model for RhaK dependant transport

Although the mechanism for RhaK dependant transport cannot be elucidated from what has been discovered to date, we have gained a great deal of insight into the nature RhaK and its involvement in transport. I propose a model where RhaK assists in transport by physically interacting at its N- and C-terminal domains with an unidentified sequence near the N- terminal region of RhaT. A putative domain required for the N-terminal domain of RhaK to interact with N-terminal region of RhaT was defined (Figure 3.4). Although these interactions appear to be required for transport, they likely are not sufficient. Evidence suggests that only the catalytically active form of RhaK is able to complement transport

(Richardson and Oresnik 2007 , Rivers and Oresnik 2013). Interestingly the stoichiometry and mechanism of the CUT II class transporter is unknown (Schneider 2001). Recall, CUT II transporters are defined as a fusion of two ABC domains, where the second fused ABC domain is catalytically non-functional due to an amino acid substitution (Schneider 2001). Although the mechanism of the CUT II class transporter is unknown, it is generally assumed to be the same as the CUT I mechanism. It is striking that the CUT II ABC protein has a non-functional second ATPase domain, especially because this CUT II ABC rhamnose transporter requires a kinase protein, and likely one with a functional ATPase domain. It begs the question, does the ATP binding domain from the kinase play a role in the mechanism of the transporter?

The available data are consistent with the hypothesis that the transport and phosphorylation of rhamnose are two subsequent events, and not a simultaneous transport and phosphorylation event. This is supported by the fact that the mutarotase RhaU acts between these two events (Richardson, Carpena et al. 2008). There are many possibilities as to how RhaK could physically interact with RhaT that would be consistent with the two-hybrid data available. For example, two different copies of RhaK may be each interacting with a copy of RhaT, or a single RhaK may be interacting with 2 different copies of RhaT. However, from the data available we also cannot rule out transport being dependant on a multi-protein complex where other interactions are also involved, perhaps involving multiple copies of some proteins. Consistent with this, RhaK is peripherally associated with the membrane, but this interaction is not dependant on the same region of RhaK or on RhaT at all. Mutations that eliminate the RhaK-RhaT interaction do not interfere with the RhaK membrane association. Additionally, RhaK is still found in the membrane fraction of RhaT

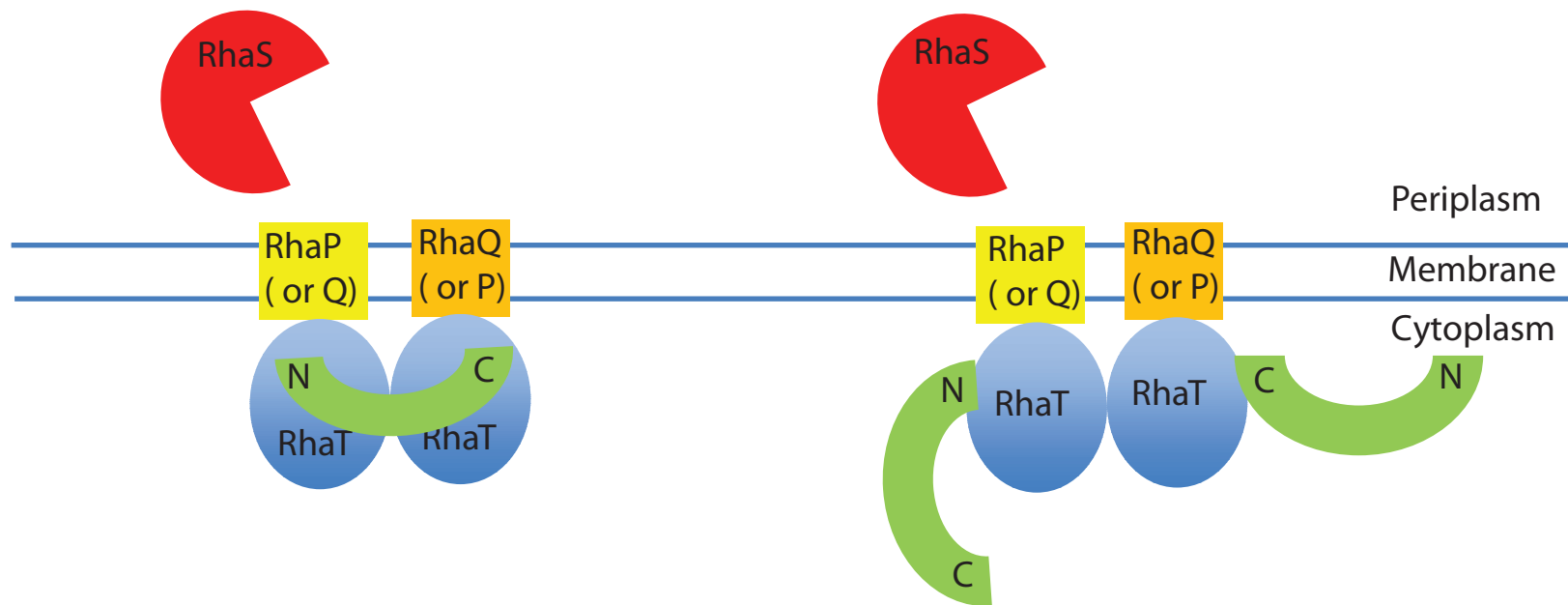


Figure 5.1. RhaK and RhaT interactions consistent with the two-hybrid data

mutants (Figure 2.1). This suggests that if this association is biologically relevant it must involve a protein other than RhaT. If this is not an artifact, it suggests either multiple separate interactions involving RhaK or a multi protein complex.

5.1.7 RhaK_{Sm} and RhaK_{Rl} differ in catabolic function

Upon examination of syntenous operons containing RhaK like proteins, a phylogeny based on their amino acid sequence was generated. It shows that RhaK from *R. leguminosarum* and *S. meliloti* diverge into two different root clades (Figure 3.4). This differs from a previous published phylogeny that suggests these two proteins branch to the same clade, along with the homolog from *Agrobacterium tumefaciens* (Richardson and Oresnik 2007). A new phylogeny was generated from a larger sequence data base now available, using only RhaK like proteins from syntenous operons found in α -proteobacteria. From this it was shown that all *Sinorhizobium* species cluster to one clade, all *Mesorhizobium* to another, and all other sequences into the third clade (Figure 3.4). This demonstrates that the *S. meliloti* kinase and *R. leguminosarum* kinase are actually very different from each other when compared within the α -proteobacteria rhamnase catabolic kinases. In fact, the *R. leguminosarum* kinase is more closely related to the ones found in *Brucella sp.* and *Ochrabacter sp.* than *S. meliloti*. This differs from universal RpoD trees generated. Although initial screens showed the *S. meliloti* RhaK (RhaK_{Sm}) could in fact complement the transport and the kinase function of the *R. leguminosarum* RhaK (RhaK_{Rl}), cosmid complementation analysis anomalies hinted at possible functional differences. Upon direct biochemical analysis, it was shown that the RhaK_{Rl} had activity directly phosphorylating rhamnase, whereas RhaK_{Sm} does not (Table 4.4). Since heterologous complementation is possible and

RhaK_{Sm} cannot phosphorylate rhamnose directly, the primary catabolic function of RhaK, even RhaK_{RI} must not be phosphorylating rhamnose, but phosphorylating rhamnulose as in the canonical rhamnose catabolic pathway (Takagi and Sawada 1964, Badía, Baldomà et al. 1989). This is supported by the fact that both the *R. leguminosarum* and the *S. meliloti* RhaI are highly conserved, and very closely related to the well characterized RhaI found in *Pseudomonas stutzeri*. The *P. stutzeri* isomerase has been shown to catalyse the conversion of rhamnose to rhamnulose. The *P. stutzeri* isomerase structure has been determined, important residues have been identified, and the catalytic mechanism has been predicted (Leang K and Birger Granström T 2004, Yoshida, Yamada et al. 2007). It seems exceedingly unlikely that a sugar with a phosphate at carbon one could fit in or bind to this protein. In addition, direct enzyme assays show that the *R. leguminosarum* and the *S. meliloti* isomerase can convert rhamnose to rhamnulose. This begs the question- what role does the phosphorylated product generated directly from rhamnose serve? It may simply be an oddity of the evolution of RhaK_{RI}, and serve no biological role. It is also possible this product plays a role in regulation. The *R. leguminosarum* rhamnose catabolic operon contains a DeoR-type negative regulator (Richardson, Hynes et al. 2004). This class of regulator is known to respond to phosphorylated catabolic intermediates. Although heterologous complementation of both the kinase and the isomerase were possible, unusual anomalies were encountered when cosmid complementations were attempted. If the cosmid contained a transposon insert in any gene of the operon it could not complement a mutation in the heterologous *S. meliloti* operon. The heterologous genes appear to code proteins that carry out the same primary biochemical functions. They also can heterologously complement mutations when individually over expressed. Because of this, it seems likely these anomalies are due to

differences in regulation. Perhaps the secondary enzymatic function of RhaK_{RI} is relevant in this differential regulation, as DeoR-type negative regulators are known to interact with phosphorylated catabolic intermediate effectors.

5.2 Future work on RhaK dependant rhamnose transport

Although a great deal of insight into RhaK and its role in transport has been discovered, there is much work that still needs to be done to determine the nature of the RhaK-RhaT interaction, as well as the mechanism of the transporter.

The next step in addressing the role of RhaK in transport would be to address the question of how RhaT interacts with RhaK. This could be accomplished by the reverse of what was previously done; an insertional mutagenesis of RhaT. Any of these new mutations affecting RhaT and its interaction with RhaK could be informative in regards to the region of RhaT involved in the protein-protein interaction. However, this would be very time consuming as we do not have a positive screen to carry this out. Each allele that did not restore growth on rhamnose as a sole carbon source would need to be screened for the ability to restore transport, and then assayed for an interaction with RhaK using the two hybrid system. Even then, these mutations may not be informative. Looking for a negative result may only identify proteins so grossly altered that they no longer fold correctly.

Alternatively, as we have a kinase-positive, transport-negative allele of *rhaK* (*rhaK72*, *rhaK73*), a positive screen could be used to identify kinase independent transport positive variants of RhaT. This could be accomplished by generating a bank of random *rhaT* mutants (either via in-frame-inserts or chemical mutagenesis) and co-transforming these with *rhaK72* into a *rhaT*, *rhaK* strain of *R. leguminosarum*. After this, selecting for growth on

rhamnose as a sole carbon source will yield strains where growth will be dependent on the new *rhaT* allele. Transport will either be independent of RhaK, or be a compensatory mutation that allows this new RhaT variant to better interact with RhaK72. Either way, these variants would be informative in regards to the region of RhaT involved in the protein-protein interaction with RhaK.

A bacterial genetics approach was used to address questions about RhaK and its role in transport. This method has been successful in determining a lot about RhaK dependant transport in *R. leguminosarum*. Going forward I believe any major advancements in understanding this model will require structural biology. Any structural information inferred so far has been based off homology modelling. I believe the next important step in this work is to generate a crystal structure for RhaK. This will allow us to confirm all our assumptions about the location of the potential protein-protein binding sites of RhaK, and perhaps further target residues confirmed to be surface exposed. It will also allow us to address how rhamnose and rhamnulose can both bind RhaK. Additionally, any mechanism studies involving this model or other CUT II transporters will require complex structural biology.

Finally, it is still unknown why RhaR_{Sm} appears to be a dominant negative regulator of the *R. leguminosarum* rhamnose catabolic operons. Although this question has been preliminarily explored using beta-galactosidase fusions (Table 4.5), this data may not reflect expression in a wild-type background. This is because the inserts themselves may be affecting gene expression. These inserts render the genes they are in non-functional, and are typically polar on the rest of the same transcript. The regulator RhaR likely responds to a phosphorylated catabolic intermediate (Mortensen, Dandanell et al. 1989). This means the presence of the inserts themselves may be affecting what catabolic intermediates build up,

and thus affect gene expression. Because DeoR type negative regulators respond to phosphorylated intermediates, it may be difficult to determine the roles of the respective RhaR proteins in regulation. This is especially true in *R. leguminosarum* where it appears that RhaK_{RI} can phosphorylate two different sugars. Because of this qPCR would make the best way to further address regulation questions.

Assuming this affect is not simply due to RhaR_{Sm} having a greater affinity for the *R. leguminosarum* DNA sequence, a systematic characterization of gene expression could be informative. Gene expression in heterologous gene replacement backgrounds could be compared to wild type expression levels using qPCR. This could be used to determine what heterologous rhamnose catabolic gene combination leads to RhaR_{Sm} having a dominant negative phenotype in *R. leguminosarum*.

References

- Abramson, J., I. Smirnova, V. Kasho, G. Verner, H. R. Kaback and S. Iwata.** (2003). Structure and Mechanism of the Lactose Permease of *Escherichia coli*. *Science* 301(5633): 610-615.
- Amouyal, M., L. Mortensen, H. Buc and K. Hammer.**(1989). Single and double loop formation when *deoR* repressor binds to its natural operator sites. *Cell* 58(3): 545-551.
- Anderson, M. J., B. DeLaBarre, A. Raghunathan, B. O. Palsson, A. T. Brunger and S. R. Quake.** (2007). Crystal structure of a hyperactive *Escherichia coli* glycerol kinase mutant Gly²³⁰→ Asp obtained using microfluidic crystallization devices. *Biochemistry* 46(19): 5722-5731.
- Anderson, R. L. and V. L. Sapico.** (1975). D-fructose (D-mannose) kinase. *Methods Enzymol* 42: 39-43.
- Badía, J., L. Baldomà, J. Aguilar and A. Boronat.** (1989). Identification of the *rhaA*, *rhaB* and *rhaD* gene products from *Escherichia coli* K-12. *FEMS Microbiol Lett.* 53(3): 253-257.
- Baldani, J. I., R. W. Weaver, M. F. Hynes and B. D. Eardly.** (1992). Utilization of carbon substrates, electrophoretic enzyme patterns, and symbiotic performance of plasmid-cured clover rhizobia. *Appl. Environ. Microbiol.* 58(7): 2308-2314.
- Baldoma, L. and J. Aguilar.** (1988). Metabolism of L-fucose and L-rhamnose in *Escherichia coli*: aerobic-anaerobic regulation of L-lactaldehyde dissimilation. *J. Bacteriol.* 170(1): 416-421.
- Barriere, C., M. Veiga-da-Cunha, N. Pons, E. Guédon, S. A. F. T. van Hijum, J. Kok, O. P. Kuipers, D. S. Ehrlich and P. Renault.** (2005). Fructose utilization in *Lactococcus lactis* as a model for low-GC gram-positive bacteria: its regulator, signal, and DNA-binding site. *J. Bacteriol.* 187(11): 3752-3761.
- Berger, E. A.** (1973). Different Mechanisms of Energy Coupling for the Active Transport of Proline and Glutamine in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 70(5): 1514-1518.
- Berger, E. A. and L. A. Heppel.** (1974). Different mechanisms of energy coupling for the shock-sensitive and shock-resistant amino acid permeases of *Escherichia coli*. *J. Biol. Chem.* 249(24): 7747-7755.
- Beringer, J. E.** (1974). R factor transfer in *Rhizobium leguminosarum*. *J. Gen. Microbiol.* 84: 188-198.
- Beringer, J. E., J. L. Beynon, A. V. Buchanan-Wollason and A. W. B. Johnston.** (1978). Transfer of the drug resistance transposon Tn5 to *Rhizobium*. *Nature* 276: 633-634.

Biemans-Oldehinkel, E., M. K. Doeven and B. Poolman. (2006). ABC transporter architecture and regulatory roles of accessory domains. *FEBS Letters* 580(4): 1023-1035.

Biemans-Oldehinkel, E. and B. Poolman. (2003). On the role of the two extracytoplasmic substrate-binding domains in the ABC transporter OpuA. *EMBO J.* 22(22): 5983-5993.

Bisseling, T., R. vanden Bos, A. van Kammen, M. vander Ploeg, v. Duijn and A. Houwers. (1977). Cytofluorometrical determination of the DNA contents of bacterioids and corresponding broth-cultured *Rhizobium* bacteria. *J. Gen. Microbiol.* 01, : 79–84.

Boos, W. and A. Böhm. (2000). Learning new tricks from an old dog. *Trends Genet.* 16: 404-409.

Boos, W. and H. Shuman. (1998). Maltose/maltodextrin system of *Escherichia coli*: transport, metabolism, and regulation. *Microbiol. Mol. Biol. Rev.* 62: 204-229.

Bordignon, E., M. Grote and E. Schneider. (2010). The maltose ATP-binding cassette transporter in the 21st century--towards a structural dynamic perspective on its mode of action. *Mol Microbiol* 77: 1354-1366.

Bournaud, C., S. M. de Faria, J. M. F. dos Santos, P. Tisseyre, M. Silva, C. Chaintreuil, E. Gross, E. K. James, Y. Prin and L. Moulin. (2013). *Burkholderia* species are the most common and preferred nodulating symbionts of the Piptadenia Group (tribe Mimoseae). *PloS one* 8(5): e63478.

Brom, S., A. G. De Los Santos, T. Stepkowsky, M. Flores, G. Dávila, D. Romero and R. Palacios. (1992). Different plasmids of *Rhizobium leguminosarum* *bv.* *phaseoli* are required for optimal symbiotic performance. *J. Bacteriol.* 174(16): 5183-5189.

Brouwer, M., M. G. L. Elferink and G. T. Robillard. (1982). Phosphoenolpyruvate-dependent fructose phosphotransferase system of *Rhodopseudomonas sphaeroides*: purification and physicochemical and immunochemical characterization of a membrane-associated enzyme I. *Biochemistry* 21(1): 82-88.

Burgess, B. K. (1984). Structure and reactivity of nitrogenase—an overview. *Advances in Nitrogen Fixation Research*, Springer: 103-114.

Chen, J., G. Lu, J. Lin, A. L. Davidson and F. A. Quioco. (2003). A tweezers-like motion of the ATP-binding cassette dimer in an ABC transport cycle. *Mol. Cell* 12 651–661.

Chen, S., M. L. Oldham , A. L. Davidson and J. Chen.(2013). Carbon catabolite repression of the maltose transporter revealed by X-ray crystallography. *Nature* 499 364-370.

Chevenet, F., C. Brun, A. L. Banuls, B. Jacq and R. Christen. (2006). TreeDyn: towards dynamic graphics and annotations for analyses of trees. *BMC Bioinformatics* 7: 439.

- Chiu, T. H. and D. S. Feingold.** (1964). The purification and properties of L-rhamnulokinase. *Biochim. Biophys. Act.* 92(3): 489-497.
- Crooks, G. E., G. Hon, J. Chandonia and S. E. Brenner.** (2004). WebLogo: a sequence logo generator. *Genome Res* 14: 1188-1190.
- D’Haeze, W. and M. Holsters.** (2002). Nod factor structures, responses, and perception during initiation of nodule development. *Glycobiology* 12(6): 79R-105R.
- Dassa, E.** (2003). Phylogenetic and functional classification of ABC (ATP-binding cassette) systems. *ABC Proteins From Bacteria to Man*. London, Academic Press: 3-35.
- Dassa, E. and P. Bouige.** (2001). The ABC of ABCs: a phylogenetic and functional classification of ABC systems in living organisms. *Res. Microbiol.* 152: 211-229
- Dassa, E. and M. Hofnung.** (1985). Sequence of gene *malG* in *E. coli* K12: homologies between integral membrane components from binding protein-dependent transport systems. *EMBO J.* 4(9): 2287.
- Davidson, A. L. and J. Chen.** (2004). ATP-binding cassette transporters in bacteria. *Annu Rev Biochem* 73: 241-268.
- Davidson, A. L. and H. Nikaido.** (1991). Purification and characterization of the membrane-associated components of the maltose transport system from *Escherichia coli*. *J. Biol. Chem.* 266(14): 8946-8951.
- Davis, R. W., D. Botstein and J. R. Roth.** (1980). A manual for genetic engineering: Advanced Bacterial Genetics. Cold Spring Harbor, New York, Cold Spring Harbor Laboratory.
- Dawson, R. J. P. and K. P. Locher.** (2006). Structure of a bacterial multidrug ABC transporter. *Nature* 443(7108): 180-185.
- Dean, M., Y. Hamon and G. Chimini.** (2001). The human ATP-binding cassette (ABC) transporter superfamily. *Journal of Lipid Research* 42(7): 1007-1017.
- Deutscher, J., C. Francke and P. W. Postma.** (2006). How Phosphotransferase System-Related Protein Phosphorylation Regulates Carbohydrate Metabolism in Bacteria. *Microbiol. Mol. Biol. Rev.* 70. (40).
- Di Luccio, E., B. Petschacher, J. Voegtli, H. Chou, H. Stahlberg, H. Nidetzky and D. K. Wilson.** (2007). Structural and kinetic studies of induced fit of xylulose kinase from *Escherichia coli*. *J. Mol. Biol.* 365: 783-798.

- Dische, Z. and E. Borenfreund.** (1951). A new spectrophotometric method for the detection and determination of keto sugars and trioses. *J Biol Chem* 192: 583-587.
- Durham, D. R. and P. V. Phibbs.** (1982). Fractionation and characterization of the phosphoenolpyruvate: fructose 1-phosphotransferase system from *Pseudomonas aeruginosa*. *J. Bacteriol.* 149(2): 534-541.
- Schneider, E. and S. Hunke.** (1998). ATP-binding cassette (ABC) transport systems: functional and structural aspects of the ATP-hydrolysing subunits/ domains. *FEMS Microbiol. Rev.* 22: 1-20.
- Edgar, R. C.** (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32: 1792-1797.
- Egan, S. M. and R. F. Schleif.** (1994). DNA-dependent renaturation of an insoluble DNA binding protein: Identification of the RhaS binding site at *rhaBAD*. *J. Mol. Biol.* 243(5): 821-829.
- Ehrhardt, D. W. and E. M. Atkinson.** (1992). Depolarization of alfalfa root hair membrane potential by *Rhizobium meliloti* Nod factors. *Science* 256(5059): 998-1000.
- Ehrhardt, D. W., R. Wais and S. R. Long.** (1996). Calcium spiking in plant root hairs responding to *Rhizobium* nodulation signals. *Cell* 85(5): 673-681.
- Eitinger, T., D. A. Rodionov, M. Grote and E. Schneider.** (2011). Canonical and ECF-type ATP-binding cassette importers in prokaryotes: diversity in modular organization and cellular functions. *FEMS Microbiol Rev* 35: 3-67.
- Engels, V. and V. F. Wendisch.** (2007). The DeoR-type regulator SugR represses expression of ptsG in *Corynebacterium glutamicum*. *J. Bacteriol.* 189(8): 2955-2966.
- Esseling, J. J., F. G. P. Lhuissier and A. M. C. Emons.** (2003). Nod factor-induced root hair curling: continuous polar growth towards the point of nod factor application. *Plant Physiol.* 132(4): 1982-1988.
- Finan, T. M., E. Hartweg, K. Lemieux, K. Bergman, G. C. Walker and E. R. Signer.** (1984). General transduction in *Rhizobium meliloti*. *J. Bacteriol.* 159: 120-124.
- Finan, T. M., A. M. Hirsch, J. A. Leigh, E. Johansen, G. A. Kuldau, S. Deegan, G. C. Walker and E. R. Signer.** (1985). Symbiotic mutants of *Rhizobium meliloti* that uncouple plant from bacterial differentiation. *Cell* 40: 869-877.
- Finan, T. M., I. Oresnik and A. Bottacin.** (1988). Mutants of *Rhizobium meliloti* defective in succinate metabolism. *J. Bacteriol.* 170(8): 3396-3403.

- Finan, T. M., S. Weidner, K. Wong, J. Buhrmester, P. Chain, F. J. Vorholter, I. Hernández-Lucas, A. Becker, A. Cowie, J. Gouzy, B. Golding and A. Pühler.** (2001). The complete sequence of the 1,683-kb pSymb megaplasmid from the N₂-fixing endosymbiont *S. meliloti*. Proc. Natl. Acad. Sci. USA 98: 9889-9894.
- Fry, J., M. Wood and P. S. Poole.** (2001). Investigation of myo-inositol catabolism in *Rhizobium leguminosarum* bv. *viciae* and its effect on nodulation competitiveness. Mol. Plant-Microbe Interact. 14: 1016-1025.
- Gang, L., J.M. Westbrooks, A.L. Davidson, J. Chen.** (2005). ATP hydrolysis is required to reset the ATP-binding cassette dimer into the resting-state conformation. Proc. Natl. Acad. Sci. USA 102 17969–17974.
- Gadsby, D. C., P. Vergani and L. Csanady.** (2006). The ABC protein turned chloride channel whose failure causes cystic fibrosis. Nature 440 (7083): 477-483.
- Gage, D. J.** (2004). Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. Microbiol. Mol. Biol. Rev. 68(2): 280-300.
- Galibert, F., T. M. Finan, S. R. Long, A. Pühler, P. Abola, F. Ampe, F. Barloy- Hubler, M. J. Barnett, A. Becker, P. Boistard, G. Bothe, M. Boutry, L. Bowser, J. Buhrmester, E. Cadieu, D. Capela, P. Chain, A. Cowie, R. W. Davis, S. Dréano, N. A. Federspiel, R. F. Fisher, S. Gloux, T. Godrie, A. Goffeau, B. Golding, J. Gouzy, M. Gurjal, I. Hernández-Lucas, A. Hong, L. Huizar, R. W. Hyman, T. Jones, D. Kahn, M. L. Kahn, S. Kalman, D. H. Keating, E. Kiss, C. Komp, V. Lelaure, D. Masuy, C. Palm, M. C. Peck, T. M. Pohl, D. Portetelle, B. Purnelle, U. Ramsperger, R. Surzycki, P. Thébault, M. Vandenbol, F. J. Vorholter, S. Weidner, D. H. Wells, K. Wong , K. C. Yeh and J. Batut.** (2001). The composite genome of the legume symbiont *Sinorhizobium meliloti*. Science 293: 668–672.
- Garces, F., F. J. Fernández, A. M. Gómez, R. Pérez-Luque, E. Campos, R. Prohens, J. Aguilar, L. Baldoma, M. Coll and J. Badía.** (2008). Quaternary Structural Transitions in the DeoR-Type Repressor UlaR Control Transcriptional Readout from the l-Ascorbate Utilization Regulon in *Escherichia coli*. Biochemistry 47(44): 11424-11433.
- Geddes, B. A., G. Hauser and I. J. Oresnik.** (2013). Phylogenetic analysis of erythritol catabolic loci within the Rhizobiales and Proteobacteria. BMC Microbiology 13: 46.
- Geddes, B. A. and I. J. Oresnik.** (2012). Genetic characterization of a complex locus necessary for the transport and catabolism of erythritol, adonitol, and L-arabitol in *Sinorhizobium meliloti*. Microbiology 158: 2180-2191.

Geddes, B. A. and I. J. Oresnik. (2014). Physiology, genetics, and biochemistry of carbon metabolism in the alphaproteobacterium *Sinorhizobium meliloti*. *Can J Microbiol.* 60(8): 491-507.

Geddes, B. A., B. S. Pickering, N. J. Poysti, H. Yudistira, H. Collins and I. J. Oresnik. (2010). A locus necessary for the transport and catabolism of erythritol in *Sinorhizobium meliloti*. *Microbiology* 156: 2970-2981.

Geddes, G. A. and I. J. Oresnik. (2012). Inability to catabolize galactose leads to increased ability to compete for nodule occupancy in *Sinorhizobium meliloti*. *J. Bacteriol* 194: 5044-5505.

Gilson, E., G. Alloing, T. Schmidt, J. P. Claverys, R. Dudler and M. Hofnung. (1988). Evidence for high affinity binding-protein dependent transport systems in gram-positive bacteria and in *Mycoplasma*. *EMBO J.* 7(12): 3971.

Gordon, D. M., M. H. Ryder, K. Heinrich and P. J. Murphy.(1996). An experimental test of the rhizopine concept in *Rhizobium meliloti*." *Appl Environ Microbiol* 62: 3991-3996.

Grote, M., Y. Polyhach, G. Jeschke, H.-J. Steinhoff, E. Schneider and E. Bordignon. (2009). Transmembrane Signaling in the Maltose ABC Transporter MalFGK₂-E Periplasmic MalF-P2 Loop Communicates Substrate Availability To The Atp-Bound MalK Dimer. *J. Biol. Chem.* 284(26): 17521-17526.

Grueninger, D. and G. E. Schulz. (2006). Structure and Reaction Mechanism of 1-Rhamnulose Kinase from *Escherichia coli*. *J. Mol. biol.* 359(3): 787-797.

Hanson, P. I. and S. W. Whiteheart. (2005). AAA+ proteins: have engine, will work." *Nat Rev Mol Cell Biol* 6(7): 519-529.

Hengge, R. and W. Boos. (1983). Maltose and lactose transport in *Escherichia coli* Examples of two different types of concentrative transport systems. *Biochim. Biophys. Act.* 737(3): 443-478.

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

Higgins, C. F. and K. J. Litton. (1992). ABC transporters: from microorganisms to man *Annu. Rev. Cell Biol.* 8 67–113.

Higgins, C. F. and K. J. Litton. (2004). The ATP switch model for ABC transporters. *Nat. Struct. Mol. Biol.* 11: 918-926.

Hoischen, C., J. Levin, s. Pitaknarongphorn, J. Reizer and J. M. H. Saier. (1996). Involvement of the central loop of the lactose permease of *Escherichia coli* in its allosteric

regulation by the glucose-specific enzyme IIA of the phosphoenolpyruvate-dependent phosphotransferase system. *J. Bacteriol.* 178: 6082–6086.

Horlacher, R., K. B. Xavier, H. Santos, J. DiRuggiero, M. Kossmann and W. Boos. (1998). Archaeal binding protein-dependent ABC transporter: molecular and biochemical analysis of the trehalose/maltose transport system of the hyperthermophilic archaeon *Thermococcus litoralis*. *J. Bacteriol.* 180(3): 680-689.

House, B. L., M. W. Mortimer and M. L. Kahn. (2004). New recombination methods for *Sinorhizobium meliloti* genetics. *Appl Environ Microbiol.* 70: 2806-2815.

Hurley, J. H. (1996). The sugar kinase/heat shock protein 70/actin suprefamily: Implications of conserved structure for mechanism. *Annu Rev Biophys Biomol Struct* 25: 137-162.

Hurley, J. H., H. R. Faber, D. Worthylake, N. D. Meadow, S. Roseman, D. W. Pettigrew and S. J. Remington. (1993). Structure of the regulatory complex of *Escherichia coli* III^{glc} with glycerol kinase. *Science* 259: 673-677.

Hynes, M. F. (1990). The role of plasmids in competition between strains of *Rhizobium leguminosarum*. *Nitrogen Fixation: Achievements and Objectives.* G. P.M., L. E. Roth, G. Stacey and W. E. Newton, Chapman and Hall, New York: 262.

Hynes, M. F. and M. P. O'Connell. (1990). Host plant effect on competition among strains of *Rhizobium leguminosarum*. *Can J Microbiol.* 36(12): 864-869.

Jacob, A. I., S. A. I. Adhamn, D. S. Capstick, S. R. D. Clark, T. Spence and T. C. Charles.(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: 979-989.

Jiang, G., A. H. Krishnan, Y. W. Kim, T. J. Wacek and H. B. Krishnan. (2001). A functional myo-inositol dehydrogenase gene is required for efficient nitrogen fixation and competitiveness of *Sinorhizobium fredii* USDA191 to nodulate soybean (*Glycine max* [L] Merr.). *J Bacteriol* 183: 2595-2604.

Jones, K. M., H. Kobayashi, B. W. Davies, M. E. Taga and G. C. Walker. (2007). How rhizobial symbionts invade plants: the *Sinorhizobium–Medicago* model. *Nat. Rev. Microbiol.* 5(8): 619-633.

Jones, P. M. and A. M. George. (2004). The ABC transporter structure and mechanism: perspectives on recent research. *Cell Mol Life Sci.* 61(6): 682-699.

Jones, S. and J. M. Thornton. (1996). Principles of protein-protein interactions. *Proc. Natl. Acad. Sci. USA* 93(1): 13-20.

- Kaneko, T., Y. Nakamura, S. Sato, E. Asamizu, T. Kato, S. Sasamoto, A. Watanabe, K. Idesawa, A. Ishikawa, K. Kawashima, T. Kimura, Y. Kishida, C. Kiyokawa, M. Kohara, M. Matsumoto, A. Matsuno, C. Takeuchi, M. Yamada and S. Tabata.** (2000). Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. DNA Res. 7: 331-338.
- Karimova, G., A. Ullmann and D. Ladant.** (2000). A bacterial two-hybrid system that exploits a cAMP signaling cascade in *Escherichia coli*. Methods Enzymol. 328: 59-73.
- Karimova, G., A. Ullmann and D. Ladant.** (2002). A bacterial two-hybrid system based on a cyclic AMP signaling cascade. Protein-Protein Interactions A Molecular Cloning Manual. E. Golemis. Cold Spring Harbor, New York, Cold Spring Harbour Laboratory Press: 477-488.
- Kelley, L. A. and M. J. E. Sternberg.** (2009). Protein structure prediction on the web: a case study using the Phyre server. Nature Protocols 4: 363-371.
- Kereszt, A., P. Mergaert and E. Kondorosi.** (2011). Bacteroid Development in Legume Nodules: Evolution of Mutual Benefit or of Sacrificial Victims? Molecular Plant-Microbe Interactions 24(11): 1300-1309.
- Khare, D., M. Oldham, C. Orelle, A. L. Davidson and J. Chen.** (2009). Alternating access in maltose transporter mediated by rigid-body rotations. Mol Cell 33: 528-536.
- Knee, E. M., F. Gong, M. Gao, M. Teplitski, A. R. Jones, A. Foxworthy, A. J. Mort and W. D. Bauer.** (2001). Root mucilage from pea and its utilization by rhizosphere bacteria as a sole carbon source. Mol. Plant-Microbe Interact. 14: 775-784.
- Kohler, P. R. A., J. Y. Zheng, E. Schoffers and S. Rossbach.** (2010). Inositol catabolism, a key pathway in *Sinorhizobium meliloti* for competitive host nodulation. Appl Environ Microbiol 76: 7972-7980.
- Korndörfer, I. P., W.-D. Fessner and B. W. Matthews.** (2000). The Structure of Rhamnose Isomerase from *Escherichia coli* and its Relation with Xylose Isomerase Illustrates a Change Between Inter and Intra-subunit Complementation During Evolution. J. Mol. Biol. 300(4): 917-933.
- Laemmli, U. K.** (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680-685.
- Landmesser, H., A. Stein, B. Blüschke, M. Brinkmann, S. Hunke and E. Schneider.** (2002). Large-scale purification, dissociation and functional reassembly of the maltose ATP-binding cassette transporter MalFGK₂ of *Salmonella typhimurium*. Biochim. Biophys. Acta. 1565(1): 64-72.

- Leang, K., G. Takada, Y. Fukai, K. Morimoto, T. B. Granström and K. Izumori.** (2004). Novel reactions of l-rhamnose isomerase from *Pseudomonas stutzeri* and its relation with d-xylose isomerase via substrate specificity. *Biochim. Biophys. Act.* 1674(1): 68-77.
- Leang, K., G. Takada, A. Ishimura, M. Okita and K. Izumori.** (2004). Cloning, nucleotide sequence, and overexpression of the L-rhamnose isomerase gene from *Pseudomonas stutzeri* in *Escherichia coli*. *Appl. Environ. Microbiol.* 70(6): 3298-3304.
- Leipe, D. D., Y. I. Wolf, E. V. Koonin and L. Aravind.** (2002). Classification and evolution of P-loop GTPases and related ATPases. *J. Mol. Biol.* 317(1): 41-72.
- Long, S. R.** (1989). *Rhizobium* genetics. *Annu. Rev. Gen.* 23(1): 483-506.
- Lu, G., J. M. Westbrook, A. L. Davidson and J. Chen.** (2005). ATP hydrolysis is required to reset the ATP-binding cassette dimer into the resting-state conformation. *Proc Natl Acad Sci USA* 102: 17969-17974.
- Madsen, L. H., L. Tirichine, A. Jurkiewicz, J. T. Sullivan, A. B. Heckmann, A. S. Bek, C. W. Ronson, E. K. James and J. Stougaard.**(2010). The molecular network governing nodule organogenesis and infection in the model legume *Lotus japonicus*. *Nat Commun.* 1: 10.
- Marger, M. D. and M. H. Saier Jr.** (1993). A major superfamily of transmembrane facilitators that catalyze uniport, symport and antiport. *Trends in Biochem. Sci.* 18(1): 13-20.
- Mauchline, T. H., J. E. Fowler, A. K. East, A. L. Sartor, A. H. F. Hosie, P. S. Poole and T. M. Finan.** (2006). Mapping the *Sinorhizobium meliloti* 1021 solute-binding protein-dependent transportome. *Proc Natl Acad Sci USA* 103: 17933-17938.
- Mergaert, P., T. Uchiumi, B. Alunni, G. Evanno, A. Cheron, O. Catrice, A.-E. Mausset, F. Barloy-Hubler, F. Galibert and A. Kondorosi.** (2006). Eukaryotic control on bacterial cell cycle and differentiation in the *Rhizobium*–legume symbiosis. *Proc. Natl. Acad. Sci. USA* 103(13): 5230-5235.
- Miller, J. H.** (1972). *Experiments in Molecular Genetics*. Cold Springs Harbor, NY, Cold Springs Harbor Laboratory.
- Moënne-Loccoz, Y. and R. Weaver.** (1995). Plasmids influence growth of rhizobia in the rhizosphere of clover. *Soil Biol. and Biochem.* 27(8): 1001-1004.
- Moralejo, P., S. M. Egan, E. Hidalgo and J. Aguilar.** (1993). Sequencing and characterization of a gene cluster encoding the enzymes for L-rhamnose metabolism in *Escherichia coli*. *J. Bacteriol.* 175(17): 5585-5594.

- Moreira, I. S., P. A. Fernandes and M. J. Ramos.** (2007). Hot spots—A review of the protein–protein interface determinant amino-acid residues. *Proteins: Structure, Function, and Bioinformatics* 68(4): 803-812.
- Morse, M. L., K. L. Hill, J. B. Egan and W. Hengstenberg.** (1968). Metabolism of lactose by *Staphylococcus aureus* and its genetic basis. *J. Bacteriol.* 95(6): 2270-2274.
- Mortensen, L., G. Dandanell and K. Hammer.** (1989). Purification and characterization of the deoR repressor of *Escherichia coli*. *EMBO J.* 8(1): 325.
- Nikaido, H.** (2002). How are the ABC transporters energized? *Proc. Natl. Acad. Sci. USA* 99(15): 9609-9610.
- Nikaido, H. and M. H. Saier Jr.** (1992). Transport proteins in bacteria: common themes in their design. *Science* 258(5084): 936-942.
- Oldham, M. L., D. Khare, F. A. Quijcho, A. L. Davidson and J. Chen.** (2007). Crystal structure of a catalytic intermediate of the maltose transporter. *Nature* 450: 515-522.
- Oldroyd, G. E. D. and J. A. Downie.** (2008). Coordinating nodule morphogenesis with rhizobial infection in legumes. *Ann. Rev. Plant Biol* 50: 519-546.
- Oldroyd, G. E. D., J. D. Marray, P. S. Poole and J. A. Downie.** (2011). The rules of engagement in the legume-Rhizobial symbiosis. *Annu Rev Genet* 45: 119-144.
- Oono, R. and R. F. Denison.** (2010). Comparing symbiotic efficiency between swollen versus nonswollen rhizobial bacteroids. *Plant Physiol.* 154(3): 1541-1548.
- Oresnik, I. J., T. C. Charles and T. M. Finan.** (1994). Second site mutations specifically suppress the Fix⁻ phenotype of *Rhizobium meliloti ndvF* mutations on alfalfa: identification of a conditional *ndvF*-dependent mucoid colony phenotype." *Genetics* 136(4): 1233-1243.
- Oresnik, I. J. and D. B. Layzell.** (1994). Composition and distribution of adenylates in soybean (*Glycine max L.*) nodule tissue. *Plant Physiol.* 104(1): 217-225.
- Oresnik, I. J., L. A. Pacarynuk, S. A. P. O'Brien, C. K. Yost and M. F. Hynes.** (1998). Plasmid encoded catabolic genes in *Rhizobium leguminosarum bv. trifolii*: evidence for a plant-inducible rhamnase locus involved in competition for nodulation. *Mol. Plant-Microbe Interact.* 11: 1175-1185.
- Panagiotidis, C. H., W. Boos and H. A. Shuman.** (1998). The ATP-binding cassette subunit of the maltose transporter MalK antagonizes MalT, the activator of the *Escherichia coli mal* regulon. *Mol. Microbiol.* 30(3): 535-546.

- Park, Y., Y. J. Cho, T. Ahn and C. Park.** (1999). Molecular interactions in ribose transport: the binding protein module symmetrically associates with the homodimeric membrane transporter. *EMBO J.* 18(15): 4149-4156.
- Perret, X., C. Staehelin and W. J. Broughton.** (2000). Molecular basis of symbiotic promiscuity. *Microbiol. Mol Biol Rev.* 64(1): 180-201.
- Peters, J. W., K. Fisher and D. R. Dean.** (1995). Nitrogenase structure and function: a biochemical-genetic perspective. *Annu Rev Microbiol.* 49(1): 335-366.
- Pettigrew, D. W., G. B. Smith, K. P. Thomas and D. C. Dods.** (1998). Conserved active site aspartates and domain-domain interactions in regulatory properties of the sugar kinase superfamily. *Arch. Biochem. Biophys.* 349: 236-245.
- Pickering, B. S. and I. J. Oresnik.** (2008). Formate-dependent autotrophic growth in *S. meliloti*. *J Bacteriol* 190: 6409-6418.
- Pinedo, C. A., R. M. Bringhurst and D. J. Gage.** (2008). *Sinorhizobium meliloti* mutants lacking PTS enzymes HPr or EIIA are altered in diverse processes including carbon metabolism, cobalt requirements and succinoglycan production. *J. Bacteriol.*
- Postma, P. W. and J. W. Lengeler.** (1985). Phosphoenolpyruvate: Carbohydrate Phosphotransferase System of Bacteria. *Microbiol. Rev.* 49(3): 232-269.
- Power, J.** (1967). The L-rhamnose genetic system in *Escherichia coli* K-12. *Genetics* 55(3): 557.
- Poysti, N. J., E. D. Loewen, Z. Wang and I. J. Oresnik.** (2007). *Sinorhizobium meliloti* pSymB carries genes necessary for arabinose transport and catabolism. *Microbiology* 153: 727-736.
- Poysti, N. J. and I. J. Oresnik.** (2007). Characterization of *Sinorhizobium meliloti* triose phosphate isomerase genes. *J Bacteriol* 189: 3445-3451.
- Prell, J., G. Mulley, F. Haufe, J. White, A. Williams, R. Karunakaran, J. Downie and P. Poole.** (2012). The PTSNtr system globally regulates ATP-dependent transporters in *Rhizobium leguminosarum*. *Mol. Microbiol.* 84(1): 117-129.
- Quentin, Y. and G. Fichant.** (2000). ABCdb: an ABC transporter database. *J Mol Microbiol Biotechnol.* 2(4): 501-504.
- Ramachandran, V. K., A. K. East, R. Karunakaran, J. A. Downie and P. Poole.** (2011). Adaptation of *Rhizobium leguminosarum* to peas, alfalfa, and sugar beet rhizospheres investigated by comparative transcriptomics. *Genome Biology* 12: R106.

- Ray, W. K. and T. J. Larson.** (2004). Application of AgaR repressor and dominant repressor variants for verification of a gene cluster involved in N-acetylgalactosamine metabolism in *Escherichia coli* K-12. *Mol. Microbiol.* 51(3): 813-826.
- Rea, P. A.** (2007). Plant ATP-binding cassette transporters. *Annu Rev Plant Biol* 58: 347-375.
- Richardson, J. S., X. Carpena, J. Switalta, R. Perez-Luque, L. J. Donald, P. C. Loewen and I. J. Oresnik.** (2008). RhaU of *Rhizobium leguminosarum* is a rhamnose mutarotase. *J Bacteriol* 190: 2903-2910.
- Richardson, J. S., M. F. Hynes and I. J. Oresnik.** (2004). A genetic locus necessary for rhamnose uptake and catabolism in *Rhizobium leguminosarum* *bv. trifolii*. *J. Bacteriol.* 186: 8433-8442.
- Richardson, J. S. and I. J. Oresnik.** (2007). L-rhamnose transport in *Rhizobium leguminosarum* is dependent upon RhaK, a sugar kinase. *J Bacteriol* 189: 8437-8446.
- Richet, E., A. L. Davidson and N. Joly.** (2012). The ABC transporter MalFGK₂ sequesters the MalT transcription factor at the membrane in the absence of cognate substrate. *Mol Microbiol* 85: 632-647.
- Rivers, D. and I. J. Oresnik.** (2013). Carbohydrate Kinase (RhaK)-Dependent ABC Transport of Rhamnose in *Rhizobium leguminosarum* Demonstrates Genetic Separation of Kinase and Transport Activities. *J. bacteriol.* 195(15): 3424-3432.
- Rodionova, I. A., X. Li, V. Thiel, S. Stolyar, K. Stanton, J. K. Fredrickson, D. A. Bryant, A. L. Osterman, A. A. Best and D. A. Rodionov.** (2013). Comparative genomics and functional analysis of rhamnose catabolic pathways and regulons in bacteria. *Front Microbiol.* 4.
- Ryu, K.-S., J.-I. Kim, S.-J. Cho, D. Park, C. Park, H.-K. Cheong, J.-O. Lee and B.-S. Choi.** (2005). Structural Insights into the Monosaccharide Specificity of *Escherichia coli* Rhamnose Mutarotase. *J. Mol. Biol.* 349(1): 153-162.
- Saier, M. H., B. U. Feucht and S. Roseman.** (1971). Phosphoenolpyruvate-dependent Fructose Phosphorylation in Photosynthetic Bacteria. *J. Biol. Chem.* 246(24): 7819-7821.
- Samanta, S., T. Ayvaz, M. Reyes, H. A. Shuman, J. Chen and A. L. Davidson.** (2003). Disulfide cross-linking reveals a site of stable interaction between C-terminal regulatory domains of the two MalK subunits in the maltose transport complex. *J Biol Chem* 278: 35265-35271.

Sambrook, J. and D. W. Russell. (2001). *Molecular Cloning A Laboratory Manual* 3rd ed. Cold Spring Harbour, Cold Spring Harbour Laboratory Press.

Saurin, W., M. Hofnung and E. Dassa (1999). Getting in or out: early segregation between importers and exporters in the evolution of ATP-binding cassette (ABC) transporters. *J Mol Evol* 48(1): 22-41.

Sawada, H. and Y. Takagi. (1964). The metabolism of L-rhamnose in *E. coli* III. L rhamnULOse-phosphate aldolase. *Biochim. Biophys. Acta* 64: 26-32.

Schneider, E. (2001). ABC transporters catalyzing carbohydrate uptake. *Res. Microbiol.* 152 303–310.

Schneider, E. and S. Hunke. (1998). ATP-binding cassette (ABC) transport systems: functional and structural aspects of the ATP-hydrolysing subunits/domains. *FEMS Microbiol. Rev.* 22: 1-20.

Schrödinger, L. L. C. (2012). The PyMOL Molecular Graphics System.

Schroeder, B. K., B. L. House, M. W. Mortimer, S. N. Yurgel, S. C. Maloney, K. L. Ward and M. L. Kahn. (2005). Development of a functional genomics platform for *Sinorhizobium meliloti*: Construction of an ORFeome. *Appl Environ Microbiol.* 71: 5858-5864.

Schuurman-Wolters, G. K. and B. Poolman. (2005). Substrate Specificity and Ionic Regulation of GlnPQ from *Lactococcus lactis* An Atp-Binding Cassette Transporter With Four Extracytoplasmic Substrate-Binding Domains. *J. Biol. Chem.* 280(25): 23785-23790.

Seok, Y.-J., J. Sun, H. R. Kaback and A. Peterkofsky. (1997). Topology of allosteric regulation of lactose permease. *Proc. Natl. Acad. Sci. USA* 94: 13515–13519.

Simossis, V. A., J. Kleinjung and J. Heringa. (2005). Homology-extended sequence alignment. *Nucleic Acids Res.* 33: 816-824.

Sondej, M., J. Z. Sun, Y.-J. Seok, H. R. Kaback and A. Peterkofsky. (1999). Deduction of consensus binding sequences on proteins that bind IIA^{Glc} of the phosphoenolpyruvate:sugar phosphotransferase system by cysteine scanning mutagenesis of *Escherichia coli* lactose permease. ." *Proc. Natl. Acad. Sci. USA* 96: 3525–3530.

Sondej, M., J. L. Vazquez-Ibar, A. Farshidi, A. Peterkofsky and H. R. Kaback. (2003). Characterization of a lactose permease mutant that binds IIA^{Glc} in the absence of ligand. *Biochemistry* 42: 9153–9159.

- Sondej, M., A. B. Weinglass, A. Peterkofsky and H. R. Kaback.** (2002). Binding of enzyme IIA^{Glc}, a component of the phosphoenolpyruvate:sugar phosphotransferase system, to the *Escherichia coli* lactose permease. *Biochemistry* 41: 5556–5565.
- Stover, C. K., X. Q. Pham, A. L. Erwin, S. D. Mizoguchi, P. Warrenner, M. J. Hickey, F. S. Brinkman, W. O. Hufnagle, D. J. Kowalik, M. Lagrou, R. L. Garber, L. Goltry, E. Tolentino, S. Westbrook-Wadman, Y. Yuan, L. L. Brody, S. N. Coulter, K. R. Folger, A. Kas, K. Larbig, R. Lim, K. Smith, D. Spencer, G. K. Wong, Z. Wu, I. T. Paulsen, J. Reizer, M. H. Saier, R. E. Hancock, S. Lory and M. V. Olson.** (2000). Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. *Nature* 406: 959-964.
- Sullivan, J. T., B. D. Eardly, P. Van Berkum and C. W. Ronson.** (1996). Four unnamed species of nonsymbiotic rhizobia isolated from the rhizosphere of *Lotus corniculatus*. *Appl. Environ. Microbiol.* 62(8): 2818-2825.
- Sullivan, J. T. and C. W. Ronson.** (1998). Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a phe-tRNA gene. *Proc. Natl. Acad. Sci. USA* 95(9): 5145-5149.
- Sundararajan, T. A., A. M. C. Rapin and H. M. Kalckar.** (1962). Biochemical observation on *E. coli* mutants defective in uridine diphosphoglucose. *Proc. Natl. Acad. Sci. USA* 48: 2187-2192.
- Szmelcman, S., M. Schwartz, T. J. Silhavy and W. Boos.** (1976). Maltose transport in *Escherichia coli* K12. *European Journal of Biochemistry* 65(1): 13-19.
- Takagi, Y. and H. Sawada.** (1964). The metabolism of L-rhamnose in *E. coli* I. L-rhamnose isomerase. *Biochim. Biophys. Acta* 92: 10-17.
- Takagi, Y. and H. Sawada.** (1964). The metabolism of L-rhamnose in *Escherichia coli* II l-Rhamnulose kinase. *Biochim. Biophys. Act.* 92: 18-25.
- Talavera, G. and J. Castresana.** (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst Biol* 56: 564-577.
- Tate, C. G., J. A. Muiry and P. J. Henderson.** (1992). Mapping, cloning, expression, and sequencing of the *rhaT* gene, which encodes a novel L-rhamnose-H⁺ transport protein in *Salmonella typhimurium* and *Escherichia coli*. *J. Biol. Chem.* 267(10): 6923-6932.
- Terpolilli, J. J., G. A. Hood and P. S. Poole.** (2012). What determines the efficiency of N₂-fixing Rhizobium-legume symbioses? *Adv Microb Physiol.* 60: 326.

- Thomas, G. H.** (2010). Homes for the orphans: utilization of multiple substrate-binding proteins by ABC transporters. *Mol. Microbiol.* 75(1).
- Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin and D. G. Higgins.** (1997). The CLUSTAL-X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25: 4876-4882.
- Tobin, J. F. and R. F. Schleif.** (1990). Purification and properties of RhaR, the positive regulator of the l-rhamnose operons of *Escherichia coli*. *J. Mol. Biol.* 211(1): 75-89.
- Triplett, E. W. and M. Sadowsky.** (1992). Genetics of competition for nodulation of legumes. *Annu. Rev. Microbiol.* 46: 399-428.
- Udvardi, M. and P. S. Poole.** (2013). Transport and metabolism in legume-rhizobia symbioses. *Annu. Rev. Plant Biol.* 64: 781-805.
- Valentin-Hansen, P., B. Albrechtsen and J. E. L. Larsen.** (1986). DNA-protein recognition: demonstration of three genetically separated operator elements that are required for repression of the *Escherichia coli deoCABD* promoters by the DeoR repressor. *EMBO J.* 5(8): 2015.
- Valentin-Hansen, P., P. Højrup and S. Short.** (1985). The primary structure of the DeoR repressor from *Escherichia coli* K-12. *Nucleic Acids Res.* 13(16): 5927-5936.
- van der Heide, T. and B. Poolman.** (2002). ABC transporters: one, two or four extracytoplasmic substrate-binding sites? *EMBO reports* 3(10): 938-943.
- Vanderlinde, E. M., M. F. Hynes and C. K. Yost.** (2013). Homoserine catabolism by *Rhizobium leguminosarum* bv. *viciae* 3841 requires a plasmid-borne gene cluster that also affects competitiveness for nodulation. *Environ Microbiol.*
- Vincent, J. M.** (1970). A manual for the practical study of root-nodule bacteria. Oxford, England, Blackwell Scientific Publications.
- Walker, J. E., M. Saraste, M. J. Runswick and N. J. Gay.** (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* 1(8): 945.
- Wang, D., J. Griffitts, C. Starker, E. Fedorova, E. Limpens, S. Ivanov, T. Bisseling and S. Long.** (2010). A nodule-specific protein secretory pathway required for nitrogen-fixing symbiosis. *Science* 327(5969): 1126-1129.
- Watanabe, S. and K. Makino.** (2009). Novel modified version of nonphosphorylated sugar metabolism—an alternative l-rhamnose pathway of *Sphingomonas* sp. *FEBS journal* 276(6): 1554-1567.

Watanabe, S., M. Saimura and K. Makino. (2008). Eukaryotic and bacterial gene clusters related to an alternative pathway of nonphosphorylated L-rhamnose metabolism. *J. Biol. Chem.* 283(29): 20372-20382.

Wiebo, J., M. Marek-Kozaczuk, A. Kubik-Komar and A. Skorupska. (2007). Increased metabolic potential of *Rhizobium* spp. is associated with bacterial competitiveness. *Can J Microbiol* 53: 957-967.

Willems, A. (2006). The taxonomy of rhizobia: an overview. *Plant and Soil* 287(1-2): 3-14.

Xavier, K. B., L. O. Martins, R. Peist, M. Kossmann, W. Boos and H. Santos. (1996). High-affinity maltose/trehalose transport system in the hyperthermophilic archaeon *Thermococcus litoralis*. *J. Bacteriol.* 178(16): 4773-4777.

Yan, N. (2013). Structural advances for the major facilitator superfamily (MFS) transporters. *Trends in Biochem. Sci.* 38(3): 151-159.

Yoshida, H., M. Yamada, Y. Ohyama, G. Takada, K. Izumori and S. Kamitori. (2007). The Structures of l-Rhamnose Isomerase from *Pseudomonas stutzeri* in Complexes with l-Rhamnose and d-Allose Provide Insights into Broad Substrate Specificity. *Journal of molecular biology* 365(5): 1505-1516.

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

Young, J. M., L. D. Kuykendall, E. Martinez-Romero, A. Kerr and H. Sawada. (2001). A revision of *Rhizobium* Frank 1889, with an emended description of the genus, and the inclusion of all species of *Agrobacterium* Conn 1942 and *Allorhizobium undicola* de Lajudie et al. 1998 as new combinations: *Rhizobium radiobacter*, *R. rhizogenes*, *R. rubi*, *R. undicola* and *R. vitis*. *Int J Syst Evol Microbiol.* 51(1): 89-103.

Young, J. P. W., L. C. Crossman, A. W. Johnston, N. R. Thomson, Z. F. Ghazoui, K. H. Hull, M. Wexler, A. R. J. Curson, J. D. Todd, P. S. Poole, T. H. Mauchline, A. K. East, M. A. Quail, C. Churcher, C. Arrowsmith, I. Cherevach, T. Chillingworth, K. Clarke, A. Cronin, P. Davis, A. Fraser, Z. Hance, H. Hauser, K. Jagels, S. M. Moule, K., H. Norbertczak, S. Whitehead and J. Parkhill. (2006). The genome of *Rhizobium leguminosarum* has recognizable core and accessory components." *Genome Biology* 7: R34.