Characterization of L-Rhamnose Catabolism in *Rhizobium leguminosarum*

By

Jason Scott Richardson

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

Ph. D.

Department of Microbiology
University of Manitoba
Winnipeg, Manitoba, Canada

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION

Characterization of L-Rhamnose Catabolism in *Rhizobium leguminosarum*

BY

Jason Scott Richardson

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirement of the degree

DOCTOR OF PHILOSOPHY

Abstract

Rhizobium leguminosarum is a gram-negative soil bacterium that can form nitrogen-fixing nodules on various species of clover. *R. leguminosarum* mutants unable to catabolize L-rhamnose do not compete effectively for nodule occupancy versus wild-type strains. The L-rhamnose catabolic locus encodes two divergently transcribed transcripts encoding *rhaRSTPQUK* and *rhaDI* spanning 10,959 base pairs. This locus is induced by L-rhamnose and regulated by a DeoR-type negative regulator RhaR. Regulation was also observed by a terminator-like sequence located between *rhaS* and *rhaT*. RNA dot blot analysis demonstrated that this terminator-like sequence is correlated with transcript attenuation only under non-inducing conditions.

Respectively, *rhaSTPQ* encodes a periplasmic binding protein, ATPase and two transmembrane proteins; the components of an ABC transporter. Transport assays utilizing tritiated L-rhamnose demonstrated that uptake of L-rhamnose was inducible and dependent upon the presence of this ABC transporter. Unlike other characterized ABC transporters, import of L-rhamnose is also dependant upon a kinase encoded by *rhaK*. Strains carrying *rhaK* mutations were unable to transport L-rhamnose. An *in vitro* assay for RhaK activity showed that: RhaK activity is consistent with RhaK catalyzing a primary step in L-rhamnose catabolism; and the biochemical activity of RhaK is necessary for L-rhamnose transport. The absence of a functional rhamnose kinase did not affect transcription of the ABC transporter components.

rhaU encodes a mutarotase that catalyses α- to β-anomerization of L-rhamnose C-1 carbon hydroxyl moiety. A Δ*rhaU* strain revealed a slower growth phenotype versus the wild-type strain although the ultimate cell yield was equivalent. Transport of L-

rhamnose and the rate of L-rhamnose phosphorylation were unaffected in the absence of RhaU.

Catabolism of L-rhamnose differs from previously described methyl-pentose catabolic pathways. Mutarotation by RhaU follows transport but occurs before phosphorylation by RhaK, suggesting RhaK has two distinct uncoupled roles: transport; and phosphorylation of L-rhamnose.

Acknowledgments

I would like to thank Dr. Ivan Oresnik for his tremendous support and guidance throughout my 'career' as a graduate student. I owe you a debt of gratitude for taking a chance on me and giving me an opportunity to pursue my dream! You provided an ideal environment to work hard while still having fun while doing it. Thank you for enabling me to attend many out of province conferences and for your genuine interest in my overall graduate student experience. I would also like to thank the other members of my committee Dr. Court and Dr. Fernando for their support in and out of committee meetings, thanks to Dr. Charles for traveling here to be an external committee member. I am also appreciative of many other staff members who have given me advice on various experiments and procedures especially Dr. Peter Loewen and Jacek Switala for their collaborative efforts with various aspects of my project. I would like to acknowledge NSERC for the grants awarded to Dr. Ivan Oresnik that supported me and my work. I can not thank Mark Miller Williams, Nathan Poysti, Erin Loewen and Brad Pickering enough for making me laugh out loud to the point of tears almost every day, thanks guys! I would like to thank the project students that come through the lab helping to progress our projects. I would like to thank my friends and family for the support they have given me, without you I would not have been able to accomplish my goals. To Carolyn, thanks for your love and patience, thanks for supporting me in more ways than you know! Elly thanks for keeping me relatively sane through the thesis writing process.

Dedication

To my parents:

Thank you for your support in every facet of my life. I am blessed to have parents that give so much love to their children. Mom, thanks for giving me great advice. I credit you for instilling in me an attitude to never give up, for teaching me not to settle for anything less than exactly what I want, and for helping me become whatever I dreamed of becoming. Dad, thanks for always saying how proud you are of me, and for understanding that going golfing, to movies or enjoying a glass of wine together is so much more than just that! Thanks!

Table of Contents

List of Tables

List of Figures

Lists of Abbreviations

Å = angstrom
ABC = ATP binding cassette
ADP = adenosine diphosphate
AMP = adenosine monophosphate
AMP-PNP = adenylylimidodiphosphate
Ap = ampicillin
ATP = adenosine triphosphate
BLAST = basic local alignment search tool
bp = base pair
bv. = biovar
°C = degrees Celsius
CDART = conserved domain architecture retrieval tool
CFTR = cystic fibrosis conductance regulator
Ci = curie
CSPD = disodium 3-(4-methoxyspiro(1,2-dioxetane-3,2(5-chloro)tricyclo decan)-4-yl)phenyl phosphate
ddH_2O = double distilled water
DEPC = diethyl pyrocarbonate
DNA = deoxyribonucleic acid
dpm = disintegrations per minute
EDTA = ethylenediaminetetraacetic acid
FBP = fructose 1,6-bisphosphate
FGGY = ATPase belonging to the sugar kinase-hsp70-actin superfamily
FPLC = fast protein liquid chromatography
g = gravity
gFW = gram fresh weight
Gm = gentamicin
HEPES = 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
hsp = heat shock protein
IPTG = isopropyl β-D-1-thiogalactopyranoside
kb = kilobase
Km = kanamycin
L = liter
LAOBP = lysine-arginine-ornithine-binding-protein
LB = Luria-Bertani broth
LPS = lipopolysaccharide
m = meter
M = molar units (mol/L)
Mg^{2+} = magnesium ion
NAD^+(P) = nicotinamide adenine dinucleotide (phosphate)
NBD = nucleotide binding domain
Nm = neomycin
OD = optical density
OH = hydroxyl moiety

Lists of Abbreviations (continued)

OM = outer membrane
ORF = open reading frame
PBP = periplasmic binding proteins
PCI = phenol chloroform isoamylalcohol
PCR = polymerase chain reaction
P-gp = P-glycoprotein
PHYLIP = phylogeny inference package
P-loop = pyrophosphate loop
PRALINE = profile alignment
psi = pound per square inch
PSI-BLAST = position-specific iterated BLAST
PTS = phosphoenolpyruvate:glycose phosphotransferase system
Rf = retardation factor
RNA = ribonucleic acid
rpm = revolutions per minute
SDS = sodium dodecyl sulphate
SDS-PAGE = sodium dodecyl sulfate-polyacrylamide gel electrophoresis
Sm = streptomycin
SSC = sodium chloride-sodium citrate
TAE = tris-acetate-EDTA
Tc = tetracycline
TM = melting temperature
TMD = transmembrane domains
Tris = 2-amino-2-hydroxymethyl-1,3-propanediol
TY = tryptone yeast
VMM = Vincent's minimal media

List of Copyrighted Material for which Permission was Obtained

Figure 1.2. Model of ATPase Rad50 dimer association (**Page 30**). Reprinted from Biochimica et biophysica acta, Volume 1561, Kerr, I. D., Structure and association of ATP-binding cassette transporter nucleotide-binding domains, pp. 47-64, copyright 2002, with permission from Elsevier.

Figure 1.3. Consensus of the NBD fold and 3-D structure of ABC transporter ABC proteins NBD monomer (**Page 32**). Reprinted from Cellular and Molecular Life Sciences (CMLS), volume 61(6), Jones, P. M. and George, A. M., The ABC transporter structure and mechanism: perspectives on recent research, pp. 682-699, copyright 2004, with permission from Birkhäuser Publishing Ltd.

Figure 1.4. ABC ATPase catalytic core (**Page 33**). Reprinted from Cellular and Molecular Life Sciences (CMLS), volume 61(6), Jones, P. M. and George, A. M., The ABC transporter structure and mechanism: perspectives on recent research, pp. 682-699, copyright 2004, with permission from Birkhäuser Publishing Ltd.

Figure 1.5. TMD to ABC protein interface of BtuC-BtuD vitamin B_{12} ABC transporter (**Page 41**). Reprinted by permission of Federation of the European Biochemical Societies from ABC transporter architecture and mechanism: implications from the crystal structures of BtuCD and BtuF, by Locher, K. P. and Borths, E., FEBS Letters, volume 564(3), pp. 264-268, copyright 2004.

Figure 1.7. Active site configuration of L-rhamnose mutarotase YiiL in *E. coli* (**Page 51**). Reprinted from the Journal of Molecular Biology, volume 349(1), Ryu, K. S., Kim, J. I., Cho, S. J., Park, D., Park, C., Cheong, H. K., Lee, J. O. and Choi, B. S., Structural insights into the monosaccharide specificity of *Escherichia coli* rhamnose mutarotase, pp. 153-162, copyright 2005, with permission from Elsevier.

Figure 1.8. Representation of actin highlighting the common motifs of the kinase-hsp70-actin protein superfamily (**Page 56**). Reprinted from The Proceedings of the National Academy of Sciences of the United States of America, volume 89, Bork, P., C. Sander, and A. Valencia, An ATPase domain common to prokaryotic cell cycle proteins, sugar kinases, actin, and Hsp70 heat shock proteins, pp. 7290-7294, copyright 1992, with permission by Peer Bork.

Figure 1.9. Diagrammatic representation of the topology of the conserved core in members of the kinase-hsp70-actin protein superfamily (**Page 57**). Reprinted, with permission, from the Annual Review of Biophysics and Biomolecular Structure, Volume 25 © 1996 by Annual Reviews www.annualreviews.org.

Figure 1.10. Glycerol and ADP binding to glycerol kinase within binding pocket (**Page 63**). Reprinted (abstracted/excerpted) with permission from Science, volume 259(5059), pp. 673-677, Hurley, J. H., Faber, H. R., Worthylake, D., Meadow, N. D., Roseman, S., Pettigrew, D. W. and Remington, S. J., Structure of the regulatory complex of

Escherichia coli IIIGlc with glycerol kinase, Copyright 1993, AAAS. Reprint permission also granted by James Hurley.

Figure 1.11. A β-hairpin in glycerol kinase containing crucial for γ-phosphate contact of ATP (**Page 64**). Reprinted from Biochemical and Biophysical Research Communications, volume 259(3), Mao, C., Ozer, Z., Zhou, M. and Uckun, F. M., X-Ray structure of glycerol kinase complexed with an ATP analog implies a novel mechanism for the ATP-dependent glycerol phosphorylation by glycerol kinase, pp. 640-644, Copyright 2004, with permission from Elsevier.

Figure 1.12. Model for catalytic mechanism of phosphotransferase in glycerol kinase (**Page 65**). Reprinted from Biochemical and Biophysical Research Communications, volume 259(3), Mao, C., Ozer, Z., Zhou, M. and Uckun, F. M., X-Ray structure of glycerol kinase complexed with an ATP analog implies a novel mechanism for the ATP-dependent glycerol phosphorylation by glycerol kinase, 640-644, copyright 2004, with permission from Elsevier.

Chapter 2. Section 2.1-2.4 text, **Tables 2.1-2.5**, and **Figures 1.1-1.12 (Page 69-103).** Reproduced from the Journal of Bacteriology, volume 186, Richardson, J. S., M. F. Hynes, and I. J. Oresnik, A genetic locus necessary for rhamnose uptake and catabolism in *Rhizobium leguminosarum* bv. *trifolii*, pp. 8433-8442, Copyright 2004, with permission from the American Society for Microbiology.

Chapter 1

Literature Review

1.1 Rhizobia and symbiosis

1.1.1 Nitrogen fixation

Nitrogen fixation is the process of converting relatively inert atmospheric nitrogen (N_2) into more biologically accessible forms of nitrogen, like ammonia (NH_3) and nitrate (NO_3^-) that can be converted to protein and nucleic acids. The majority of fixed nitrogen on earth is now generated industrially by humans for use in agriculture as fertilizers and for generation of explosives (Vitousek 1997). Natural mechanisms of nitrogen fixation occur abiotically by lightning and mostly by biological nitrogen fixation (Vitousek 1997).

1.1.2 Biological nitrogen fixation

In biological nitrogen fixation, dinitrogen is converted to ammonia in the following reaction: $N_2 + 8\ H^+ + 8\ e^- + 16\ ATP \rightarrow 2\ NH_3 + H_2 + 16\ ADP + 16\ P_i$. This energetically expensive reaction is enzymatically catalyzed by nitrogenase enzymes. All biological nitrogen fixing prokaryotes have the genes encoding these nitrogenase enzymes. Nitrogenases are inhibited by oxygen, so nitrogen fixation must occur in a relatively low anoxic environment (Kuzma *et al.* 1993).

1.1.3 Symbiotic association of rhizobia and legumes

A large, diverse group of prokaryotes including bacteria and archaea can fix nitrogen biologically (Widmer *et al.* 1999, Zehr *et al.* 2003). The majority of the nitrogen fixing prokaryotes belongs to the α-proteobacteria *Rhizobiaceae*. They form symbiotic associations through root nodules with leguminous plants. It is within these nodules that oxygen levels are kept low enabling nitrogenase enzymes to function.

Symbiotic association involves the rhizobia entering the host plant root cells where the bacterium receives shelter and sugars; in return the rhizobia provide the plant with fixed nitrogen (Hirsch 1992, Quispel 1974, Verma 1988).

There are also examples of filamentous *Actinobacteria* from the order *Actinomycetales* called *Frankia* that form symbiotic association and fix nitrogen via root nodules with woody shrubs or trees of actinorhizal plants (Benson and Silvester 1993), in addition to heterocyst-forming cyanobacteria capable of fixing nitrogen (Smith 1982). It has been suggested that non-rhizobia species have acquired nodulation genes by horizontal gene transfer thereby enabling them to nodulate and fix nitrogen in host plants (Moulin *et al.* 2001, Sy *et al.* 2001).

1.1.4 Introduction to rhizobia

The term rhizobia has been assigned to several types of soil bacteria with a common ability to form symbiotic associations with leguminous plants leading to nitrogen fixation within root nodules. Rhizobia research has led to a better understanding regarding signal exchanges between plant and microbe, genes involved in nodulation and nitrogen fixation, and competition in the rhizosphere to name a few. Rhizobia consist of bacteria from several genera including *Rhizobium, Sinorhizobium, Mesorhizobium, Agrobacterium* and *Bradyrhizobium*.

1.1.5 Genetic organization of rhizobia symbiosis genes

The genetic organization within each genus differs. In addition to a chromosome, *Rhizobium, Sinorhizobium, Agrobacterium* and *Mesorhizobium* all contain stable

plasmids that range in size from 100 to 1,700 kb, accounting for about 40% of the genome (Honeycutt *et al.* 1993). The location of symbiosis genes differ for the varying types of rhizobia; they may be clustered in the same region or dispersed, they may be located on plasmids or scattered within the chromosome and plasmids (Long 2001). For example the symbiotic genes necessary for nodulation are found either on large plasmids or within symbiotic islands in rhizobia (Baron and Zambryski 1995, Freiberg *et al.* 1997, Kaneko *et al.* 2000, Kaneko *et al.* 2002, Sullivan and Ronson 1998). It is common for a *Rhizobium* species to contain a plasmid encoding many of the symbiotic genes and some strains may also contain up to ten other >100 kb stable plasmids (Thurman 1985). These large plasmids are likely beneficial in increasing fitness to the strains within the soil environment, however the function of much of the coding regions located on these plasmids is still unknown (Baldani *et al.* 1992, Shaw 1987). Sequencing projects by several groups have identified the genome size and arrangement of strains representing members of the rhizobia.

Rhizobium leguminosarum bv. *viciae* nodulates *Vicia, Pisum, Lathyrus,* and *Lens. Vicia faba* is commonly known as fava bean, *Pisum sativum* is the garden pea, and, *Lens* species are often referred to as lentils. The genome of *R. leguminosarum* bv. *viciae* strain Rlv3841 was sequenced; it contains a 5,057,142 bp chromosome and six plasmids called pRL7-12 and contain 151,564 bp, 147,463 bp, 352,782 bp, 488,135 bp, 684,202 bp, and 870,021 bp respectively (Young *et al.* 2006).

Rhizobium etli nodulates the common bean (*Phaseolus vulgaris*). *R. etli* strain CFN42 consists of a 4,381,608 bp chromosome and six plasmids called p42a-f containing

194,229 bp, 184,338 bp, 250,948 bp, 371,255 bp, 505,334 bp, and 642,517 bp respectively (Gonzalez *et al.* 2006).

Sinorhizobium is known to nodulate the roots of *Medicago*, *Melilotus* and *Trigonella* species. *Medicago sativa* is the most prominent member and is more commonly known as alfalfa. *Sinorhizobium meliloti* strain Rm1021 consists of a 3,654,135 bp chromosome and two plasmids called pSymA and pSymB containing 1,354,226 bp and 1,683,333 bp respectively (Barnett *et al.* 2001, Capela *et al.* 2001, Finan *et al.* 2001, Galibert *et al.* 2001).

Mesorhizobium forms determinant nodules on several species of *Lotus*. A strain of *Mesorhizobium loti*, MAFF303099 has been sequenced, the genome consists of 7,036,071 bp chromosome and two plasmids called pMLa and pMLb that contain 351,911 bp and 208,315 bp respectively (Kaneko *et al.* 2000). *Mesorhizobium loti* genome sequencing has identified a 611-kb region that contains genes for nodulation and nitrogen fixation that based on codon usage analysis was predicted along with the plasmids to have originated from other genetic systems (Kaneko *et al.* 2000).

Bradyrhizobium forms nodules on soybeans (*Glycine max*). *Bradyrhizobium japonicum* strain USDA110 was sequenced and the genome consists of a single chromosome containing 9,105,828 bp (Kaneko *et al.* 2002).

Characterization of the plasmids has identified several genes necessary for symbiosis. In *S. meliloti,* the genes for nodulation and nitrogen fixation (Banfalvi *et al.* 1981, Hooykaas 1981, Johnston 1978, Rosenberg *et al.* 1981), genes involved in exopolysaccharide production necessary for symbiotic associations with indeterminate nodules (Finan 1986, Hynes 1986), and C_4 dicarboxylate transport (Finan *et al.* 1988,

Watson *et al.* 1988, Yarosh *et al.* 1989) necessary for bacteroid carbon catabolism within the nodule, are all examples of plasmid-encoded genes. Strains of *S. meliloti* and *R. leguminosarum* have been shown to contain genes encoding bacteriocin production and resistance (Hirsch 1979, Hirsch 1980, Oresnik *et al.* 1999) as well as genes encoding proteins necessary for the catabolism of several carbon compounds (Baldani *et al.* 1992, Charles and Finan 1991, Charles 1990). These genes are also located on plasmids that improve the fitness of these strains in the competitive rhizosphere environment (Martinez-Romero and Rosenblueth 1990, Toro and Olivares 1986, Triplett and Sadowsky 1992).

1.1.6 Plasmid curing experiments

In the early 1990s plasmid curing experiment were used to determine if the presence or absence of plasmids would have an affect on the fitness and competitive ability of the rhizobia in the rhizosphere versus wild-type strains. In experiments where plasmid-cured strains of *R. leguminosarum* or *R. etli* were co-inoculated with wild-type strains on host seedlings, the plasmid-cured strains in some cases, were not as competitive as wild-type strains for nodule occupancy (Brom *et al.* 1992, Hynes 1990a, Hynes 1990b).

1.1.7 R. leguminosarum strain Rlt100 (wild-type)

Rhizobium leguminosarum bv. *trifolii* strain Rlt100, was previously known as W14-2; the genome consists of the chromosome and four plasmids called pRtrW14-2a-d, they are 150, 170, 260, and 460 MDa in size as predicted by their migration position

within a modified Eckhardt (Eckhardt 1978) agarose gel in electrophoresis experiments (Baldani *et al.* 1992, Hynes and McGregor 1990). Strain Rlt100, unlike *R. leguminosarum* bv. *viciae* strain Rlv3841 has not been sequenced and is relatively poorly characterized. Rlt100 was isolated from nodules of arrowleaf clover (*Trifolium vesiculosum* Savi.) grown in east Texas (Baldani *et al.* 1992).

Rlt100 plasmid pRtrW14-2d was determined to be the symbiotic plasmid (Baldani *et al.* 1992). Baldani *et al.* tested plasmid cured derivatives of Rlt100 on a myriad of carbon sources to determine if the genes necessary for catabolism of a particular carbon source were plasmid-encoded (Baldani *et al.* 1992). They found that the pRtrW14-2a cured derivative had inhibited growth on malate and lactose as well as a loss of both motility and ability to swarm, pRtrW14-2b cured derivative affected growth on adonitol, pRtrW14-2c cured derivative had inhibited growth on L-rhamnose and sorbitol, and pRtrW14-2d cured derivative was impaired on catechol and lacked symbiotic properties (Baldani *et al.* 1992). The inability to swarm is likely due to lack of O antigen of LPS production known to be plasmid-encoded (Hynes and McGregor 1990) and was previously linked to motility (Hynes and McGregor 1990, Priefer 1989, Smit *et al.* 1989). pRtrW14-2d cured derivative strain was unable to nodulate and pRtrW14-2a cured derivative had fewer nodules versus the wild-type strain, possibly due to the lack of LPS as this phenotype was been observed in other strains lacking LPS production (Baldani *et al.* 1992, Brink *et al.* 1990, Hynes 1990a, Hynes 1990b, Hynes and McGregor 1990).

1.1.8 Characterization of plasmid-encoded L-rhamnose locus in strain Rlt100

With the observation by Baldani *et al.* that the plasmid cured derivatives of Rlt100 yielded strains that had L-rhamnose, sorbitol, adonitol, lactose, glycerol, and catechol growth phenotypes (Baldani *et al.* 1992), Oresnik *et al.* hypothesized that some of the plasmid-encoded genes may also be necessary for growth in the rhizosphere and if absent, might adversely affect nodulation competitiveness in these strains (Oresnik *et al.* 1998).

Generation of a cosmid bank of Rlt100 yielded a plasmid containing the genes necessary for growth on L-rhamnose flanked by genes involved in arabinose and sorbitol catabolism. The L-rhamnose locus was determined to be located on pRtrW14-2c (Oresnik *et al.* 1998). Transposon mutagenesis of the L-rhamnose catabolic locus identified a region of approximately 11 kb, arranged in at least two transcriptional units. Partial sequence analysis of L-rhamnose catabolic genes flanking insertional elements identified a coding sequence for a protein homologous to an *E. coli* sorbitol-6-phosphate-2-dehydrogenase (Oresnik *et al.* 1998). Sequence from another insert suggested the presence of a protein homologous to an ATP binding protein associated with transport of sugars (Oresnik *et al.* 1998).

Oresnik *et al.* observed that when strain Rlt105 containing a lesion within the putative dehydrogenase gene was co-inoculated with wild-type Rlt100 on clover *Trifolium repens* cv. Ladino seedlings, it was isolated in less than 10% of the developed nodules (Oresnik *et al.* 1998). An inoculation ratio of 10:1 (Rlt105:Rlt100) was required to achieve a 1:1 ratio of strains within nodules. Strain Rlt105 was not impaired in its ability to form nodules or fix nitrogen despite the reduced ability to compete versus the wild-type strain for nodule occupancy (Oresnik *et al.* 1998).

A strain containing a *lacZ* translational fusion within the putative dehydrogenase gene of the L-rhamnose catabolic locus was determined to be induced by both L-rhamnose and by clover root exudate, suggesting that the genes encoding the L-rhamnose catabolic locus may also be up-regulated within the rhizosphere (Oresnik *et al.* 1998).

It was proposed that while L-rhamnose catabolism does not appear to have an affect on nodulation and nitrogen fixation, it may play a role in establishing and maintaining growth within the infection thread (Oresnik *et al.* 1998). Rhamno-galacturonan is a component of dicotyledonous plant cell wall (McNeil *et al.* 1984). It has also been found in the mucilage of a number of legume plants (Knee *et al.* 2001). If it is present in a form that can be utilized by the invading strain of rhizobia, it would not be unrealistic to propose that strains able to use available nutrients may have an advantage in growth and propagation along the infection thread (Oresnik *et al.* 1998).

1.1.9 Plant root exudate

The soil environment immediately surrounding plant roots is called the rhizosphere. Legume roots exude compounds and slough cells in a leaky manner into the rhizosphere providing nutrients to soil bacteria. Some compounds found in legume plant exudate like flavones, flavanones, flavonoids and betaines have been identified as being responsible for the induction of several nodulation (*nod*) genes in the rhizobia (Cassab 1998, Cosgrove *et al.* 1997, Cosgrove *et al.* 2002, Gage 2004). These plant inducers are different for each *Rhizobium*, they are species specific and the composition and concentration of the compounds may vary with age and condition of the plant (Downie 1994, Phillips 1996, Spaink *et al.* 1998). Long and Peters determined that the alfalfa

exudate inducing compound was tetrahydroyflavone and lutoelin, whereas the soybean inducer was the isoflavone, diadzein (Long 2001). Several other plant-derived inducers of rhizobial *nod* genes have also been identified: methoxychalcone, naringenin, as well as non-flavonoid inducers such as trigonelline and stachydrine for example (Long 1996, Phillips 1994).

1.1.10 nod genes

The *Rhizobiaceae nod* genes encode approximately 25 proteins involved in the synthesis and export of Nod factor (Fisher and Long 1993, Gage 2004). Some of the *nod* genes are common to all *Rhizobium* while other *nod* genes are species-specific (Long 1996). NodD, is a transcriptional activator that acts on the *nod* box sequence where it binds the DNA upon sensing the plant derived inducer and it is involved in activation of the *nod* genes (Fisher and Long 1993).

1.1.11 Nod factor

All rhizobia produce Nod factor (Denarie *et al.* 1996, Long 1996, Schultze and Kondorosi 1998) a lipooligosaccharide with a standard chitin backbone containing a lipid attachment to the non-reducing end and host specific side groups with modified sugars, and acetyl or carbamoyl residues (Long 2001). Upon encountering Nod factor, the host plant can undergo several changes including, root hair deformation, membrane depolarization, calcium oscillations, and initiation of root cortex cell division (Gage 2004).

1.1.12 Infection thread overview

The infection thread is the method of delivery that enables rhizobia to be delivered to the root cortex for the development of the mature nitrogen fixing root nodule, this process was described eloquently in a review by Gage (Gage 2004). The process of infection thread development is initiated when bacteria come in contact with developing root hairs. Bacteria cells can contact and become trapped between two adjacent root hairs or within single root hairs that have curls (Callaham 1981, Gage 2004). Invagination of the plant call wall enables the rhizobia to enter and divide within the infection thread as the thread elongates and is directed towards the epidermal root cells (Van den bosch *et al.* 1989, van Spronsen *et al.* 1994, van Workum 1998). The rhizobia exit the infection thread and enter the developing nodule where they may differentiate into a bacteroid-like state capable of fixing nitrogen (Gage 2004).

1.1.13 Legume root hair development

To initiate root hair development the nucleus of the outermost epidermal root cell aligns itself with the inner facing cell wall (Sieberer 2000). A root hair bulge filled with cytoplasm emerges from the outermost wall and is elongated via internal turgor pressure and secretions of Golgi body derived vesicles containing membrane and cell wall components near the tip of the growing root hair (Gage 2004). Delivery of vesicles occurs via actin-dependent cytoplasmic streaming towards the tip, this delivery system accommodates rapid tip elongation (Hepler *et al.* 2001, Iwanami 1956). As the root hair extends, the nucleus follows the tip into the root hair at a distance, these are zone I root hairs (Gage 2004). In zone II root hairs a vacuole approaches the cytoplasmic filled tip,

while the nucleus no longer follows the tip (Gage 2004). Zone III root hairs have no defined location of nucleus and the vacuole occupies most of the tip (Gage 2004). Most infection threads occur in zone II root hairs, this process is discussed later in this chapter.

Actin filaments and microtubules are involved in directional root hair elongation (Carol and Dolan 2002, Hepler *et al.* 2001, Ketelaar 2001, Ridge 1992, Sieberer 2000, Vidali *et al.* 2001). Myosin motors move along actin filaments delivering vacuoles to the tip during cytoplasmic streaming (Herrmann *et al.* 1992, Staiger *et al.* 1994, Tominaga 2000, Vidali *et al.* 2001), the actin filaments are required for tip growth (Bibikova *et al.* 1999) and are organized by microtubules (Lloyd 1987, Tominaga 1997). Microtubules also maintain the structure within the cytoplasmic dense region at the tip, thereby controlling the direction without affecting tip elongation (Bibikova *et al.* 1999, Lloyd 1987, Sieberer *et al.* 2002) in addition to maintaining correct distance of the nucleus (Sieberer *et al.* 2002).

1.1.14 Rhizobia to root hair contact

Initial weak binding of rhizobia to root hair occurs by Ca^{2+}-dependent rhicadhesin proteins (Smit *et al.* 1987, Smit *et al.* 1989). Tighter binding occurs with bacterial synthesis of cellulose fibrils, anchoring the rhizobia to the root hair (Gage 2004, Smit *et al.* 1987). Host specificity is accomplished with host synthesized lectins that recognize bacteria cell wall saccharide moieties proposed to be a unique fingerprint for that symbiont (Diaz *et al.* 1995, Diaz 1986, Hirsch 1999). Host specific, rhizobia to root hair contact is believed to be necessary for increasing localized concentration of Nod factor,

that lead to tight root hair curling referred to as a shepherd's crook, and infection thread formation (Gage 2004, Hirsch 1999, van Rhijn *et al.* 2001).

1.1.15 Root Nod factor receptors

Putative *M. truncatula* receptors LYK3 and LYK4, specific for *S. meliloti* wild-type Nod factor were identified (Gage 2004, Limpens *et al.* 2003). LYK3 and LYK4 encode histidine kinases that contain an extracellular domain similar to a LysM domain identified as a chitinase (Gage 2004, Limpens *et al.* 2003). The Nod factor backbone consists of a four to five subunit chitin molecule (Gage 2004). The non-reducing end of wild-type *S. meliloti* Nod factor interacts with LYK3 or LYK4, this interaction is critical for initiating and development of infection threads (Gage 2004, Limpens *et al.* 2003).

1.1.16 Nod factor induced root hair curling

Nod factor produced in rhizobia induces root hair curling for zone II root hairs likely due to the established yet plastic physiological state of zone II root hairs (de Ruijter 1998, Heidstra *et al.* 1994, Sieberer 2000). The root hair curling of the zone II root hairs upon induction by Nod factor is initiated by swelling of the root hair tip, and re-establishment of a zone I-like state whereby the root cells are once again polarized and increased delivery of vesicles towards the tip occurs (de Ruijter 1998, Heidstra *et al.* 1994, Miller 1999, Sieberer 2000).

The point source of Nod factor from rhizobia contacting root hairs might lead to tight root hair curls (Esseling *et al.* 2003, Gage 2004, Ridge 1992, van Brussel 1992). Addition of Nod factor likely plays a role in actin filament organization or reorganization

however the exact role is unclear (Cardenas *et al.* 1998, de Ruijter 1999). Root hair microtubule arrangements are also disrupted after addition of compatible Nod factor (Gage 2004). Microtubules are rearranged to align along the long axis of the root hair, they become positioned between the root hair nucleus and the tip of the root hair, with initiation of infection thread the microtubule array bridged the infection thread tip and root hair nucleus (Timmers *et al.* 1999). It has been hypothesized that the microtubules may be involved in infection thread development similar to their involvement in root hair extension (Bibikova *et al.* 1999, Olsson *et al.* 2002, Sieberer *et al.* 2002). There is also a line of actively streaming cytoplasm that occurs between the infection thread tip and the root hair nucleus in alfalfa (Gage 2004).

1.1.17 Infection thread formation and elongation

Relatively little is known about the mechanisms in which the infection thread is elongated and the details of rhizobia growth within these infection threads. Infection thread formation occurs at the point of bacteria and cell wall contact (Callaham 1981, Napoli and Hubbell 1975, Ridge and Rolfe 1985, Turgeon 1985). *S. meliloti* and *R. leguminosarum* both form pits on the roots of their host plants (Gage 2004, Mateos *et al.* 2001). The enzymes responsible for pit formation are bound to the surface of the rhizobia cell (Gage 2004).

Preinfection threads (PITs) occur when cytoplasm within the cells of the outer cortex aligns in a central position forming cytoplasmic bridges while the nucleus aligns against the innermost cell wall, reminiscent of root hair elongation (Gage 2004, van Brussel 1992). The PITs are polarized and contain bundles of endoplasmic microtubules

(Gage 2004, Timmers *et al.* 1999, van Brussel 1992). As the infection thread fuses with the outer cortex PITs the bacteria enter the intracellular space and the weakened cell wall invaginates and the infection thread continues to grow (Gage 2004, Timmers *et al.* 1999, van Brussel 1992). Infection threads are continuous with the plant cell wall and consist of esterified and unesterified pectins, xyloglucans and cellulose, in addition to other cell wall compounds like rhamnogalacturonan (McNeil *et al.* 1984, Rae 1992).

Using strains of *S. meliloti* tagged with fluorescent protein, Gage *et al.* observed that bacteria located at the tip of the infection thread were actively growing while those at the oldest point of the infection thread were not (Gage 2002, Gage *et al.* 1996). The growth rate of the rhizobia within the infection thread matches the rate of elongation of the infection thread (Gage 2004).

1.1.18 Initiation of indeterminate nodules

As the root hair curling, infection thread initiation, infection thread elongation, and outer cortical cells are stimulated to re-enter the cell cycle (Dudley 1987, Libbenga 1973, Napoli and Hubbell 1975, Timmers *et al.* 1998, Timmers *et al.* 1999, van Brussel 1992), a second wave of induced physiological changes begin. These include activation of pericycle cells next to protoxylem poles within 12 hours of inoculation with compatible rhizobia (Timmers *et al.* 1999), followed by activation of the cells of the inner cortex (together referred to as the nodule primordium) (Gage 2004). The nodule primordium is the intracellular target of the infecting rhizobia (Gage 2004). It appears that as the infection thread enters the nodule primordium the meristem is formed from the cells of the middle cortex (Timmers *et al.* 1999). With activation of the middle cortex

cells, the indeterminate nodule meristem is then established and continues to grow outward (Gage 2004). The infection thread must follow the extending nodule to infect the area behind the meristem called the infection zone (Vasse *et al.* 1990). In alfalfa, *S. meliloti* infection threads are directed towards the meristem using polarized cells (Gage 2004). The rhizobia are able to differentiate and a zone of fixation is established behind the infection zone and biological nitrogen fixation can occur.

1.1.19 Control of infection threads

There are few nodules formed relative to the number of infection threads initiated post inoculation (Fahraeus 1957, Nutman 1962, Purchase 1958, Vasse 1993). The nodules are responsible for decreased infection thread rates (Bhuvaneswari *et al.* 1981, Caetano-Anolles and Gresshoff 1991a, Caetano-Anolles and Gresshoff 1991b, Caetano-Anolles 1991, Kosslak and Bohlool 1984, Nutman 1952) and addition of an ammonia or nitrate source has a similar inhibitory effect (Caetano-Anolles and Gresshoff 1991b, Gage 2004, Heidstra *et al.* 1997, Streeter 1988). The plant controls infection rates and how many infection threads are destined for nodules (Gage 2004). Plant directed infection thread termination occurs as plant cells inhibit the bacteria from exiting the host cells; often there are increased concentrations of plant defense chemicals used to inhibit viral, bacterial and fungal pathogens (Vasse 1993). This response does not occur with successful infection threads (Gage 2004).

Ethylene is a hormone produced by leguminous plants and it has a part in controlling the number of nodules generated (Lee and Larue 1992, Peters and Crist-Estes 1989). Ethylene inhibits the number of infections, elongation of the infections, how many

infections result in nodules, protophloem cortical cell activation (Oldroyd *et al.* 2001, Penmetsa and Cook 1997, Penmetsa *et al.* 2003). Ethylene inhibits early plant responses to rhizobia, by interrupting Nod factors affect to induce calcium spiking thereby lessening the calcium spiking within the root hairs in addition to decreasing the number of effective progressing infections threads (Gage 2004, Oldroyd *et al.* 2001).

1.1.20 Rhizobium leguminosarum L-rhamnose catabolism perspective

Experimentation of many aspects of the plant host and the bacteria symbiont have led to better understanding of these symbiotic associations and chemical exchanges leading to nitrogen fixation. A large amount of information has been obtained to further our understanding of how the plant and bacteria communicate and respond the each other in the nodulation and nitrogen fixation process. Sequencing projects have aided in phylogenetic characterization and have served to identify symbiotic plasmids, islands and common genetic elements that identify *Rhizobium* members. However, there are still many aspects of the symbiotic association and fitness of strains in the rhizosphere that remain less clear, for example aspects of infection thread elongation, the growth of rhizobia within the infection thread and competition within the rhizosphere.

In characterizing the rhamnose locus in *Rhizobium leguminosarum* we wish to gain insight regarding the ability of the wild-type strain to use any available nutrient source within the rhizosphere or within the infection thread may lead to a competitive advantage in the ability to nodulate it host plant. This characterization of the rhamnose locus in strain Rlt100 identified an ABC transporter in addition to other catabolic genes responsible for L-rhamnose catabolism.

1.2 ABC transporters

1.2.1 Introduction to ABC transports

Movement of molecules across selectively permeable cell membranes occurs by several different mechanisms: simple passive diffusion of molecules directly through the lipid bilayer, by protein assisted passive diffusion, through non-gated pores, ions through gated channels, or facilitated diffusion. In order to move molecules across a concentration gradient, energy is required. Energy driven pumps, endocytosis and exocytosis, carrier-assisted transport, symport, and antiport are examples of active transport.

ABC (ATP binding cassette) transporters (Hyde *et al.* 1990) are among the most studied energy dependent methods of active transport (Saier 2000). They have been conserved throughout evolution and are found in archaea, bacteria and eukarya (Ames *et al.* 1990, Higgins 1992). A classification system based on phylogenetic analysis and functional properties of sequence related transporters in all three kingdoms has yielded more than 1000 members with known sequences (Saier 1999). ABC transport proteins make up one of the largest protein families, In *E. coli* there are 67 ABC transport proteins constituting about 5% of the entire genome (Blattner *et al.* 1997, Linton and Higgins 1998, Mauchline *et al.* 2006), whereas around 50 human transport proteins have been classified into 7 subfamilies, representing one of the four major gene families in humans (Altenberg 2003, Higgins *et al.* 1986, Tatusov *et al.* 1997). They have been classified as importers, exporters and as having non-transport functions, using the free energy derived from binding and hydrolysis of ATP (Berger and Heppel 1974, Dassa 2003, Higgins 1992). It is believed that ABC transporter proteins may be found in all cells of all species (Jones and George 2004).

1.2.2 ABC transporter proteins have diverse functions

In humans, as least 15 ABC genes have been associated with a disease phenotype (Dean and Allikmets 2001). Mutations in ABC genes can lead to diseases like cystic fibrosis (gut and lung) (Hanrahan *et al.* 2003, Riordan *et al.* 1989), Tangiers disease (cardiovascular) (Rust *et al.* 1999), Dublin-Johnson syndrome (liver) (Konig *et al.* 1999), Stargardt disease (sight disorder) (Allikmets *et al.* 1997), amongst others.

Mammalian ABC proteins are involved in maintenance of the blood-brain barrier, transport of hormones, antigen processing, bacterial immunity, parasite resistance, detoxification, cholesterol transport, lipid transport and in developmental stem cell biology (Ames *et al.* 1990, Bunting 2002, Fath 1993, Higgins 1992, Martinoia *et al.* 2002, Young and Holland 1999). ABC transporters have also been identified in resistance of plasmodium to antimalarial compounds (Ouellette *et al.* 2001), herbicide resistance in plants (Rea 1998), multidrug resistance fungi, yeasts, and parasites, bacteria and mammals (Borst and Elferink 2002, Gottesman 2002, Tsuruo *et al.* 1981, van Veen and Konings 1997). Over-expression of ABC components in cancer patients, can confer resistance of tumor cells to chemotherapeutic agents (Bates 2003, Borst and Elferink 2002, Gottesman 2002).

In microorganisms, ABC transporters are involved in many cellular activities such as transport of small inorganic ions, sugars, amino acids, vitamins, large polypeptides, complex polysaccharides and lipids (Ames *et al.* 1990, Dassa and Bouige 2001). They are involved in conferring resistance to antibiotics (Poelarends 2002) and antifungal agents (Bauer *et al.* 1999), in maintenance and repair of DNA and gene regulation (Aravind *et*

al. 1999), osmotic homeostasis, cell division, and pathogenesis (Jones and George 2004) to list a few.

1.2.3 Organization of ABC complexes

Typical ABC transporters consist of four core peptides, two transmembrane domains (TMD) and two nucleotide binding domains (NBD) also referred to as ABC proteins or ABC ATPase proteins (Higgins *et al.* 1986). These core components can be organized as separate polypeptides or as multidomain polypeptides, functioning as homo- or heterodimers spanning the cytoplasmic membrane (Linton and Higgins 1998). The term ABC protein, used in the text refers to the protein responsible for hydrolysis of ATP and should not be confused with ABC transporter proteins, referring to all the proteins involved in the ABC transporter complex.

Most eukaryotic ABC transporters function as exporters of hydrophobic molecules from the cytoplasm (Dean and Allikmets 2001). Eukaryotic and bacterial ABC exporters are usually organized as fused multidomain polypeptides (Dassa and Bouige 2001). It has been hypothesized that endocytosis has replaced ABC transporters as the primary transport mechanism in eukaryotic organisms and this is reflected in their overall reduction of ABC transport proteins (Higgins and Linton 2003).

In prokaryotes, ABC transporters involved in import usually have separate polypeptides for each transporter component (Saier 2000). Prokaryote and archaea importers have additional auxiliary proteins called a periplasmic binding proteins (PBP) that confer directionality, substrate specificity, and high affinity for substrate (Wilkinson and Verschueren 2003).

1.2.4 Brief overview of ABC transport

In gram-negative bacterial ABC importers, substrate enters a cell through the outer membrane via porins, a substrate specific periplasmic binding protein binds the substrate and presents the bound substrate to a periplasmically exposed TMD (Nikaido and Hall 1998). Interaction of the PBP with bound substrate to the TMD stimulates conformational changes through the TMD to the ABC proteins located in the cytoplasm, associated with the TMD. The process of translocation of substrate from the periplasm to the cytoplasm is intricate, it involves dimer formation of the ABC proteins triggering ATP hydrolysis on the cytoplasmic side (Figure 1.1) (Davidson *et al.* 1992). The known details of the translocation process will be discussed in detail later in this chapter.

1.2.5 ABC transporter components

1.2.5.1 Outer membrane

The outer membrane protects gram-negative bacteria from harsh environments, it is asymmetric and consists of an inner facing leaflet containing phospholipids similar to the cytoplasmic membrane and the outer leaflet usually consisting of lipopolysaccharides (LPS) (Figure 1.1) (Koebnik *et al.* 2000, Pugsley 1993). Outer membrane proteins organized within the outer membrane as pores called porins, that enable selective permeability of solutes (Koebnik *et al.* 2000, Pugsley 1993). Due to the hydrophobic nature of the membrane the proteins making up the porins are hydrogen bonded β-pleated sheets organized as closed barrels (Koebnik *et al.* 2000).

Several families of outer membrane proteins have been identified in *E. coli* including OmpA, OmpX, phospholipase A, general porins (OmpF, PhoE), substrate

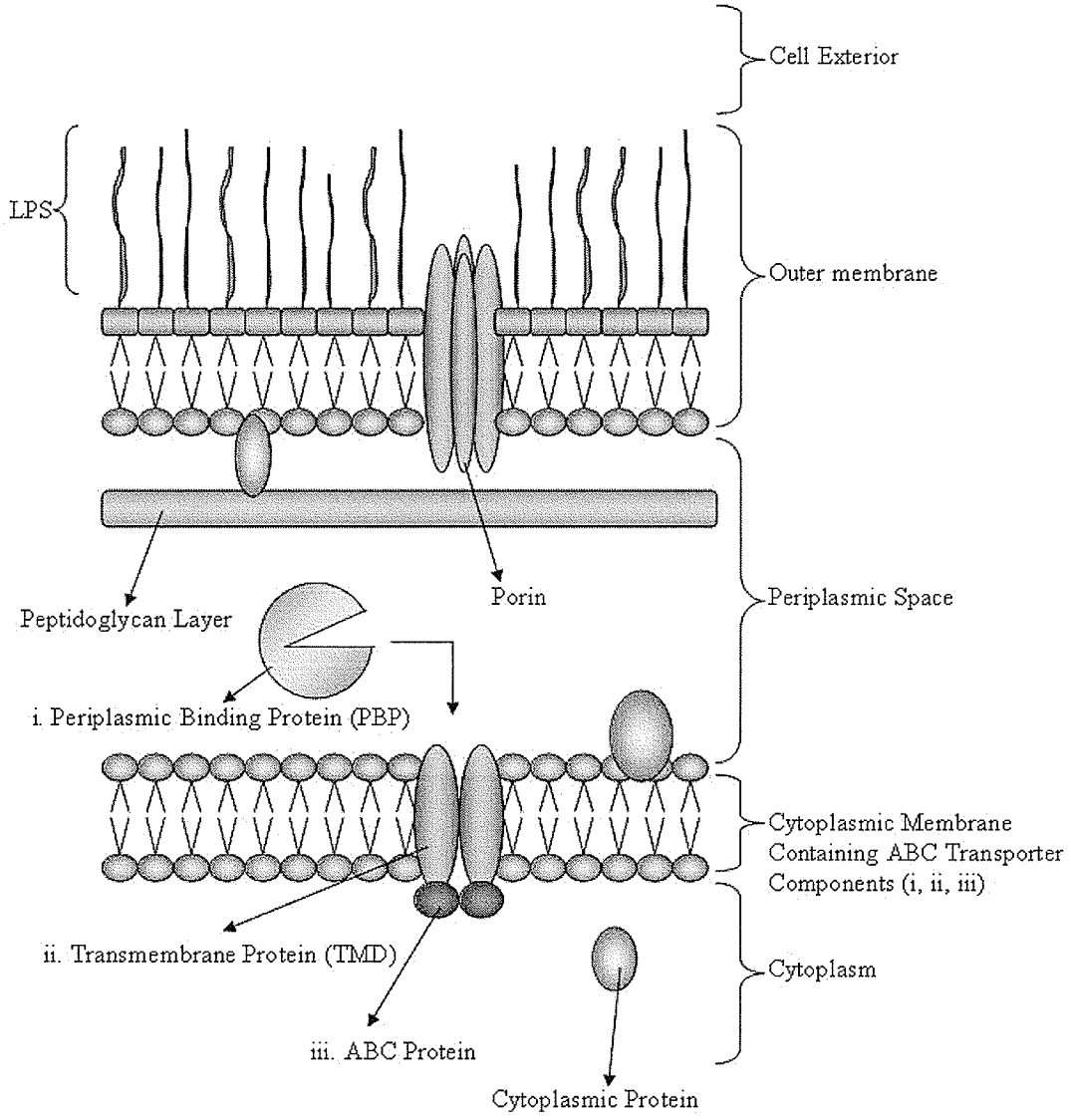


Figure 1.1. Cross-section overview of the structure of a gram-negative bacterial cell

specific (LamB, ScrY) and TonB-dependent iron siderophore transporters (FhuA, FepA). They are involved in the physical linkage of outer membrane to peptidoglycan, neutralizing host defense mechanisms, hydrolysis of phospholipids, diffusion of ions and other small molecules, maltose and maltodextrin uptake, and uptake of iron-siderophore complexes, respectively as reviewed by Koebnik (Koebnik *et al.* 2000). Some of these outer membrane proteins can be strongly induced when they are required, for example maltoporin (LamB) (Boos and Shuman 1998), can have 40, 000 copies per cell, upon induction with maltose or maltodextrins (Luckey and Nikaido 1980a, Luckey and Nikaido 1980b).

The method of L-rhamnose entry through the outer membrane of *R. leguminosarum* has yet to be elucidated. Perhaps the best clue for porin-based outer membrane L-rhamnose translocation in *R. leguminosarum* comes from research on the *Brucella-Rhizobium* porin (BRP) family. Two homologous outer membrane proteins, Omp2a and Omp2b have been proposed in cell permeability (Paquet *et al.* 2001). Omp2a has been proposed to have a role in sugar permeability, via pore formation, in liposome swelling assays in *Brucella* (Paquet *et al.* 2001). Another member of the BRP family is OmpIIIA of *R. leguminosarum* biovar *viciae* that is described as an 36-kDa outer membrane protein that may act as a pore (de Maagd *et al.* 1992). Interestingly, this protein was also identified as an outer membrane protein antigen group III, that exhibits repressed expression in the bacteroid state during symbiosis (de Maagd *et al.* 1989, de Maagd *et al.* 1992) where the primary source of nutrients provided by the plant is C_4 dicarboxylic acids and the transport for other sugars is presumably decreased (Finan *et al.* 1988, Watson *et al.* 1988, Yarosh *et al.* 1989).

1.2.5.2 Periplasmic binding protein (PBP)

The role of the periplasmic binding protein (PBP) in ABC importers, is to confer substrate specificity by recognition of substrate in the periplasmic space and delivery of the substrate to an associated TMD for transport (Locher 2004, Quiocho and Ledvina 1996). Gram-negative, gram-positive bacteria and archaea contain PBPs; however, eukaryotic ABC transport systems function as exporters, and do not contain known PBP analogues.

PBP of gram-positive bacteria and archaea are lipoproteins that must be anchored to the cytoplasmic membrane surface because of the physiological difference of the membrane structure as compared to the outer and cytoplasmic membranes of gram-negative bacteria. This can be accomplished by fatty acids covalently coupled to an N-terminal cysteine residue (Gilson *et al.* 1988, Horlacher *et al.* 1998, Sutcliffe and Russell 1995), or in some archaea via a carboxy-terminal transmembrane segment (Elferink *et al.* 2001). Recently, a fused PBP-TMD-ABC protein that may function as a single unit per transporter (Borths *et al.* 2002) was predicted to be fairly widespread based on genome sequences of several gram-positive bacteria ABC-like operons (van der Heide and Poolman 2000). Both bacterial and eukaryotic exporters have TMDs and ABC proteins fused as single polypeptides (Locher 2004).

PBPs have a high affinity towards their respective substrates, enabling cells to accumulate substrates at high levels against concentration gradients (Higgins 1992, Miller *et al.* 1983). In induced gram-negative bacteria the concentration of PBP can be in the mM range (Ames *et al.* 1996). PBPs exhibit similar overall organization despite divergence in sequence; this is likely attributed to the diversity in the types of substrates

that PBPs are able to recognize. PBPs consist of two domains with similar topology. These contain a central β-sheet flanked by α-helices connected by segments of the polypeptide that acts as a hinge (Bjorkman and Mowbray 1998). PBPs have a ligand-binding, protein fold allowing the two domains of some PBPs to engulf the substrate upon binding, accounting for the proteins high affinity and selectivity (Borths *et al.* 2002, Karpowich *et al.* 2003, Newcomer *et al.* 1981, Shilton *et al.* 1996). There are also examples of PBP like BtuF, a vitamin B_{12} importer, where the substrate specific sites within the hinge region contact the substrate; however the size of the substrate prevents the substrate from being engulfed completely (Locher 2004, Locher *et al.* 2002).

Molecules of crystallized maltose binding protein (MBP) not bound to substrate, were reported to be at an equilibrium between the open and closed forms, in solution (Bjorkman and Mowbray 1998, Flocco and Mowbray 1994, Sack *et al.* 1989, Sharff *et al.* 1993, Wolf *et al.* 1994). The substrate first interacts with the ligand-binding surface of an open PBP subunit, then contacts the opposite subunit thus shifting the equilibrium to the closed form (Quiocho *et al.* 1997). The selectivity of the PBP is based on interactions between protonation states (Luecke and Quiocho 1990, Pflugrath and Quiocho 1985), size and/or steric hindrance (Hu *et al.* 1997) of similar substrates. In contrast to high level of specificity there are PBP that can accommodate similar substrates. For example the lysine-arginine-ornithine-binding-protein (LAOBP), that as the name implies, can bind substrates with similar structures (Nikaido and Ames 1992, Oh *et al.* 1994).

While the binding of several PBPs to substrates has been characterized with crystal structure data, the interaction of PBP to TMD is not clear. To increase transport efficiency it is necessary that the TMDs are able to distinguish between substrate-bound

and unbound forms of PBP. Several independent studies using PBP mutants that have not affected PBP-substrate binding, but whose transport is adversely affected can be targeted to one portion of the PBP opposite the hinge region, close to the substrate-binding pocket, and span the two subunits only when the PBP is in the closed orientation (Binnie *et al.* 1992, Bjorkman and Mowbray 1998, Eym *et al.* 1996, Hor and Shuman 1993, Oh *et al.* 1993, Prossnitz *et al.* 1988). Each closed PBP subunit is believed to interact with a different TMD subunit (Hor and Shuman 1993, Zhang *et al.* 1992).

1.2.5.3 Membrane spanning proteins

Transmembrane proteins or domains (TMD) have proven to be more difficult to over-express and purify in active forms than soluble proteins due to their amphipathic nature (Dahl *et al.* 2004, Ostermeier and Michel 1997). The complications arise because over-expression of membrane proteins often leads to toxicity within the cell and misfolding of the over-expressed protein within the membrane disrupts their role in providing directional transport (Linton 2003). Generating TMD crystallization data is also difficult due to flexibility within the bilayer that is not conducive to high resolution X-ray imaging (Linton 2003). Most of the early topological analysis was accomplished with a combination of hydropathy plot computer analysis and with the use of translational fusions (Linton 2003).

In general, it is the overall architecture that is conserved. Phylogenetic analysis of TMDs has suggested that families of similar topology may have evolved independently of each other (Busch and Saier 2002, Dahl *et al.* 2004, Saier 1999, Saier 2000). TMDs do not exhibit significant sequence similarity unless they are transporting chemically related

substrates (Saurin *et al.* 1999). TMDs span the membrane by predicted α-helices 6 times per domain, for a total of 12 per transporter for exporters. For ABC importers the number of membrane spanning domains varies from 6 to 20 (Locher 2004), an example of these variations can be observed with the crystallization of BtuCD which has ten membrane spanning α-helices per TMD (Locher *et al.* 2002). It is believed that 8 transmembrane segments, 4 per monomer, might be the minimum to generate a pore large enough to allow substrate through the cytoplasmic membrane (Young and Holland 1999).

The topological structure of most importer TMDs appears to exhibit an aqueous chamber present on the periplasmically exposed side generated by the arrangement of the TMDs (Locher and Borths 2004). The pathway through the transport channel is not continuous as residues within loops of the TMDs impede the access through the channel (Locher 2004).

While not much is known about the conformational changes that occur within the TMDs during the transport process, a conserved motif called the EAA loop (**EAA**-X_3-**G**-X_9-**I**-**X**-**LP**) where X is an arbitrary residue (Kabsch and Holmes 1995, Saraste *et al.* 1990) is conserved in many bacterial ABC importers (Cotten *et al.* 1996, Currier *et al.* 1992, Dassa and Hofnung 1985a, Dassa and Hofnung 1985b, Holland 2003, Locher 2004, Mourez *et al.* 1997). This motif is located within a cytoplasmic loop of the TMD, it is believed to be a contact point with a helical domain of the ABC protein and is required for the signal pathway between the PBPs and the ABC protein for bacterial importers (Locher *et al.* 2002, Mourez *et al.* 1997). It has been suggested that EAA-loop of certain ABC importers (Mourez *et al.* 1998), the fourth intracellular loop of CFTR (Cotten *et al.*

1996), and the first cytoplasmic loop of ABC drug exporters (Currier *et al.* 1992) might be the only structurally conserved motif of TMDs in ABC transporters (Linton 2003).

RhaP and RhaQ are putative membrane spanning proteins encoded by the L-rhamnose locus in *R. leguminosarum*. Based on amino acid residue alignments designed to detect conserved motifs both membrane spanning proteins appear to contain a version of the TMD EAA loop motif (**EAA**-X_3-**G**-X_9-**I**-X-**LP**). For RhaP, this appears to be **TAA**-(VYA)-**G**-(IDTGWTKF)-**L**-(AFV)-**L**, and for RhaQ the residues are **FAA**-(RFS)-**G**-(IPVERVKS)-**I**-(LFL)-**L**, they are predicted to contain 9 and 8 putative membrane spanning segments respectively, with the RhaQ motif is clearly predicted to be located in a cytoplasmic loop based on membrane topology predictions.

1.2.5.4 ABC protein

Most of the original phylogenetic analyses of ABC transporter components were done with ABC proteins. ABC proteins contain conserved domains common to all ABC transporters. The nucleotide binding domain (NBD) is a conserved tertiary fold necessary for ATP binding and hydrolysis (Schneider and Hunke 1998). It has enabled researchers to classify and identify many putative ABC transporters. Akin to ABC transporter complexes, ABC proteins have been classified as being involved in export, as orphans with no membrane association, or involved in import (Aravind *et al.* 1999, Holland and Blight 1999).

The conserved NBD motifs were also identified in ATP synthase, myosin, adenylate kinase proteins (Walker *et al.* 1982), in receptor signaling, phosphoryl transfer reactions, motility, proton efflux, membrane transport, and DNA

translation/maintenance/repair (Jones and George 2004). NBDs were shown to bind and hydrolyze ATP (Higgins *et al.* 1985, Hobson *et al.* 1984), and were necessary for the transport of substrate (Bishop *et al.* 1989, Mimmack *et al.* 1989). High resolution structures of several NBD have been accomplished: *E. coli* RbsA ribose uptake (Armstrong 1998), *S. typhimurium* HisP histidine uptake (Hung *et al.* 1998), *T. litoralis* MalK maltose uptake (Diederichs *et al.* 2000), *M. jannaschii* MJ0796 an unknown function (Yuan *et al.* 2001), *M. jannaschii* MJ1267 amino acid transport (Karpowich *et al.* 2001), *H. sapiens* TAP1 antigen presentation (Gaudet and Wiley 2001), *P. furiosus* Rad50 double stranded DNA repair (Hopfner *et al.* 2000), *T. maritima* SMC double stranded DNA repair (Lowe *et al.* 2001). They have all contributed significantly to the characterization of the NBDs.

ABC proteins are hydrophilic cytoplasmic proteins (Figure 1.2A). Each ABC protein has three subdomains, arm I β-sheets, arm II α-helices, and the area joining arm I and arm II (Karpowich *et al.* 2001). ATP binding and hydrolysis only occurs with ABC protein dimers (Locher 2004). X-ray crystallography analysis suggests the ABC protein monomers are positioned in a head-to-tail arrangement upon dimerization (Diederichs *et al.* 2000, Hung *et al.* 1998, Smith *et al.* 2002). Conserved NBD motifs called Walker A and Walker B of one monomer are in close physical contact upon dimerization with the signature sequence of the opposite monomer generating a domain to domain interaction with bound ATP (Figure 1.2). This was observed for Rad50 (Hopfner *et al.* 2000), MutS (Lamers *et al.* 2000, Obmolova *et al.* 2000), BtuCD (Locher *et al.* 2002), MJ0796 variant (Smith *et al.* 2002), and corroborating sequence and biochemical data was also generated

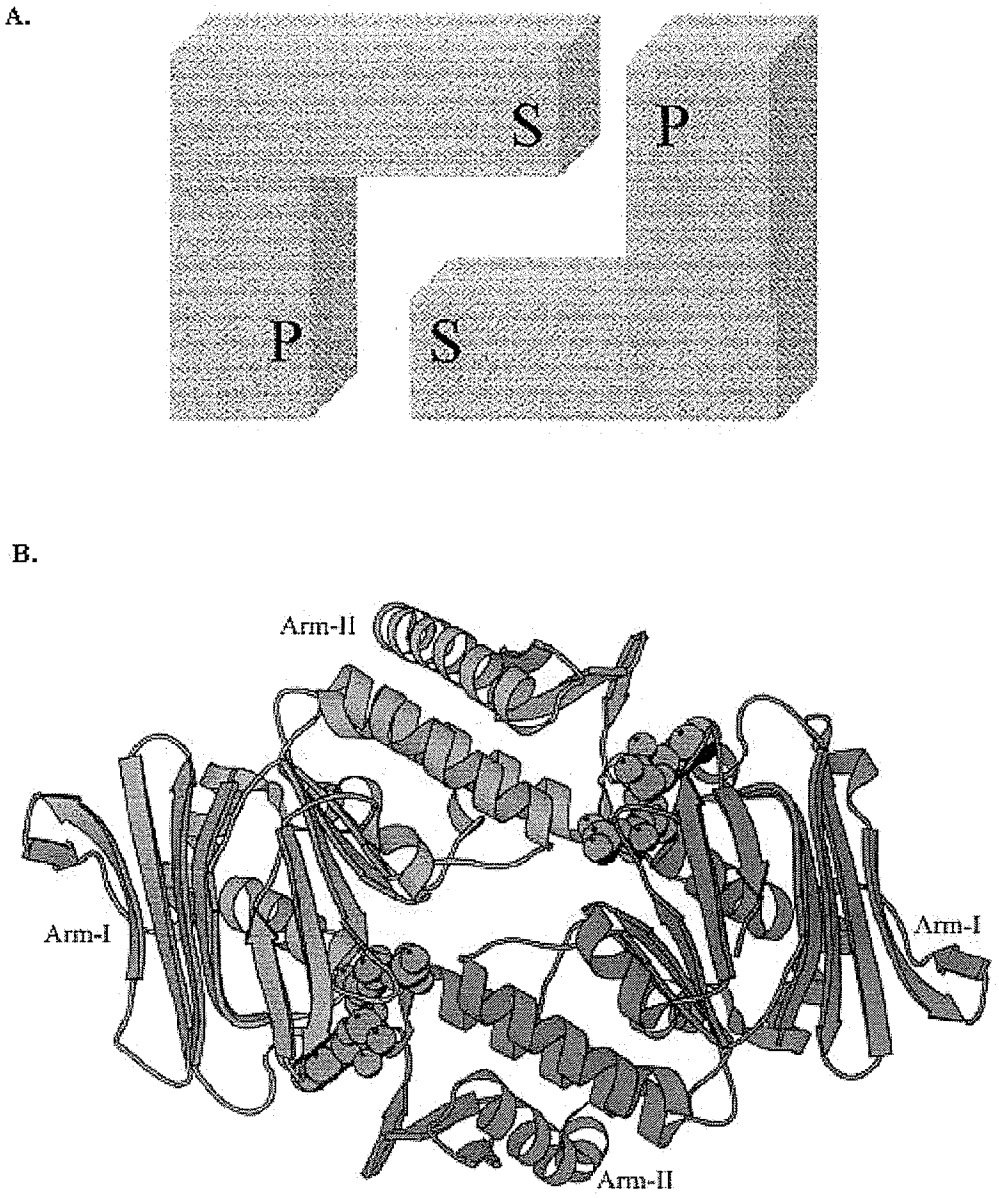


Figure 1.2. Model of ATPase Rad50 dimer association. **A**. Typical head-to-tail arrangement of ABC-type protein dimers. 'P' refers to the phosphate binding loop located within arm I, while 'S' refers to the signature sequence motif within arm II. **B**. The arms of Rad50 are indicated and displayed with a ribbon format. Two ATP molecules are indicated in green, sandwiched between two monomers colored in orange and blue, forming a dimer (Kerr 2002). Reprinted from Biochimica et biophysica acta, Volume 1561, Kerr, I. D., Structure and association of ATP-binding cassette transporter nucleotide-binding domains, pp. 47-64, copyright 2002, with permission from Elsevier.

to support this arrangement (Fetsch 2002, Hopfner *et al.* 2000, Jones and George 1999, Junop *et al.* 2001, Loo *et al.* 2002, Qu and Sharom 2001).

ABC proteins contain a central core of β-sheets surrounded by α-helices (Smith and Rayment 1996). This α/ β motif is similar to motifs found in F1-ATP synthase proteins, RecA, UvrB, helicase, Rad50 DNA repair and maintenance proteins, MutS, and SMC (Jones and George 2004) shown to be involved in the binding of ATP. Arm I, also called β-subdomain consists of antiparallel β-sheets (Figure 1.3) (Karpowich *et al.* 2001). Walker A and Walker B motifs are located within the central region of the NBD (Walker *et al.* 1982). They are separated by distance within the sequence but are in close proximity to each other in the tertiary structure.

Walker A, also called the P-loop (pyrophosphate loop), consists of a glycine rich loop (**G**-X_4-**G**-X-**G**-**K**-**S**-**T**), it is involved in electrostatic interactions of the β- and γ-phosphate of the bound nucleotide phosphate (ATP), the lysine residue especially, appears to be critical in the binding of ATP (Jones and George 2004). The geometry of the ATP molecule is held by hydrogen bonding between the nitrogen atoms in the main chain of the peptide and the phosphates of the ATP (Figure 1.4) (Higgins and Linton 2004, Jones and George 2004).

Walker B motif is a very hydrophobic β-sheet ending with an acidic residue, usually aspartate, that interacts with Mg^{2+} ion associated with the ATP molecule, also helping to establish and maintain the geometry of the active site and it is believed to contribute to the catalytic base in ATP hydrolysis (Higgins and Linton 2004, Locher 2004). The ABC β-subdomain moves the ribose and adenine of ATP away from the ATP binding core enabling the second NBD monomer to interact via the signature sequence

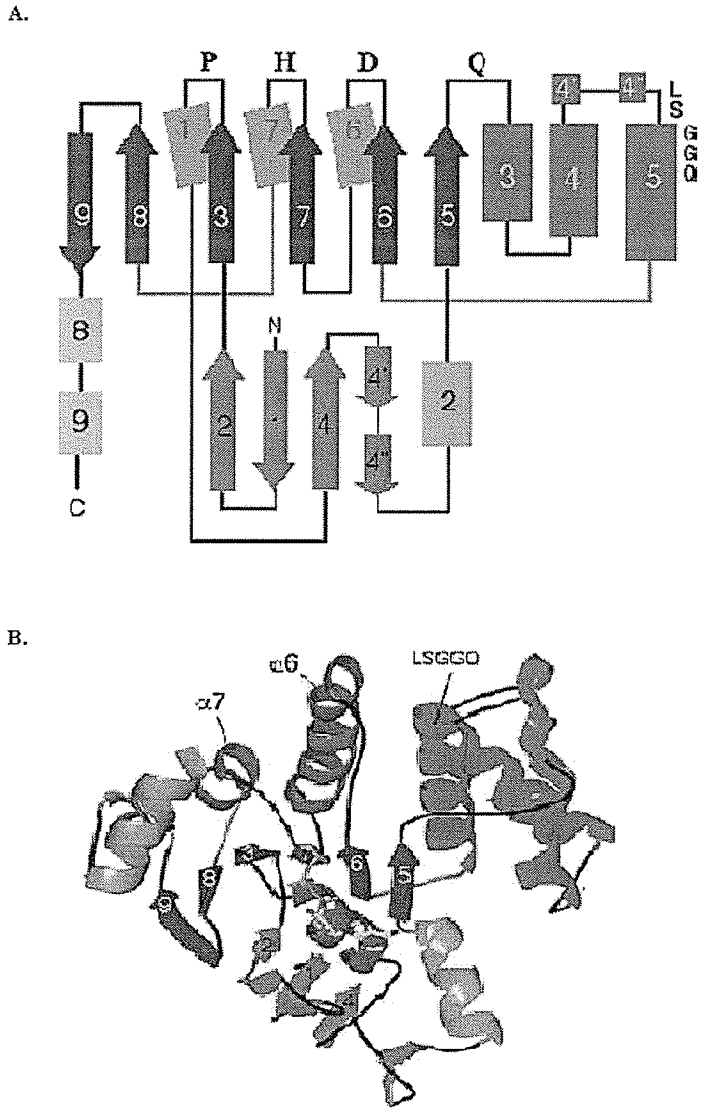


Figure 1.3. Consensus of the NBD fold and 3-D structure of ABC transporter ABC proteins NBD monomer. **A**. β-stands and α-helices are shown as arrows and rectangles, respectively. The dark blue and the light blue represent the β-stands and α-helical portions of the NBD core, respectively, while the antiparallel β-subdomain is green and the α-subdomain is red. The P-, H-, D-, and Q-loops are labeled accordingly based on the loop name. **B**. HisP (Hung *et al.* 1998) ribbon diagram with ATP inserted, the colors are as per (A) (Jones and George 2004). Reprinted from Cellular and Molecular Life Sciences (CMLS), volume **61**(6), Jones, P. M. and George, A. M., The ABC transporter structure and mechanism: perspectives on recent research, pp. 682-699, copyright 2004, with permission from Birkhäuser Publishing Ltd.

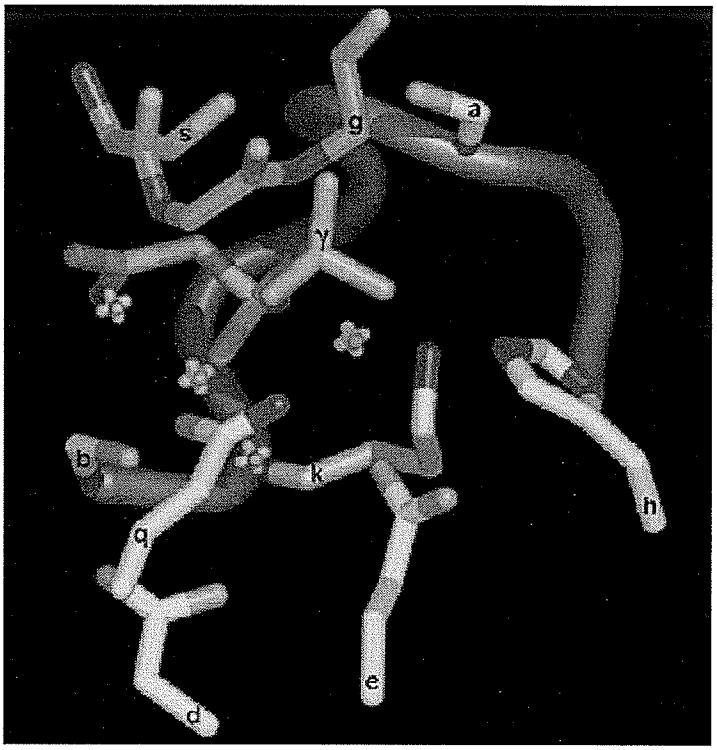


Figure 1.4. ABC ATPase catalytic core. Conserved residues within catalytic core of MJ0796 (E171Q variant) (Smith *et al.* 2002) ABC ATPase protein. Colors are as follows: P-loop of one monomer (grey tube), oxygen atoms (red), nitrogen (blue), phosphate (orange), residues belong to the P-loop (yellow), signature sequence of opposite monomer (green), catalytic Mg^{2+} ion (cyan), water molecules (light red), and nucleophillic water in center (dark red). Letters are as follows: Walker A (a, b, k = serine, serine, lysine), Walker B (d = aspartate), β-strand 5, 6, 7 (q, e, h = glutamine, glutamate, histidine), C-motif (g, s = glycine, serine) (Jones and George 2004). Reprinted from Cellular and Molecular Life Sciences (CMLS), volume **61**(6), Jones, P. M. and George, A. M., The ABC transporter structure and mechanism: perspectives on recent research, pp. 682-699, copyright 2004, with permission from Birkhäuser Publishing Ltd.

with the semi-exposed partially bound ATP molecule (Karpowich *et al.* 2001, Kerr 2002).

Arm II, also called the α-subdomain, is made up of a more structurally diverse α-helical domain (Ames *et al.* 1992) containing the ABC signature motif **L-S-G-G-Q** (Bianchet *et al.* 1997), also called the C-motif, which also participates in ATP association (Hopfner *et al.* 2000, Locher *et al.* 2002, Smith *et al.* 2002). In the absence of ATP the physical distance between Walker A motif and the signature sequence is increased in the active site and the shared interface between the subunits is limited (Hopfner *et al.* 2000, Locher *et al.* 2002), explaining why physiological dimers could only be observed in the presence of ATP (Locher 2004).

The conserved Q-, D-, and H-loops are named after their conserved residues; they are located in the hinge region between the two arms of the ABC protein. These active site residues are involved in contacting the γ-phosphate of the bound ATP and they are believed to be involved in dimerization and conformational changes during the catalytic cycle (Jones and George 2004). While these motifs are predicted to constitute important switch regions, their roles are not completely understood (Jones and George 2004).

The Q-loop, also called γ-phosphate linker (Yuan *et al.* 2001) was determined through phylogenetic analysis to contain a conserved glutamine followed by a flexible loop (Hopfner *et al.* 2000). The glutamine residue within the Q-loop is projected into the active site during the ATP hydrolysis catalytic cycle where it engages the Mg^{2+} ion, followed by its retraction after ATP hydrolysis (Jones and George 2002, Yuan *et al.* 2001). The Q-loop was also proposed to be involved in γ-phosphate sensing within the NBD, whereby conformation changes occur upon ATP binding and hydrolysis leading to

interface signaling between the TMD and the NBDs active site (Hunke *et al.* 2000a, Hunke *et al.* 2000b, Jones and George 2002, Jones and George 2004, Urbatsch *et al.* 2000).

The flexible D-loop has been identified as being critical for catalytic activity and communication within the active site. It has been proposed to affect the orientation of the catalytic base to enable or prevent ATP hydrolysis (Locher *et al.* 2002).

The H-loop participates in the interaction of Mg^{2+} ion of ATP, by presentation of a water molecule involved in the hydrolysis process (Altenberg 2003). The signature sequence H-loop, Q-loop and stacking aromatics coordinate bound nucleotide or nucleophillic water, but the exact role is unknown (Higgins and Linton 2004).

When the ABC protein involved in histidine uptake, HisP, was crystallized with ATP, several contact points were observed, Walker A lysine, serine, and threonine, Walker B aspartate and glutamic acid, the Q- and H-loops contacted ATP through interactions with a water molecule and a tyrosine interacts with the adenosine ring of ATP (Hung *et al.* 1998). A similar arrangement was observed for Rad50 protein, here the crystal structure revealed an induced fit (Karpowich *et al.* 2001) following the addition of ATP in which an ordering of the Walker A, Walker B, and Q-loop, in addition of a rotation of the α-helical domain to the β-sheet domain was reported (Hopfner *et al.* 2000).

In *R. leguminosarum* strain Rlt100 the ABC protein RhaT, has a high conservation of the Walker A, Walker B, signature sequence, Q-, D-, and H-loop conserved motifs.

1.2.6 ATP binding and ABC protein dimerization

The binding of an ATP molecule within the NBD of ABC protein monomer occurs with electrostatic interaction of the β- and γ-phosphate of ATP (Jones and George 2004), with the glycine rich P-loop of Walker A in arm I, likely focused at the conserved lysine residue, along with contact of the Mg^{2+} ion of ATP with an acidic residue, usually aspartate, within the β-sheet of Walker B (Locher 2004). The geometry of the ATP molecule is maintained by the acidic residue contact with a Mg^{2+} ion of ATP, and hydrogen bonding between the nitrogen atoms in the main chain of the peptide and the phosphates of the ATP (Altenberg 2003). The γ-phosphate of the ATP molecule causes an induced fit of the ATP by altering the conformation of the Q-, D-, H- and P-loops, thereby changing the surface of the monomer (Higgins and Linton 2004, Karpowich *et al.* 2001).

Rotation of α-helical arm II in relation to arm I β-sheets, is necessary for generating the correct alignment of the signature sequence of the opposite monomer to properly engage the NBD active site in a head-to-tail arrangement (Hopfner *et al.* 2000, Hopfner and Tainer 2003).

A proposed mechanism of ATP hydrolysis by ABC proteins involves an nucleophilic amino acid within the NBD acting as a base, donating electrons to a water molecule thereby increasing the potential for the water to attack the bond between the β- and γ-phosphate of the ATP (Altenberg 2003, Brönsted 1927). The conserved aspartate in Walker B interacting with the Mg^{2+} ion of ATP (Locher 2004), and a glutamate at the N-terminus of the D-loop (Hopfner *et al.* 2000) are believed to act as a base in the hydrolysis reaction (Hung *et al.* 1998). An oxygen atom within the peptide backbone near

the C-terminus interacts with a nucleophillic water molecule in the opposite active site (Jones and George 2004).

NBDs are flexible in the absence of ATP (Chen *et al.* 2003, Verdon *et al.* 2003), but when ATP is bound it induces a rigid arm I and arm II rotation with respect to each other, aligning Walker A, B, and signature sequence (Chen *et al.* 2003, Hopfner *et al.* 2000, Smith *et al.* 2002). This conformational change may occur by the interactions of γ-phosphate with conserved glutamine of the signature sequence motif, and conserved glutamine in the Q-loop with bound nucleotide and hydrolytic water (Higgins and Linton 2004, Smith *et al.* 2002).

Transition states of ATP hydrolysis have been reported using ATP analogues and hydrolysis deficient mutants (Chen *et al.* 2003, Hopfner *et al.* 2000, Smith *et al.* 2002). In this transition state, the signature sequence of one monomer interacts with ATP, moving several angstroms towards the opposite NBD pushing the ATP molecule into the P-loop (Locher *et al.* 2002) leading to tighter dimer interaction. ATP hydrolysis can only occur is both subunits are properly aligned to dimerize.

1.2.7 ATP hydrolysis

One model of ATP derived energy of transport suggests that the power stroke of transport is the ATP-dependent dimerization of ABC proteins that leads to the transport of substrate. Following ATP hydrolysis, electrostatic and conformational changes lead to product release and dimer separation resetting the transporter, pushing and pulling the TMD as they cycle (Chang and Roth 2001, Davidson 2002, Hopfner *et al.* 2000, Horlacher *et al.* 1998, Moody *et al.* 2002, Smith *et al.* 2002, Thomas and Hunt 2001,

Urbatsch *et al.* 2003). Crystal structure data from maltose MalK, suggests two ATP induces the formation of the closed dimer, and separation occurs after hydrolysis (Chen *et al.* 2003).

The other mechanism suggested in part by kinetic data generated by P-gp transport, whereby each monomer of the dimer acts 180° out of phase with the other ensuring the correct activity at only one NBD at a time. Here, either or both of the NBDs are loaded with an ATP molecule and is engaged by the opposite signature sequence at an orientation conducive to dimerization. With the rotation of the α-helical arm II when prompted by the TMD, one of the subdomains at a time is able to engage and disengage alternately during ATP hydrolysis (Jones and George 2004). This mechanism suggests that the ABC domains remain in contact during the transport reaction, and form the tight dimer interface as ATP is bound and hydrolyzed (Locher 2004, Locher *et al.* 2002). It was predicted that minor changes at the NBD could be amplified within the TMD-ABC interface (Locher 2004). Alternating ATP hydrolysis mechanism by NBDs was shown biochemically (Lamers *et al.* 2003, Senior 1998), coupled to an alternating mechanism in the substrate specific dimer (Lamers *et al.* 2003, van Veen *et al.* 2000, Wang *et al.* 2000). This asymmetric pattern of ATP hydrolysis suggests that rather than have one ATP hydrolyzed in one site and have the monomer dissociate wasting the other bound ATP, a process where asymmetric hydrolysis occurs and the first site either contains hydrolysis products or is empty (Higgins and Linton 2004, Jones and George 2004, Locher 2004). Biochemical data from Mdl1p, a mitochondrial transporter from *Saccaromyces cerevisiae* supports this model (Janas *et al.* 2003). If ATP hydrolysis occurs in an alternating fashion the outward rotation of the α-helical domain moving the signature sequence away from

the NDB would allow the hydrolysis products to be released (Jones and George 2004, Yuan *et al.* 2001). Displacement of ADP from Walker A with signal transmitted from TMDs may allow ATP to enter the NBD (Verdon *et al.* 2003), or ATP might be able to access the NBD P-loop at any time with low affinity, and signals from the TMD would cause induced fit of the opposite signature sequence leading to tight dimerization (Higgins and Linton 2004). Following ATP hydrolysis electrostatic repulsion of the Walker A/ADP to the signature sequence/ P_i destabilizes the ATP dimer (Smith *et al.* 2002). The α-helical, arm II rotates back relative to arm II, P_i and ADP release and the open dimer conformation is established (Higgins and Linton 2004). It appears that P_i is released before ADP based on vanadate trapping experiments, these also indicate that the open dimer conformation is not restored with only the release of P_i (Higgins and Linton 2004, Janas *et al.* 2003, Payen *et al.* 2003, Senior and Gadsby 1997, Sharma and Davidson 2000).

1.2.8 Transport mechanism

Vanadate is an non-hydrolysable ATP analogue that mimics the γ-phosphate, essentially freezing the transporter in a transition state by trapping the vanadate in the ABC-ATPase during hydrolysis (Chen *et al.* 2001). When vanadate was used to study the maltose transporter, the PBP became tightly associated with the TMD components and the affinity for maltose was decreased. Together this suggests that the PBP and ABC protein communicate via the TMD and PBP stabilizes the transition state in ATP hydrolysis (Chen *et al.* 2001).

An α-subdomain of the ABC protein, interacts with the TMDs and is involved in transmitting conformational changes in the substrate translocation process (Hyde *et al.* 1990). ABC proteins have the capability to interact with the cytoplasmic EAA loops of TMD via a segment located between the Walker A and Walker B motif of the ABC protein (α-helix 1 to α-helix 5 of NBD) that is highly variable in sequence, length, and structure (Figure 1.5) (Chang and Roth 2001, Higgins *et al.* 1986, Hunke *et al.* 2000b, Locher 2004, Mourez *et al.* 1997). This variability is probably due to the interactions they have with the highly variable TMDs EAA loops. There is low amino acid conservation at the transmission interface between the TMDs and the ABC protein, however the structure is conserved (Locher and Borths 2004). Mutations within the conserved architecture of the TMD-ABC interface adversely affect the protein folding, assembly, and the coupling of ATP hydrolysis to transport (Schmitt *et al.* 2003). This is observed in the CFTR (cystic fibrosis transmembrane conductance regulator) where mutations within the interface leads to 70% of cystic fibrosis cases (Locher 2004, Riordan 1993).

The rotation of α-helical arm II of ABC protein relative to arm I appears to be influenced by the TMD. It has been suggested that substrate binding to the TMDs leads to allosteric mechanism whereby the α-domain of ABC protein is prompted to undergo a conformational change by rotation bringing the signature sequence closer to the opposite NBDs P-loop, forming a closed dimer interface (Jones and George 2002, Loo *et al.* 2003, Szakacs *et al.* 2000, Yuan *et al.* 2001). TMDs stimulate ABC protein dimerization which in turn stimulated ATP hydrolysis. While there is structural conservation of the NBD core, it was observed that the α-domain exhibits greater positional variation (Groeger and Koster 1998, Jones and George 2004, Karpowich *et al.* 2001, Yuan *et al.* 2001).

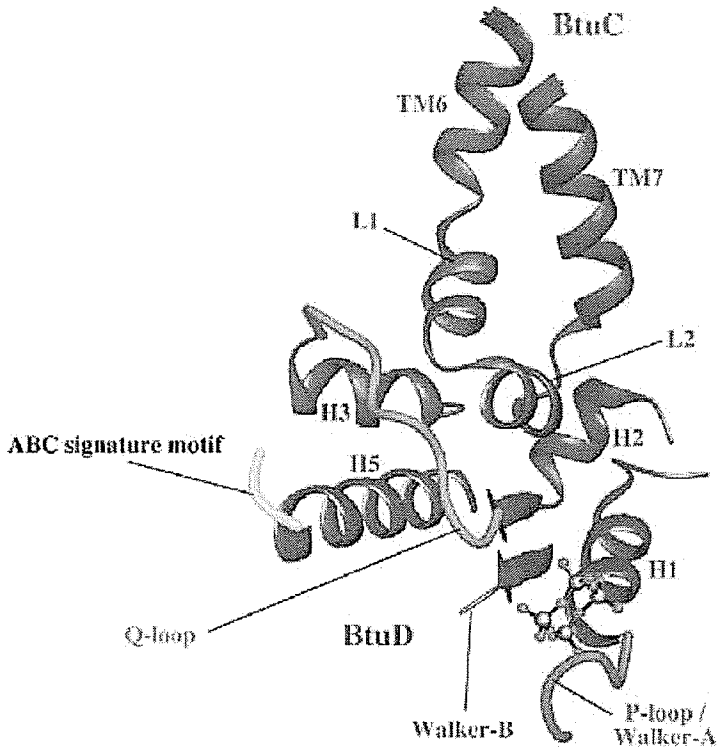


Figure 1.5. TMD to ABC protein interface of BtuC-BtuD vitamin B_{12} ABC transporter. BtuD ABC proteins (blue), Q-loop (green) contacts BtuC TMD (red) at around the L1 and L2 α-helices. ATP is bound to the NBD core, the signature sequence is shown in yellow. Reprinted by permission of Federation of the European Biochemical Societies from ABC transporter architecture and mechanism: implications from the crystal structures of BtuCD and BtuF, by Locher, K. P. and Borths, E., FEBS Letters, volume **564**(3), pp. 264-268, copyright 2004.

The trigger that initiates ATP hydrolysis is unknown. It is possible that it is the release of substrate into the TMD that triggers this or ATP hydrolysis may occur automatically with dimer formation with the correct geometry (Higgins and Linton 2004). Active site geometry is mediated by the signature sequence and induces global conformational changes resulting in substrate transport (Jones and George 1999, Jones and George 2004).

1.2.9 ABC unknowns

While much focus has been the energetics of transport many fundamental questions are yet to be answered. It is understood that substrate bound PBP interacts with the TMD through conserved residues however, we still do not fully understand the mechanism by which the binding of substrate bound PBP triggers changes in the TMDs (Locher 2004), leading to the control of the ATPase activity in the NBDs (Jones and George 2004). Nor do we fully understand the mechanism by which the energy derived from the binding and hydrolysis of ATP is translated to conformational changes in the TMDs with respect to substrate translocation (Jones and George 2004). It is known that the interface between the TMDs and the ABC protein are critical for communication within the ABC complex; however specifics concerning this interface are not known. While the overall process of transport is understood many of the intricate interactions and internal communications that appear to lead to overall global conformational changes within the complex to allow transport have proven much more difficult to elucidate.

1.2.10 Rhizobia ABC transporters

Early estimates for *Rhizobium leguminosarum* bv. *viciae* predicted that strain Rlv3841 contained 183 complete ABC operons, distributed throughout the six plasmids and chromosome for a total of 11.2% of the total protein complement; ABC transport genes constitute 27% of the coding capacity of pRL9 (Young *et al.* 2006). A recent publication suggests that there are 269 ABC genes encoded by this strain (Mauchline *et al.* 2006).

In the *Rhizobium etli* genome the ABC transporter family of proteins represents the largest constituent of protein families; they are distributed throughout the genome and have representatives on all six plasmid, with the greatest concentration found on p42f after the chromosome (Gonzalez *et al.* 2006).

Sinorhizobium meliloti strain Rm1021 pSymA contains as many as one in seven genes predicted to be involved in transport (Barnett *et al.* 2001). Genes encoding transport systems account for 12% of the coding capacity of the entire genome, making them the largest class of genes in the *Sinorhizobium meliloti* genome (Galibert *et al.* 2001). There are 200 ABC genes in S. meliloti, 146 are involved in uptake functions and 18 function as exporters (Mauchline *et al.* 2006).

In *Mesorhizobium loti* strain MAFF303099, 10.6% of the chromosome is dedicated to transport and binding proteins, pMLa and pMLb contain 12.8% and 6% for a total of 10.5% of the entire genome dedicated to transport and associated binding proteins (Kaneko *et al.* 2000). There are a total of 216 ABC genes encoded by *M. loti* (Mauchline *et al.* 2006).

Bradyrhizobium japonicum strain USDA110 contains 9% of the genome dedicated to transport proteins, most are from the ABC transport family (Kaneko *et al.* 2002). There are 240 ABC genes found in *B. japonicum* (Mauchline *et al.* 2006).

Agrobacterium tumefaciens contains 15% of the genome dedicated to transport and 60% of these belong to the ABC transporter family, for a total of 153 complete transporters in addition to "orphan" subunits (Wood *et al.* 2001).

In the soil environment where the nutrients are generally limiting, a bacterium with an arsenal of mechanisms to increase nutrient uptake would have increased fitness. While the localization of ABC transport genes differs amongst the rhizobia, many rhizobia genomes contain a relatively large amount of coding space dedicated to this family of proteins.

1.3 Mutarotases

1.3.1 Introduction to mutarotases

Monosaccharide in solution can assume several forms and there is a dynamic equilibrium between the different conformations of the monosaccharide (Drew 1998). They may exist in a furanose (one oxygen four carbon ring structure), pyranose (one oxygen five carbon ring structure) or open chain conformations. The C-1 carbon atom of a cyclic monosaccharide may adopt different anomeric states based on the orientation of hydroxyl groups associated with the C-1 carbon atom generating α- or β-anomers of the sugar (Drew 1998, Hucho and Wallenfels 1971). Anomers are stereoisomers of cyclic monosaccharides that are generated depending on the orientation of the hydroxyl group relative to the anomeric carbon. If the hydroxyl group is *in trans* to the carbon it is considered to be in the α-anomeric orientation, if it is in a *cis* position, it is considered to be in a β-anomeric conformation (Figure 1.6).

Kinetically, spontaneous α- or β-anomeric conversion can occur rapidly for monosaccharides in furanose conformations. This is attributed to the lower energy barrier for ring opening (Kim *et al.* 2001, Ryu *et al.* 2004). In contrast, spontaneous α- to β-anomeric conversion of pyranose monosaccharide occurs slowly. Glucose was demonstrated to spontaneously convert at a rate of 0.015 conversions per minute (Levy and Cook 1954, Livingstone 1977, Swain 1952).

Enzymes involved in initial monosaccharide modification in catabolic pathways often have a preference for a specific anomeric configuration of the monosaccharide (Ryu *et al.* 2004, Sigrell *et al.* 1999). For example, galactokinase uses only α-galactose

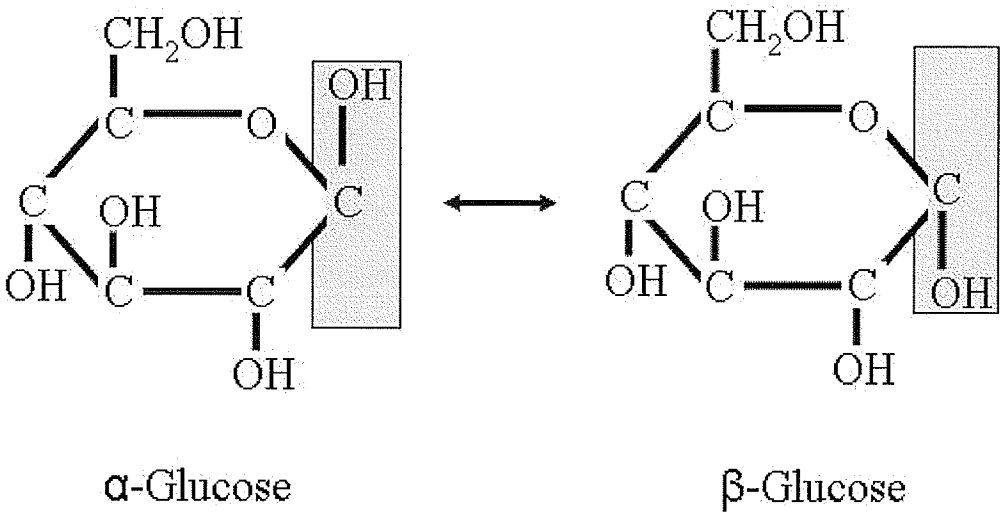


Figure 1.6. Mutarotation between α-glucose and β-glucose in the cyclic pyranose-form.

to generate α-galactose-1-phosphate (Howard and Heinrich 1965, Wilkinson 1949), or glucose and galactose dehydrogenase prefer the β-anomer of each sugar (Pauly and Pfleiderer 1975, Wallenfels and Kurz 1962), if the preferred form is not readily available, efficient use of that particular monosaccharide may be inefficient (Bouffard *et al.* 1994).

Mutarotation is the interconversion of the monosaccharide between the anomeric forms. The role of an enzymatic mutarotase in α- or β-anomerization of the hydroxyl groups appears to be necessary to support fast energy generation by supplying the preferred anomeric form of the monosaccharide even if that form is available at low concentrations in the dynamic equilibrium (Ryu *et al.* 2004). While the catalytic mechanisms differ slightly for the classes of mutarotases, each follow acid-base catalysis (Brönsted 1927) involving a highly conserved histidine residue to generate a protonation of the oxygen atom in the furanose or pyranose ring alone or in concert with another histidine or large polar residue functioning as a base to abstract a hydrogen from the C-1 carbon, initiating nucleophillic attack and sugar ring rearrangement (Kim *et al.* 2003, Ryu *et al.* 2004, Ryu *et al.* 2005, Thoden *et al.* 2003).

Below is a brief description of three previously characterized mutarotase of *E. coli*: galactose GalM (Beebe *et al.* 2003, Beebe and Frey 1998, Bettenbrock and Alpert 1998, Bouffard *et al.* 1994, Holden *et al.* 2003, Thoden and Holden 2002, Thoden *et al.* 2002, Thoden *et al.* 2003, Timson and Reece 2003), fucose FucU (Kim *et al.* 2003, Ryu *et al.* 2004) and L-rhamnose YiiL (Ryu *et al.* 2004, Ryu *et al.* 2005).

1.3.2 Galactose mutarotation GalM

In most organisms, lactose is split by β-galactosidase to form glucose and β-D galactose. Enzymes within the Leloir pathway then convert β-D-galactose to the more metabolically useful glucose-1-phosphate where phosphoglucomutase catalyzes the isomerization of glucose-1-phosphate to glucose-6-phosphate to be used in glycolysis. Here, β-D-galactose is epimerized to α-D-galactose by galactose mutarotase, GalM (Bouffard *et al.* 1994, Erlandson *et al.* 2001, Wallenfels *et al.* 1965). Subsequent phosphorylation, uridylyltransferase, and epimerization of α-D-galactose by the actions of galactokinase (Ballard 1966), galactose-1-phosphate uridylyltransferase (Kalckar 1960), and UDP-galactose 4-epimerase (Wilson and Hogness 1969) respectively, generate glucose-1-phosphate as reviewed by Holden *et al.* (Frey 1996, Holden *et al.* 2003).

In humans, defective galactokinase, galactose-1-phosphate uridylyltransferase, or UDP-galactose 4-epimerase can lead to galactosemia, a rare genetic metabolic disorder where GALT (galactose-1-phosphate uridylyltransferase) activity is diminished, preventing it from breaking down galactose into glucose, leading to toxic levels of galactose in the blood. Symptoms of galactosemia include mental retardation, enlarged liver, renal failure and cataract formation (Novelli and Reichardt 2000, Petry and Reichardt 1998). Defects in GalM have not yet been linked to a diseased state (Timson and Reece 2003).

Galactose mutarotase has been reported (Wallenfels *et al.* 1965), defined (Bouffard *et al.* 1994) and recently characterized in *E. coli* and *Lactobacillus lactis* (Beebe and Frey 1998, Erlandson *et al.* 2001, Hucho and Wallenfels 1971, Thoden and Holden 2002, Thoden *et al.* 2002), as a mutarotase that accelerates the conversion rate

between the α- and β-anomers of both D-galactose and D-glucose (Ryu *et al.* 2004). GalM of *E. coli* consists of 346 amino acid residues and functions as a monomeric enzyme (Ryu *et al.* 2004). The substrate binding pocket of GalM is deep and contains embedded water molecules and a strong hydrogen bonding network (Ryu *et al.* 2004). While the precise mapping of the catalytic amino acid residues is still uncertain (Beebe *et al.* 2003, Thoden and Holden 2002, Thoden *et al.* 2002), kinetic analysis of GalM in both *E. coli* and *L. lactis* leads to a proposed acid-base catalytic mechanism (Brönsted 1927) whereby a highly conserved glutamate, in the active site, serves as a base to abstract the C-1 hydroxyl hydrogen while a histidine protonates the C-5 ring oxygen allowing epimerization of β-D-galactose to α-D-galactose (Beebe and Frey 1998, Thoden *et al.* 2003). α-D-Galactose is the substrate for the galactokinase, the next step in the Leloir pathway.

1.3.3 Fucose FucU

FucU is located in the fucose regulon containing genes for fucose proton symport and catabolism including fucose permease, fucose isomerase, fuculose kinase, fuculose-1-phosphate aldolase, 1,2-propanediol oxidoreductase, regulatory protein, and FucU (Blattner *et al.* 1997, McClelland *et al.* 2001, Zhu and Lin 1988). In *E. coli,* there are two pathways involved in fucose catabolism. In the first pathway the first step in the degradation of fucose is an isomerization of L-fucose into L-fuculose, which is subsequently phosphorylated into fuculose 1-phosphate (Chen *et al.* 1987). In the second pathway, L-fucose is converted into L-fucose-1-phosphate by fucose kinase as the first step in the synthesis of oligo- and polysaccharides containing fucose (Park *et al.* 1998).

FucU homologues are ubiquitous and have been reported in both bacteria and eukaryotes. FucU was reported to have 140 amino acid residues and it shares 50% identity to the human FucU homologue (Kim *et al.* 2003). FucU contains no signal sequence so it is considered to function as a cytoplasmic protein that has a decameric quaternary structure with a shallow binding pocket (Kim *et al.* 2003). FucU is considered to be an anomeric mutarotase to fucose with additional pyranase activity at low levels to ribose (Ryu *et al.* 2005), however its primary role is to bind and accelerate the conversion between α- and β-fucopyranose (Ryu *et al.* 2005).

1.3.4 Rhamnose YiiL mutarotase

YiiL mutarotase is a mutarotase that catalyses α- and β-anomeric conversion of L-rhamnose a methyl-pentose sugar (Ryu *et al.* 2004, Ryu *et al.* 2005). While furan forms of sugars like D-ribose have been shown to interconvert between α- and β-anomeric forms in solution quite rapidly, pyran forms of sugars like L-rhamnose have slow spontaneous α to β conversion rates (Ryu *et al.* 2004, Ryu *et al.* 2005).

In *E. coli*, YiiL consists of 104 amino acid residues; crystallization studies have revealed the YiiL functions as an asymmetric dimer that binds one rhamnose molecule per subunit and is stabilized by hydrogen bonds within intermolecular β-sheet, various hydrophobic interactions, and a cation–pi interaction with a salt-bridge (Ryu *et al.* 2005). Within the binding pocket of YiiL there are histidine (His-22), tryptophan (Trp-76, Trp-77) and tyrosine (Try-18, Try-41) residues that form hydrogen bonds with all of the hydroxyl groups of the bound L-rhamnose molecule (Figure 1.7A) (Ryu *et al.* 2005). There are also two leucines (Leu-29, Leu-33), an isoleucine (Ile-25) along with a

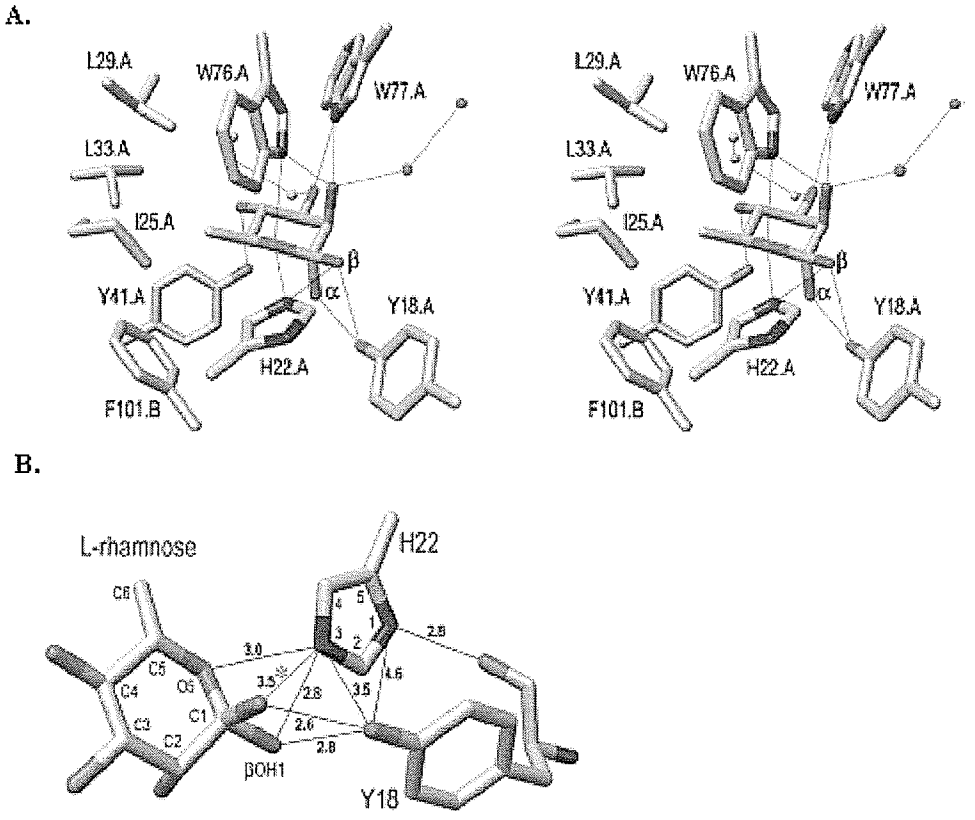


Figure 1.7. Active site configuration of L-rhamnose mutarotase YiiL in *E. coli*. **A**. Stereo-view of YiiL with several contact points (lines) within the catalytic active site of the binding pocket. **B**. Functional groups involved in catalysis are shown with projected distances between contact points. Asterisk, *, denotes the α-anomeric hydroxyl group added manually with geometry optimized (Ryu *et al.* 2005). Reprinted from the Journal of Molecular Biology, volume **349**(1), Ryu, K. S., Kim, J. I., Cho, S. J., Park, D., Park, C., Cheong, H. K., Lee, J. O. and Choi, B. S., Structural insights into the monosaccharide specificity of *Escherichia coli* rhamnose mutarotase, pp. 153-162, copyright 2005, with permission from Elsevier.

phenylalanine (Phe-101) from the adjacent monomer of the dimer structure involved in forming the hydrophobic pocket necessary to accommodate the methyl C-5-moiety of rhamnose (Ryu *et al.* 2004, Ryu *et al.* 2005). The YiiL binding pocket is deep and may contain embedded water molecules that may enhance the hydrogen bonding interaction of the bound ligand (Ryu *et al.* 2005).

The catalytic mechanism of YiiL has yet to be defined; however, the proposed mechanism of anomerization is dependent upon the ability of His-22 to donate and accept protons from L-rhamnose O-5 and C-1 hydroxyl in concert with a stabilization effect of helix capping interaction by Tyr-18 OH on the N-1 imidazole ring of His-22 via a hydrogen bond (Figure 1.7B) (Ryu *et al.* 2005, Serrano *et al.* 1992). This interaction may aid in stabilizing the short α-helix in addition to sustaining the α- and β-OH-1 orientations during catalysis (Ryu *et al.* 2004, Ryu *et al.* 2005).

1.3.5 Rhizobium Mutarotase

To our knowledge there has not been any characterization of a mutarotase of any *Rhizobium*. In Chapter 4 we discuss RhaU, a mutarotase from *R. leguminosarum* that shares some catalytic core amino acid identity to YiiL of *E. coli* (Ryu *et al.* 2004, Ryu *et al.* 2005). Putative RhaU-like genes exist in other organisms closely related to *R. leguminosarum*. These genes are generally located within loci encoding sugar catabolism; it would not be unreasonable to predict that there may be many rhizobia mutarotases specific for several types of sugars.

1.4 Kinase ATPase

1.4.1 ATPase fingerprints

Folding motifs can be conserved amongst proteins, and often as few as one or two amino acid residues can be used to identify a family or class of proteins that are similar in structure. Nucleotide binding domains are examples of these types of conserved motifs (Kabsch and Holmes 1995). Similar folds were observed for a family of proteins (Walker *et al.* 1982) including adenylate kinase (Schulz *et al.* 1974), F1-ATPase, myosin (Rayment 1993), dynein (Gibbons *et al.* 1991, Ogawa 1991), elongation factor EF-Tu (Jurnak 1985, la Cour *et al.* 1985), as well as ABC proteins associated with ABC transporters described above. The signature for these proteins (as I have discussed in detail throughout section 1.2.5.4 through section 1.2.7) is the pyrophosphate loop (P-loop) and consists of Gly-X4-Gly-X-Gly-Lys-Ser-Thr, where X is an arbitrary residue (Kabsch and Holmes 1995, Saraste *et al.* 1990). The presence of the P-loop suggests ATP binding (Kabsch and Holmes 1995, Saraste *et al.* 1990). There is also other nucleotide binding domain that is very different from the P-loop type-motif called the sugar kinase-hsp70-actin protein superfamily of ATPases.

1.4.2 Introduction to the sugar kinase-hsp70-actin ATPase protein superfamily

Many sugar kinases, 70 kDa heat shock-related proteins (hsp70) and actin are evolutionarily related and belong to the same class of protein (Bork *et al.* 1992). These proteins exhibit very different biological functions and overall, have very little sequence similarity (Flaherty *et al.* 1990, Flaherty *et al.* 1991, Kabsch *et al.* 1990).

In bacteria and archaea the majority of carbohydrate catabolism occurs in the cytoplasm, phosphorylation of sugars, via sugar kinases, following transport across cytoplasmic membranes prevents the phosphorylated sugar from returning back across the cytoplasmic membrane (Wilkinson and Verschueren 2003).

70 kDa heat shock-related proteins are involved in protein folding and chaperone activity within cells, repetitive substrate binding and release in an ATP dependent cycle enable proteins to be folded and refolded correctly (Bukau and Horwich 1998, Craig 1989, Georgopoulos 1992, Gething and Sambrook 1992, Hendrick and Hartl 1993, Pelham 1988, Pelham 1986).

Actins are ubiquitous proteins that can assemble into filaments used in the cytoskeleton. They function in muscular contraction in concert with myosin and cell motility (Kabsch and Vandekerckhove 1992, Schutt *et al.* 1993, Sheterline *et al.* 1995).

The basic function that is common to all the members of this protein superfamily is adenosine triphosphate (ATP) hydrolysis or phosphotransferase activity (Hurley 1996). Members of this class of proteins can be identified by a similar topological homology containing conserved or similar-type residues within the ATPase active site (Bork *et al.* 1992, Flaherty *et al.* 1990, Flaherty *et al.* 1991, Kabsch *et al.* 1990) with an overall arrangement of β-sheets surrounded by α-helices similar to those found in RNAseH-like folds (Bork *et al.* 1992, Flaherty *et al.* 1991). Proteins in this superfamily have demonstrated a propensity for significant rotation upon substrate binding, believed to be necessary for nucleotide phosphate hydrolysis and phosphotransferase activity (Bork *et al.* 1992, Flaherty *et al.* 1990, Flaherty *et al.* 1991, Kabsch *et al.* 1990).

1.4.3 Conserved protein structure

Members of the sugar kinase-hsp70-actin protein superfamily consist of two conserved core domains called subdomain Ia and subdomain IIa (Figure 1.8) (Anderson *et al.* 1978, Bork *et al.* 1992, Feese 1994, Flaherty *et al.* 1990, Hurley *et al.* 1993, Schutt *et al.* 1993). The topological loop arrangements of both of these domains are similar, each domain consists of βββαβαβα arrangement and are believed to have been the result of a gene duplication (Bork *et al.* 1992, Flaherty *et al.* 1990, Flaherty *et al.* 1991, Kabsch *et al.* 1990). This arrangement of five central β-pleated sheets flanked by three α-helices, is conserved in all members of this superfamily (Bork *et al.* 1992). X-ray analysis of crystal structures of these domains can be superimposed with many of the members of this family with very little deviation (Hurley 1996, Kabsch and Holmes 1995).

Some members of this protein superfamily have additional variable subdomains called subdomain Ib, and subdomain IIb, continuous with domains Ia and IIa respectively; they are involved in specialized functions corresponding to the member of the superfamily that they are located within (Figure 1.9) (Abendroth *et al.* 2004, Bork *et al.* 1992, Flaherty *et al.* 1990, Flaherty *et al.* 1991).

Analysis of several proteins of this superfamily revealed five common secondary structure motifs used to identify members (Figure 1.8) (Bork *et al.* 1992). They have been described and characterized as being conserved based on either invariant amino acids or based on the biochemical properties of like-amino acids. Invariant residues within these conserved motifs, regardless of where they are located in the physical sequence, are mostly localized within ~8 Å of the bound ATP (Kabsch *et al.* 1990) and likely constitute the ATP binding pocket. They correspond to regions of contact with the phosphates and

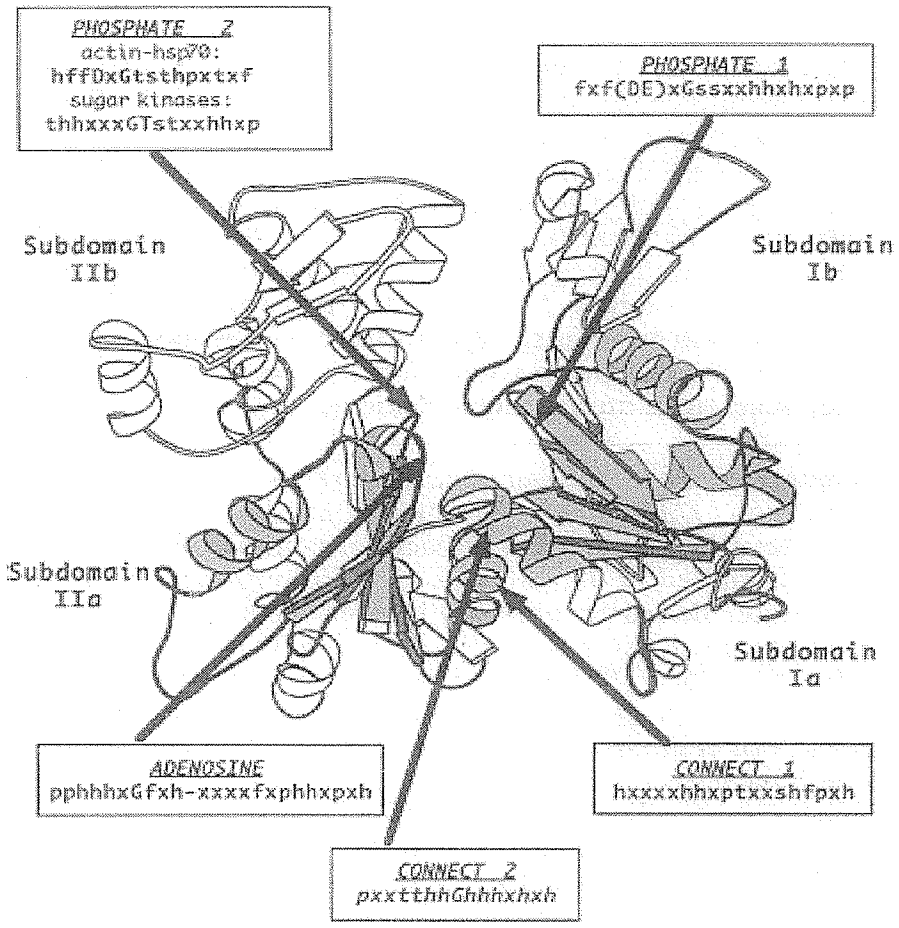


Figure 1.8. Representation of actin highlighting the common motifs of the kinase-hsp70-actin protein superfamily. Arrows highlight the conserved characteristics for the motifs in this superfamily. Letters are as follows: h = hydrophobic (VLIFWY); f = partly hydrophobic (VLIFWYMCGATKHR); t = tiny (GSAT); s = small (GSATNDVCP); p = tiny polar (GSATNDQEKHR); x = variable. Darker shading represents structure common to all members of this superfamily, domains named as per Hsp70 crystal structure (Flaherty *et al.* 1991) (Bork *et al.* 1992). Reprinted from The Proceedings of the National Academy of Sciences of the United States of America, volume 89, Bork, P., C. Sander, and A. Valencia, An ATPase domain common to prokaryotic cell cycle proteins, sugar kinases, actin, and Hsp70 heat shock proteins, pp. 7290-7294, copyright 1992, with permission by Peer Bork.

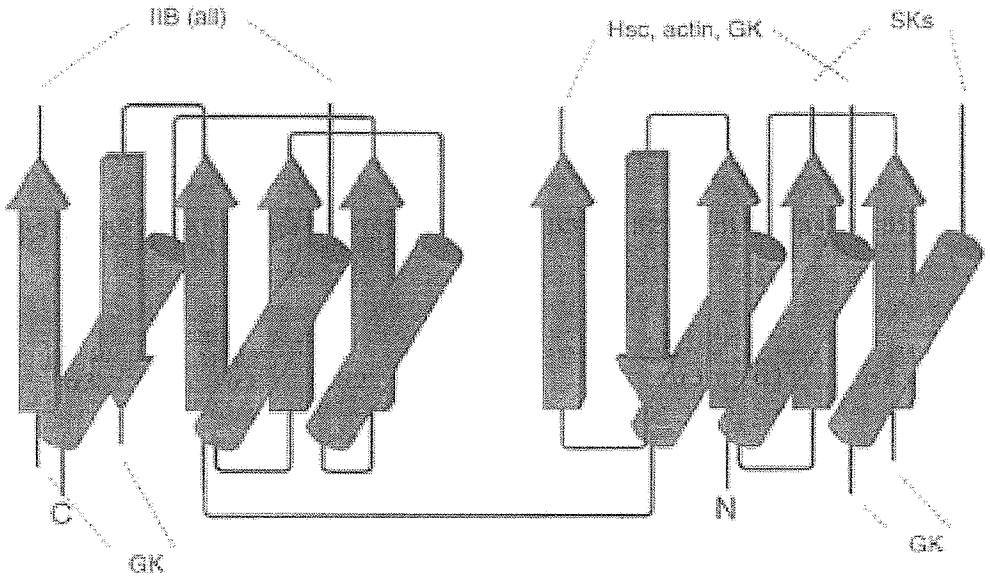


Figure 1.9. Diagrammatic representation of the topology of the conserved core in members of the kinase-hsp70-actin protein superfamily. Non-conserved regions are indicated, SK = sugar kinase (Hurley 1996). Reprinted, with permission, from the Annual Review of Biophysics and Biomolecular Structure, Volume 25 © 1996 by Annual Reviews www.annualreviews.org.

adenine of ATP and the connecting regions of subunits Ia and IIa (Bork *et al.* 1992, Kabsch and Holmes 1995).

1.4.4 Conserved motifs within the ATP-substrate binding pocket

The substrate-specific portion of the pocket is variable to account for the diverse functions of these proteins: glycerol binding interactions for glycerol kinases will be discussed later in this chapter as an example. The nucleotide phosphate binding pocket residues involved in direct interactions as well as residues necessary for catalytic functions are conserved and often invariant.

In this protein superfamily subdomains Ia and IIa connect via a large cleft containing a substrate binding site that enables the protein to bind a nucleotide phosphate like ATP, and a protein-specific substrate (Bork *et al.* 1992). Sugar kinases bind the sugar substrate deep within the cleft "below' the binding site of ATP where subdomain Ia and subdomain IIa connect (Figure 1.8) (Hurley 1996).

β1 and β2 sheets of domain Ia are connected by a loop, the main-chain nitrogen atoms of several residues within this loop contact the phosphates of bound nucleotide phosphate (Hurley 1996). A residue from β2 of domain Ia, containing a basic moiety interacts with the β-phosphate of ADP (Hurley 1996). The peptide backbone amides of the loop between β1 and β2, within domain IIa also interact with the phosphates of ATP (Hurley 1996).

Within domain Ia, a common motif was observed for this family, (Ile/Leu/Val)-X-(Ile/Leu/Val/Cys)-Asp-X-Gly-(Thr/Ser/Gly)-(Thr/Ser/Gly)-X-X-(Arg/Lys/Cys) (Flaherty *et al.* 1991). Here Asp binds Mg^{2+} ion, the invariant small polar Gly contacts the γ-

phosphate so as not to sterically hinder the interaction to ATP, -OH from the (Thr/Ser/Gly) and (Thr/Ser/Gly) residues bond the oxygen atoms from the ATP phosphates, the (Arg/Lys/Cys) at the end form a salt bridge with the nucleotide phosphate (Kabsch and Holmes 1995). While these interactions are generally conserved in ATP substrate binding, adenosine and ribose interactions differ between these types of proteins.

Based on differences between several key residues the sugar kinase-hsp70-actin protein superfamily can be differentiated. Sequence alignments of several sugar kinases (Bork *et al.* 1992) including glycerol kinase (McCabe 1995, Pavlik *et al.* 1993) show an invariant aspartic acid (Asp245 in glycerol kinase) that distinguish sugar kinases from the hsp70 and actin classes that contain equivalent glutamate and glutamine, respectively (Pettigrew *et al.* 1998).

A second defining residue is an aspartate located within the connection between β1 and β2 of domain Ia, represented by Asp10 in glycerol kinase (Pettigrew *et al.* 1998). Asp10 interacts with Mg^{2+} ion complexed with ATP and is crucial for ATP phosphotransferase (Pettigrew *et al.* 1998).

1.4.5 Domain conformational changes

While actin (Kabsch *et al.* 1990, McLaughlin *et al.* 1993, Schutt *et al.* 1993) and glycerol kinase (Hurley *et al.* 1993) have been crystallized in the closed form only (Hurley 1996), hexokinase was described in open and closed forms (Anderson *et al.* 1978, Bennett and Steitz 1980) with movements of α3 helices from both domains Ia and IIa relative to each other and the β-sheet core (Lesk and Chothia 1984) providing the

domain rotation (Hurley 1996). Conformational changes of the domains leads to the active site being less accessible to bulk solvent (Pettigrew *et al.* 1998). Domain rotation of up to 30° have been observed for some ATPases but this amount of rotation is not observed for all members in this superfamily (Bork *et al.* 1992, van den Ent and Lowe 2000). X-ray crystal analysis of some members of this family have demonstrated closed and open forms of kinase proteins in the presence and absence of ATP (Bennett and Steitz 1978, Steitz *et al.* 1981). In glycerol bound, closed forms of glycerol kinase, ATP can move in and out of crystals freely (Hurley *et al.* 1993).

1.4.6 Extended family: FGGY sugar kinases

The FGGY-type ATPase family belongs to the sugar kinase-hsp70-actin superfamily. While not all carbohydrate kinases belong to this family, FGGY-type sugar kinases include ribulose kinase, pentulose kinase, hexose kinase, fructose kinase and rhamnulose kinase (Bork *et al.* 1993). Several other FGGY-type sugar kinases have been identified as being more closely related, they include fucose kinase, gluconokinase, glycerol kinase, and xylulose kinase (Reizer *et al.* 1991). Among these one of the best characterized proteins is glycerol kinase (Bergmeyer 1961, Bublitz and Kennedy 1954, Wieland and Suyter 1957).

1.4.7 Glycerol kinase

Glycerol kinase (Bergmeyer 1961, Bublitz and Kennedy 1954, Wieland and Suyter 1957) catalyses the formation of glycerol-3-phosphate, using ATP- Mg^{2+} and glycerol as substrates in the reaction. Glycerol kinase in *E. coli* is a regulatory enzyme

involved in the phosphoenolpyruvate:glycose phosphotransferase system (PTS), responsible for the rate limiting phosphorylation of glycerol as it crosses the cytoplasmic membrane. Glycerol kinase is induced at a transcriptional level by its glycerol-3-phosphate product, and is also allosterically regulated independently by key PTS proteins, fructose 1,6-bisphosphate (FBP) and IIA (Glc) (Lin 1976, Novotny *et al.* 1985, Ormo *et al.* 1998, Pettigrew *et al.* 1990, Postma 1984).

This enzyme functions as a dimer, each monomer's active site can independently phosphorylate glycerol (Feese 1998). The crystal structure of glycerol kinase has been solved with glycerol kinase/glycerol/ADP, glycerol kinase/glycerol-3-phosphate/ADP, and glycerol kinase complexed with a non-transferable ATP analogue (AMP-PNP), believed to resemble the initial binding position of ATP, thereby providing useful structural information just prior to the phosphorylation step (Mao *et al.* 1999).

Several residues participate in binding the glycerol molecule via hydrogen bonds, Arg83, Try135, and Asp245 (Mao *et al.* 1999). There are also interactions with glycerol through Glu84, Trp103, Phe270 (Figure 1.10) (Feese 1998, Hurley *et al.* 1993, Mao *et al.* 1999).

Interactions of the adenosine base of ADP were predicted with Ala326, Ala412, Asn415, Ile313, Leu381, and Ile384 forming the hydrophobic adenosine binding pocket (Hurley *et al.* 1993, Mao *et al.* 1999). The main-chain amine carbonyl of Gly411 and the main-chain Gly310, forms hydrogen bonds with the α-phosphate and ribose respectively (Hurley *et al.* 1993). Gly411 also forms a hydrogen bond to the O1-oxygen of the α-phosphate of ATP (Hurley 1996, Kabsch and Holmes 1995). Hydrogen bonds are formed between Gln314, Gly310, and Ala412 and ribose of ATP

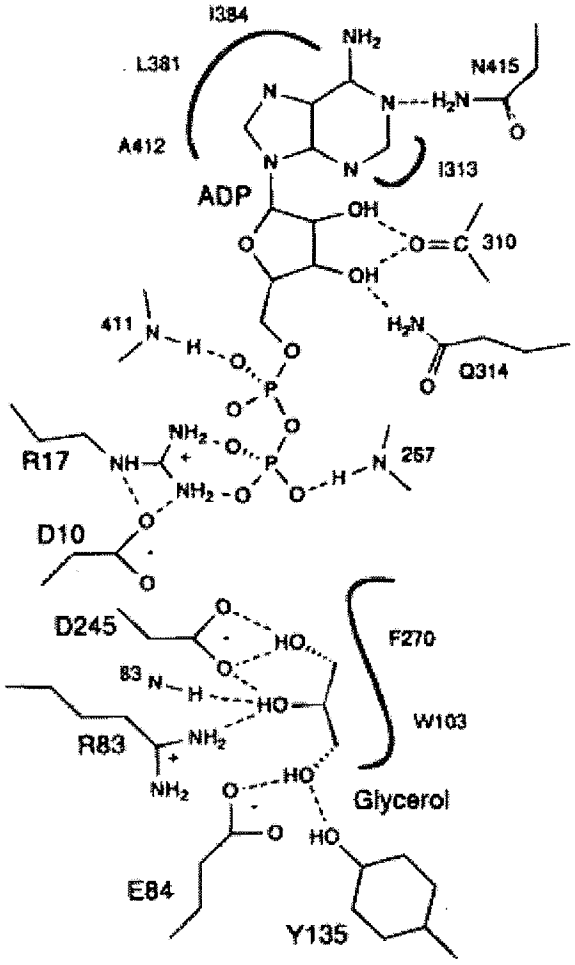


Figure 1.10. Glycerol and ADP binding to glycerol kinase within binding pocket. Residues contacting adenine of ATP: Asn415 (Asn395), Ile384 (Leu369), Leu381 (Cys366), Ala412 (Phe392); Residues contacting ribose of ATP: Gly310 (Gly310), Gln314 (Glu304); Residues contacting α-phosphate of ATP: Gly411 (Pro391); Residues contacting β-phosphate of ATP: Thr267 (Thr259), Arg17 (Lys20), **Asp10 (Asp13)**; Residues contacting glycerol: **Asp245 (Asp238)**, Arg83 (Gly81), Glu84 (Ala82), Trp103 (Tyr100), Phe270 (Val263), Tyr135 (Lys131). Residues in parenthesis are the predicted corresponding *R. leguminosarum* RhaK rhamnose kinase residues based on PRALINE data generated in Figure 3.4. Invariant residues from this family of proteins that are displayed in this figure are in bold in figure legend. Reprinted (abstracted/excerpted) with permission from Science, volume 259(5059), pp. 673-677, Hurley, J. H., Faber, H. R., Worthylake, D., Meadow, N. D., Roseman, S., Pettigrew, D. W. and Remington, S. J., Structure of the regulatory complex of *Escherichia coli* IIIGlc with glycerol kinase, Copyright 1993, AAAS. Reprint permission also granted by James Hurley.

(Hurley 1996, Kabsch and Holmes 1995) (Figure 1.10). The phosphates of ADP interact with Arg17, Gly266, Thr267, and Gly411 (Kabsch and Holmes 1995). Asp10 and Asp245 form the binding site for Mg^{2+} (Hurley 1996, Kabsch and Holmes 1995).

A β-hairpin containing Asp10, Gln11, Gly12, Thr13, Thr14, Ser15, Ser16, and Arg17 is crucial for γ-phosphate contact of ATP (Figure 1.11) (Mao *et al.* 1999), and is relatively conserved amongst members of the FGGY family. Here the amide side chains of Thr13, Thr14, and Ser15 interact with the γ-phosphate while Arg17, a large positively charged polar residue appears to mimic the role of a Mg^{2+} ion and balance the negatively charged γ-phosphate (Mao *et al.* 1999).

Using AMP-PNP, which resembles the initial binding position of ATP, to compare the crystallization data generated for ADP and glycerol-3-phosphate products, it was observed that the only significant movement in the active site was by the γ-phosphate in order to reach the 3-OH position of the glycerol, where the transfer of the phosphate occurs (Mao *et al.* 1999). The conformational changes required to generate this movement is not understood; however models of the catalytic mechanism have been proposed for glycerol kinase.

1.4.8 Model for catalytic mechanism of phosphotransferase in glycerol kinase

An ordered mechanism was proposed for glycerol kinase catalytic mechanism whereby glycerol binds within the cleft first, leading to significant conformational changes within the protein (Thorner and Paulus 1973), followed by increased affinity and binding of ATP in the binding pocket (Figure 1.12) (Mao *et al.* 1999). A proposed

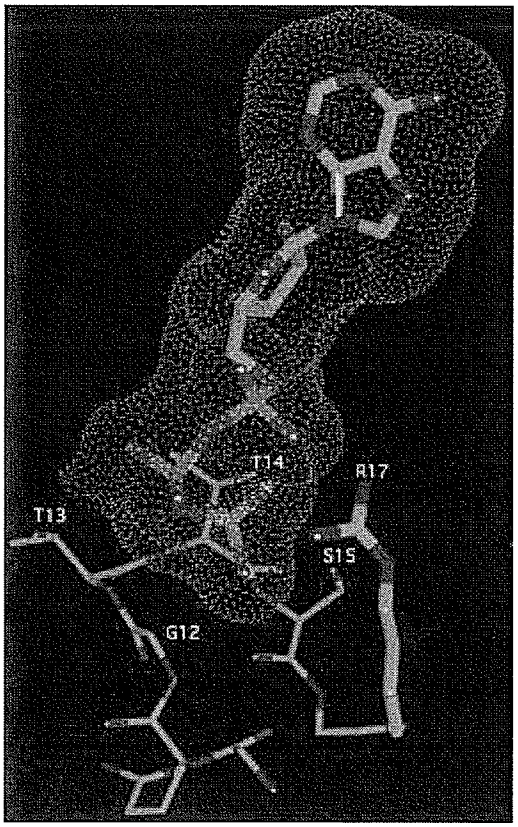


Figure 1.11. A β-hairpin in glycerol kinase containing crucial for γ-phosphate contact of ATP. ATP analogue AMP-PNP is shown within β-hairpin, colors are as follows: carbon = green; nitrogen = blue; oxygen = red; phosphate = magenta. Residues involved in for γ-phosphate contact that are displayed are listed: **Gly12 (Gly15)**, Thr13 (Lys16), Thr14 (Thr17), Ser15 (Asn18), Arg17 (Lys20). Residues in parenthesis are the predicted corresponding *R. leguminosarum* RhaK rhamnose kinase residues based on PRALINE data generated in Figure 3.4. Invariant residues from this family of proteins that are displayed in this figure are in bold in figure legend (Mao *et al.* 1999). Reprinted from Biochemical and Biophysical Research Communications, volume **259**(3), Mao, C., Ozer, Z., Zhou, M. and Uckun, F. M., X-Ray structure of glycerol kinase complexed with an ATP analog implies a novel mechanism for the ATP-dependent glycerol phosphorylation by glycerol kinase, pp. 640-644, Copyright 2004, with permission from Elsevier.

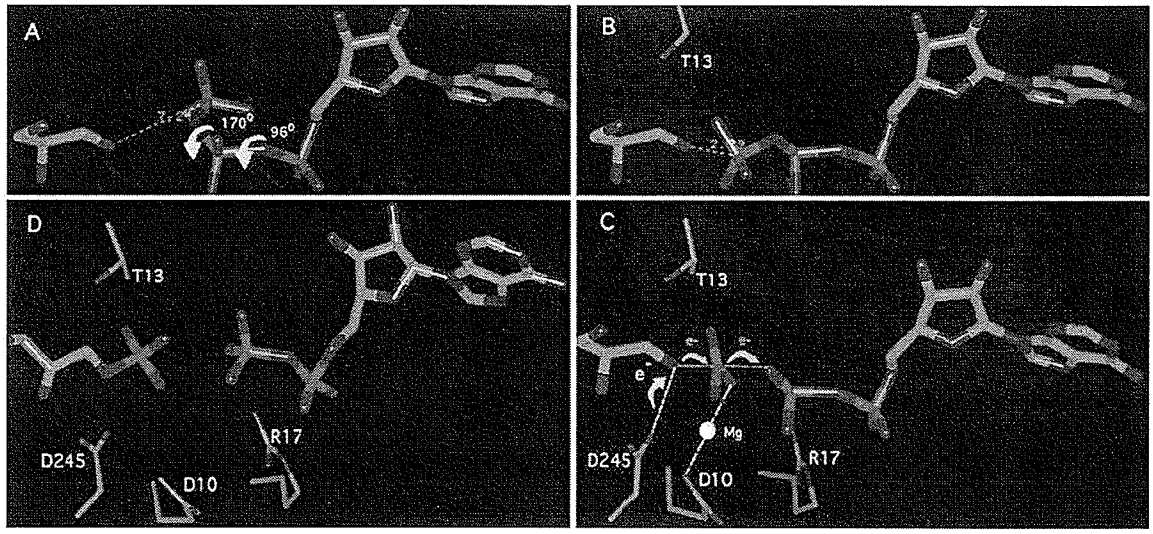


Figure 1.12. Model for catalytic mechanism of phosphotransferase in glycerol kinase. **A**. Structure of glycerol kinase/AMP-PNP/glycerol. **B**. Model of glycerol kinase/ATP/glycerol following rotation of γ-phosphate of ATP. **C**. Model of glycerol kinase in transition state before reaction occurs, electron transfer indicated and Mg^{2+} ion is shown. **D**. Post reaction crystal structure of glycerol kinase/glycerol-3-phosphate/ADP. Arg17 (Lys20), **Asp10 (Asp13), Asp245 (Asp238).** Residues in parenthesis are the predicted corresponding *R. leguminosarum* RhaK rhamnose kinase residues based on PRALINE data generated in Figure 3.4). Invariant residues from this family of proteins that are displayed in this figure are in bold in figure legend. Reprinted from Biochemical and Biophysical Research Communications, volume **259**(3), Mao, C., Ozer, Z., Zhou, M. and Uckun, F. M., X-Ray structure of glycerol kinase complexed with an ATP analog implies a novel mechanism for the ATP-dependent glycerol phosphorylation by glycerol kinase, 640-644, copyright 2004, with permission from Elsevier.

mechanism of γ-phosphate rotation and sequential contact with several threonine residues located near the binding pocket, Thr13, The14, and Thr267 enable the phosphate to swing around towards the 3-OH of glycerol decreasing the distance between the substrates to less than 3Å (Mao *et al.* 1999). In structure studies with ADP and glycerol with glycerol kinase suggest glycerol bind near the phosphates of ADP in contact with Asp245, via a hydrogen bond with 3-OH of glycerol (Hurley *et al.* 1993, Kabsch and Holmes 1995). Next, Asp245 acting as a base deprotonates the attacking OH of glycerol, while Asp10 interacts with Mg^{2+} ion of ATP-Mg stabilizing the negative charge in the transition state, for an in line attack mechanism (Figure 1.12) (Blattler and Knowles 1979, Mao *et al.* 1999). In line transfer of phosphate from ATP to sugar occurs without a phosphoenzyme intermediate (Blattler and Knowles 1979).

1.4.9 R. leguminosarum rhamnose kinase

Relatively few rhizobia sugar kinases have been characterized beyond identification and to our knowledge kinase proteins have not been identified as being associated specifically with an ABC transport complex. The rhamnose kinase gene in Rlt100 appears to encode a kinase protein, RhaK containing many conserved amino acid residues used to identify members of the sugar kinase-hsp70-actin protein superfamily. Biochemical assays confirm the presence of a phosphorylated intermediate dependent upon the presence of a functional RhaK protein. RhaK is currently in the process of being crystallized with the intention of solving the crystal structure and catalytic mechanism. Chapter 3 will focus on the role RhaK has in L-rhamnose catabolism and its function as a unique member of the sugar kinase-hsp70-actin protein superfamily.

1.5 Thesis objectives

Rhizobium leguminosarum biovar *trifolli* wild-type strain Rlt100 forms a symbiotic partnership with white clover (Baldani *et al.* 1992, Oresnik *et al.* 1998). Within root nodules *R. leguminosarum* fixes relatively inert atmospheric nitrogen into ammonia that can be incorporated into plant tissue.

Much research has been focused upon the chemical signal exchange between the rhizobia and plant host leading to the expression of genes involved in symbiosis of both parties before and during infection thread process and within the nodule as rhizobia adopt a bacteroid state. Research has also been aimed at competition within the soil and specifically within the rhizosphere. Studies characterizing the genomes of several symbiotic rhizobia have highlighted the relative importance of large stable plasmids that provide a number of genes believed to enable the bacteria to survive and thrive in the soil environment and often enable the symbiotic relationship with the host plant to occur.

To understand additional mechanisms in which plasmids may increase the fitness of the rhizobia within the rhizosphere and potentially increase the effectiveness of the symbiotic association, plasmid curing experiments enable researchers to compare plasmid cured derivative phenotypes with wild-type growth (Baldani *et al.* 1992, Oresnik *et al.* 1998). Previous experiments showed that strain Rlt101 (Rlt100 cured of plasmid c), lacked at least two genes encoding genes hypothesized to be necessary for L-rhamnose catabolism (Oresnik *et al.* 1998). Mutations introduced into one of these putative rhamnose genes, a dehydrogenase/aldolase, generated a strain that was less competitive for the nodulation of clover when co-inoculated with wild-type strain Rlt100 (Oresnik *et*

al. 1998). Strain Rlt100 was not sequenced and when my research began very little was known about L-rhamnose catabolism in any rhizobia.

To attempt to identify the mechanisms of L-rhamnose catabolism in Rlt100, my thesis objectives were:

1) Delineate the locus for rhamnose catabolism using a phenotypic screening for lack of growth on minimal media supplemented with L-rhamnose.
2) Characterize and fully sequence the entire L-rhamnose catabolic locus and assign function to the putative genes based on biochemical assays and sequence homologies.
3) Investigate the mechanism of L-rhamnose transport within the catabolic pathway.
4) Investigate and quantitate the specificity of this transporter for L-rhamnose.
5) Investigate the mechanism of regulation of the L-rhamnose catabolic locus.
6) Investigate the possibilities of unique genes or unique gene function, compared to the known and characterized *E. coli* L-rhamnose catabolic model, if they were present.

Chapter 2

A Genetic Locus Necessary for L-Rhamnose Uptake and Catabolism in *Rhizobium leguminosarum* bv. *trifolii*

Reproduced from the Journal of Bacteriology, volume 186, Richardson, J. S., M. F. Hynes, and I. J. Oresnik, A genetic locus necessary for rhamnose uptake and catabolism in *Rhizobium leguminosarum* bv. *trifolii*, pp. 8433-8442, Copyright 2004, with permission from the American Society for Microbiology.

2.1 Introduction

Rhizobium leguminosarum bv. *trifolii* is a gram-negative aerobic soil bacterium that can exist as a free-living heterotropic saprophyte, or can form nitrogen fixing nodules on various species of clovers (*Trifolium* spp.). The interaction of rhizobia with their hosts is a sequence of events that begins with an exchange of signals in the rhizosphere, followed by the invasion of the plant through infection threads and ultimately by release of the bacteria into plant cells where they differentiate into nitrogen fixing bacteroids. Through this symbiotic association the plant provides the bacteria with energy for growth and in return the *Rhizobium* provides the plant with fixed nitrogen. Although a great deal has been elucidated regarding the actual steps during nodulation (Long 1996, van Rhijn and Vanderleyden 1995), and the metabolic pathways utilized while in the bacteroid state (Poole and Allaway 2000) comparatively little is known about what influences the growth of the bacteria prior to or during their interaction with plant roots.

The rhizosphere has been described as the region directly under the influence of secreted compounds from the plant root (Bowen 1976). Competition for nodule occupancy exists between strains of *Rhizobium* within the rhizosphere and has been well documented in the literature (Dowling and Broughton 1986, Triplett and Sadowsky 1992). A large number of biotic and abiotic factors influence competition between strains for nodule occupancy. Some of the most well characterized strain-dependent factors include production of antibiotics and bacteriocins (Oresnik *et al.* 1998, Robleto *et al.* 1998, Venter *et al.* 2001), the ability to catabolize specific compounds present in the root environment, as well as the synthesis or utilization of specific vitamins (Phillips *et al.* 1998, Streit *et al.* 1996).

L-Rhamnose is a methyl-pentose sugar and is found as a constituent of pectin in the form of rhamnogalacturonan within the cell walls of dicotyledonous plants (McNeil *et al.* 1984). It has also been found in the mucilage of a number of legume plants (Knee *et al.* 2001). *R. leguminosarum* mutants unable to utilize L-rhamnose as a sole carbon source were found to be significantly impaired in their nodulation competitiveness (Oresnik *et al.* 1998). The reason for the uncompetitive nature of these mutants is not understood. It was hypothesized that if the genetic region responsible for the catabolism of rhamnose were characterized, a better understanding of why these mutants were uncompetitive could be achieved. The objective of this chapter, therefore, was to identify and characterize the genes in this region.

2.2 Materials and methods

2.2.1 Bacterial strains, plasmids, and media

Bacterial strains and plasmids used and generated in this work are listed in Table 2.1. *R. leguminosarum* strains were routinely grown at 30°C on TY as a complex medium (Beringer 1978) and VMM (Vincent 1970) as a defined medium. Glycerol, glucose, or rhamnose, (or a combination of two of these sugars) was used instead of mannitol as carbon sources in VMM, according to the needs of individual experiments. Carbon sources were filter-sterilized and added to defined media to a final concentration of 15 mM. When required, antibiotics were added to solid or liquid media at the following concentrations: tetracycline (Tc), either 10 μg ml^{-1} or 5 μg ml^{-1}; neomycin (Nm), 200 μg ml^{-1}; (Sm), 200 μg ml^{-1}; gentamicin (Gm), either 15 μg ml^{-1} or 30 μg

Table 2.1. Bacterial strains and plasmids

Strain or plasmid	Genotype or phenotype	Reference or source
Strains		
R. leguminosarum		
Rlt100	W14-2, *R. leguminosarum* Sm^r wild-type	(Baldani *et al.* 1992)
Rlt101	W14-2b, cured of plasmid a and d, deletion in plasmid c	(Moenne-Locoz 1995)
Rlt105	Rlt100 *rhaD1*::Tn*5*B20	(Oresnik *et al.* 1998)
Rlt106	Rlt100 *rhaT2*::Tn*5*B20	(Oresnik *et al.* 1998)
Rlt112	Rlt100 *rhaS17*::Tn*phoA*	(Richardson *et al.* 2004)
Rlt115	Rlt100 *rhaP19*::Tn*5*	(Richardson *et al.* 2004)
Rlt117	Rlt100 *rhaR25*::Tn*5*B20	(Richardson *et al.* 2004)
Rlt118	Rlt100 *rhaD26*::Tn*5*B20	(Richardson *et al.* 2004)
Rlt121	Rlt100 *rhaS29*::Tn*5*B20	(Richardson *et al.* 2004)
Rlt124	Rlt100 *rhaI31*::Tn*5*B20	(Richardson *et al.* 2004)
Rlt125	Rlt100 *rhaI32*::Tn*5*B20	(Richardson *et al.* 2004)
Rlt128	Rlt100 *rhaP36*::Tn*5*B20	(Richardson *et al.* 2004)
Rlt130	Rlt100 *rhaI39*::Tn*5*B20	(Richardson *et al.* 2004)
Rlt131	Rlt100 *rhaK40*::Tn*5*B20	(Richardson *et al.* 2004)
Rlt144	Rlt100 *rhaK50*::Tn*5*B20	(Richardson *et al.* 2004)
Rlt146	Rlt100 *rhaK52*::Tn*5*B20	(Richardson *et al.* 2004)
Rlt151	Rlt100 *rhaQ38*::Tn*5*B20	(Richardson *et al.* 2004)
Rlt152	Rlt100, pW3C1	(Richardson *et al.* 2004)
Rlt153	Rlt101, pMR73 (contains *rhaR25*)	(Richardson *et al.* 2004)
Rlt154	Rlt100, pMR73 (contains *rhaR25*)	(Richardson *et al.* 2004)
Rlt197	Rlt117, pMR53 (contains $rhaR^+$)	(Richardson *et al.* 2004)
Rlt211	Rlt105, *rhaK58*::pKNOCK-Tc	(Richardson *et al.* 2004)
Rlt212	Rlt124, *rhaK58*::pKNOCK-Tc	(Richardson *et al.* 2004)
E. coli		
MM294A	*pro-82 thi-1 hsdR17 supE44*	(Finan 1986)
MT607	MM294A *recA56*	(Finan 1986)
EcA101	MT607 + Tn*5*B20 insert in chromosome	(Clark *et al.* 2001)
MT614	MT607 + Tn*5* insert in the chromosome	(Finan 1986)
MT621	MT607 MM294A *malF*::Tn*PhoA*	(Yarosh *et al.* 1989)
MT616	MT607, pRK600	(Finan 1986)
DH5α	*end*A *hsd*R17 *sup*E44 *thi*-1 *rec*A1 *gyr*A96 *rel*AΔ(*arg*F-*lac*ZYA) U169ϕ80d*lac*ZΔM15	B.R.L. Inc.
S17-1	*recA* derivative of MM294A with RP4-2 (Tc::Mu::Km::Tn*7*) integrated into the chromosome	(Simon 1983)
Plasmids		
pUC19	Cloning vector, ColE1 *oriV*, Ap^r	(Yanisch-Perron 1985)
pBluescript II SK	Cloning vector, ColE1 *oriV*, Ap^r	Stratagene
pJQ200 SK^+	Suicide vector for homogenotization; P15a *ori*, *mob*, *sacB*, Gm^r	(Quandt and Hynes 1993)

pKNOCK-Tc	Suicide vector for insertional mutagenesis R6K *ori*, RP4 *oriT*, Tc^r	(Alexeyev 1999)
pRK600	pRK2013 *npt*::Tn*9* Cm^r Nm-Km^s	(Finan 1986)
pRK7813	RK2 *ori*, pUC9 polylinker, λ *cos* site, Tc^r	(Jones and Gutterson 1987)
pPH1JI	IncP plasmid, Gm^r	(Beringer 1978)
pW3A	pRK7813 cosmid from Rlt100 wild-type cosmid bank, complements Rlt101 for rhamnose utilization	(Oresnik *et al.* 1998)
pW3C1	pRK7813 cosmid from Rlt100 cosmid bank, containing entire rhamnose locus	(Oresnik *et al.* 1998)
pMR3	12-kb *Hin*dIII fragment from pW3C1 in pBluescript	(Richardson *et al.* 2004)
pMR5	*Eco*RI deletion of pMR3 in pBluescript	(Richardson *et al.* 2004)
pMR6	1.1 kb *Eco*RI subclone from pMR3 in pBluescript	(Richardson *et al.* 2004)
pMR28	*rhaK* coding sequence in pBluescript	(Richardson *et al.* 2004)
pW3AR1	pW3A *rhaD1*::Tn*5*B20	(Oresnik *et al.* 1998)
pW3AR2	pW3A *rhaT2*::Tn*5*B20	(Oresnik *et al.* 1998)
pW3AR11	pW3A *rhaS11*::Tn*5phoA*	(Richardson *et al.* 2004)
pW3AR44	pW3A *rhaI44*::Tn*5*B20	(Richardson *et al.* 2004)
pW3CR1B	pW3C1 *rhaT*12::Tn*5*	(Oresnik *et al.* 1998)
pW3CR3A	pW3C1 *rhaT*13::Tn*5*	(Oresnik *et al.* 1998)
pW3CR5	pW3C1 *rhaQ*14::Tn*5*	(Oresnik *et al.* 1998)
pW3CR7	pW3C1 *rhaK7*::Tn*5*	(Richardson *et al.* 2004)
pMR53	pRK7813 with $rhaR^+$ expressed from p_{lac}	(Richardson *et al.* 2004)
pMR65	pW3C1 *rhaS17*::Tn*PhoA*	(Richardson *et al.* 2004)
pMR67	pW3C1 *rhaP19*::Tn*5*	(Richardson *et al.* 2004)
pMR70	pW3C1 *rhaK22*::Tn*5*	(Richardson *et al.* 2004)
pMR71	pW3C1 *rhaP23*::Tn*5*	(Richardson *et al.* 2004)
pMR73	pW3C1 *rhaR25*::Tn*5B20*	(Richardson *et al.* 2004)
pMR74	pW3C1 *rhaD26*::Tn*5B20*	(Richardson *et al.* 2004)
pMR77	pW3C1 *rhaS29*::Tn*5B20*	(Richardson *et al.* 2004)
pMR79	pW3C1 *rhaI31*::Tn*5B20*	(Richardson *et al.* 2004)
pMR80	pW3C1 *rhaI32*::Tn*5B20*	(Richardson *et al.* 2004)
pMR84	pW3C1 *rhaP36*::Tn*5B20*	(Richardson *et al.* 2004)
pMR86	pW3C1 *rhaQ38*::Tn*5B20*	(Richardson *et al.* 2004)
pMR87	pW3C1 *rhaI39*::Tn*5B20*	(Richardson *et al.* 2004)
pMR88	pW3C1 *rhaK40*::Tn*5B20*	(Richardson *et al.* 2004)
pMR98	pW3C1 *rhaK50*::Tn*5B20*	(Richardson *et al.* 2004)
pMR100	pW3C1 *rhaK52*::Tn*5B20*	(Richardson *et al.* 2004)
pMR146	pKNOCK-Tc with 465 bp *Eco*RI fragment from *rhaK*	(Richardson *et al.* 2004)

Reproduced from the Journal of Bacteriology, volume 186, Richardson, J. S., M. F. Hynes, and I. J. Oresnik, A genetic locus necessary for rhamnose uptake and catabolism

ml^{-1}; ampicillin (Ap), 100 μg ml^{-1}; kanamycin (Km); 50 μg ml^{-1}. Bacterial growth was monitored spectrophotometrically at 600 nm.

2.2.2 DNA manipulations, Southern blot analysis and constructions

Standard techniques were used for plasmid isolation, restriction enzyme digestion, and agarose gel electrophoresis (Sambrook 1989). For Southern analysis, genomic DNA was isolated using a slightly modified version of the protocol outlined by Meade *et al.* (Meade *et al.* 1982), as has been previously described (Oresnik *et al.* 1994).

Genomic DNA for Southern blot hybridizations was restricted to completion with *Sal*I and electrophoresed through a 0.8% agarose gel using 1X TAE buffer. DNA was transferred onto nylon membranes (Roche Diagnostics, Laval, Quebec) overnight by alkaline transfer (Sambrook 1989), neutralized using 5X SSC and hybridized overnight with digoxygenin labeled probes. Hybridization was done at 68°C overnight and washed 2 x 5 minutes with 2X SSC, 0.1% SDS at room temperature, 2 x 15 minutes at 68°C with 0.1% SSC, 0.1% SDS. Detection was done utilizing CSPD (Roche Diagnostics, Laval, Quebec) reagent as specified by the manufacturer.

Southern blot probes were constructed by digesting a plasmid pW3C1 (isolated from Rlt100 genomic library) with *Eco*R1. The fragments used as probes were labeled using digoxygenin (Roche Diagnostics, Laval, Quebec) as suggested by the manufacturer.

Construction of pMR53 was accomplished in two steps. First, PCR amplification of *rhaR* was achieved by using the primers RhaR5' (5' ATAT**CTGCAG**AAGGAGTTGATCAATGCACGAACGCGAACGC 3') and RhaR3'

(5' ATAT**GAATTC**CCCGGGTCATCGGACCGACGAGGA 3') and pW3C1 as the template. The RhaR 5' primer contained a *Bcl*I site immediately preceding the predicted ATG start site, and a ribosome binding site with optimal spacing from the ATG. The amplification product was cloned into pBluescript II SK using the *Pst*I and *Eco*RI restriction sites that were incorporated into the primers yielding pMR41. This construct was verified by nucleotide sequencing. pMR53 was subsequently generated by subcloning the *Pst*I/*Eco*RI fragment from pMR41 into pRK7813.

Construction of pMR28 was carried out by amplifying the coding region of *rhaK* using the primers RhaK5' (5' ATA**GGATCC**ATGACCGCCAGTTCCTATC 3') and RhaK3' (5' ATA**AAGCTT**CTATGCCATCGCCGCGTA 3') utilizing pW3C1 as a template. The amplification product was cloned as a *Bam*HI/*Hin*dIII fragment into pBluescript II SK by utilizing the restriction sites that were introduced into the primers. The construct was verified by nucleotide sequencing.

2.2.3 Genetic manipulations

Generation and isolation of Tn*5*, Tn*5*B20, and Tn*pho*A inserts was carried out by using strains MT614, EcA101, or MT620 respectively as previously described (Clark *et al.* 2001, Finan 1986, Yarosh *et al.* 1989). Briefly, pW3C1, which had previously been shown to complement Rlt101 for its inability to utilize L-rhamnose as a sole carbon source (Oresnik *et al.* 1998), was introduced into each of these strains and subsequently mobilized into Rlt101. Tc^R and Nm^R transconjugants were screened for the inability to complement. Transposon inserts of interest were subsequently subcloned into pBluescript

II SK such that DNA flanking the insert could be sequenced by utilizing an IS50 primer (5' TAGGAGGTCACATGGAAGTCAGAT 3').

Gene replacement experiments were accomplished either by the use of pJQ200SK (Quandt and Hynes 1993) or by the use of the incompatible plasmid pPH1JI (Beringer 1978) essentially as described by Ruvkun and Ausubel (Ruvkun and Ausubel 1981). Correct replacement of the wild-type gene was confirmed by Southern blot analysis of genomic DNA cut with *Sal*I and probed with pW3C1.

The construction of *rhaD/rhaK* and *rhaI/rhaK* double mutants was accomplished by isolating the internal 465 bp *Eco*RI fragment from *rhaK* and ligating it into pKNOCK-Tc (Alexeyev 1999) yielding pMR146. pMR146 was subsequently conjugated into strains Rlt105 and Rlt124 which contained *rhaD1* and *rhaI31* alleles respectively. Single cross-over of pMR146 in either Rlt105 or Rlt124 were selected on TY agar containing Nm and Tc. A number of colonies were purified and checked by Southern hybridization to ensure correct integration into *rhaK*. Two strains, Rlt211 and Rlt212 containing *rhaD1/rhaK58* and *rhaI31/rhaK58* respectively were subsequently used.

2.2.4 DNA sequencing and sequence analysis

The rhamnose catabolism region was sequenced by a combination of sequencing the DNA flanking Tn*5* inserts in pW3C1 that had lost the ability to confer growth on L-rhamnose in strain Rlt101, as well as primer walking to fill in sequence gaps. Sequencing reactions were done using dye terminators at the University of Calgary Core DNA Services and analyzed on an ABI automated sequencer. Sequence data were analyzed using DNASIS (Hitachi Software Engineering Co., San Bruno, CA. U.S.A.). Database

searches were done using the BLASTX program (Altschul *et al.* 1997). Both strands were completely sequenced and the sequence has been deposited to GenBank as accession number AF085687.

2.2.5 RNA isolation, RNA probe preparation and RNA dot blots

Rapid isolation of RNA was accomplished as previously described (Ausubel 1994) with the following modifications; 500 ml of cells were used for the isolation, an additional PCI (phenol, chloroform, isoamylalcohol 24:25:1) extraction was added, and the final resuspension contained 40 units of RNaseOUT (GiboBRL, Life Technologies). The concentration of RNA for each preparation was determined by assaying an adequate dilution at 260 nm using a Shimadzu MultiSpec-1501. The quality of the prep was also visualized by running the sample on a denaturing agarose gel.

To generate gene specific RNA labeled probes utilizing digoxygenin an *in vitro* labeling kit utilizing either T3 or T7 polymerase was utilized as recommended by the manufacturer (Roche Diagnostics, Laval, QC). Briefly, constructs containing the desired open reading frames were linearized with an appropriate enzyme and RNA was synthesized using either T3 or T7 RNA polymerase. To generate a *rhaR* specific probe, pMR41 was linearized with *Pst*I and T7 RNA polymerase was used. pMR28 was linearized with *Bam*HI and T7 RNA polymerase was used to generate a *rhaK* RNA probe. Probes specific for *rhaS* and *rhaT* were generated by digesting pMR49 and pMR51 with *Sal*I and T3 RNA polymerase was used for both.

Dot blots were carried out by first quantitating the RNA (OD_{260}), adjusting the concentration of the isolated RNA such that all would be at an equivalent concentration

and serially diluting each sample (using diethyl pyrocarbonate (DEPC) treated double distilled H_2O: 20xSSC: formaldehyde in the ratio 5:3:2). These were subsequently boiled for 5 minutes, quenched on ice and spotted on a positively charged nylon membrane (Roche Diagnostics, Laval, Quebec). Samples were applied such that equivalent volumes with identical concentrations of RNA were applied to the membranes. Samples were cross-linked using ultraviolet light and hybridizations were carried out at 68°C overnight. Washes and chemiluminescent detection were carried out as suggested by the manufacturer (Roche Diagnostics, Laval, Quebec). Results were documented by autoradiography as well as chemiluminescent analysis (FluorChem V.2.01, Alpha Innotech Corporation, U.S.A). Signals were quantitated by using the linear portion of the phosphoimager data as well as by using NIH Image Version 1.62 (National Institutes of Health, U.S.A.).

2.2.6 β-Galactosidase Activity

Rhizobium cultures containing active fusions were assayed following an overnight induction with rhamnose. Cultures used for assays were first grown overnight in TY and then used to inoculate a fresh broth culture of defined medium supplemented with 15 mM glycerol and 15 mM rhamnose. Cells were grown overnight in VMM-glycerol/rhamnose broth to mid-log phase (OD 600 nm approx. 0.5) and were assayed as described by Miller (Miller 1972). Cultures were assayed in duplicate and a minimum of 3 independent cultures was used for each experiment. β-galactosidase activity was shown to be linear over the duration of the assay (data not shown).

2.2.7 Rhamnose Transport Assay

Radioactive [^{3}H] rhamnose (5 Ci/mmol) was purchased from American Radiolabeled Chemicals Ltd. St. Louis, Missouri. Strains were grown 20 hours in 5 ml cultures, and then sub-cultured into 50 ml of VMM-glycerol/rhamnose. Cells were harvested by centrifugation (5,000 x g for 10 minutes) at mid-log phase (OD_{600} = 0.5 - 0.7), washed twice in VMM-glycerol, and resuspended in this medium to a final OD_{600} of ~ 0.3 in a total volume of 10 ml. Transport assays were initiated by the addition of tritiated rhamnose to a final concentration of 2 μM (125,000 dpm) 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 the defined salts medium and the radioactivity that remained on the filter was determined using a liquid scintillation spectrophotometer (Beckman LS6500). Samples were taken every 20 seconds and continued for up to two minutes. Transport rates were generally linear over the first minute of the assay. Linearity however did vary slightly between experiments. To determine the K_M of the ABC transporter, 15 and 30 second time points were used to determine initial rates.

2.3 Results

2.3.1 Rhamnose locus contains catabolic and transporter genes

The cosmid pW3C1 was isolated from a cosmid bank constructed from the wild-type Rlt100 on the basis of its ability to complement Rlt101 for its inability to utilize rhamnose as a sole carbon source (Oresnik *et al.* 1998). To delineate the region necessary for the complementation phenotype, random transposon mutagenesis was carried out on

pW3C1 utilizing Tn*5*, Tn*5*B20 and Tn*phoA*, and a total of 56 independent mutants were isolated, mapped and recombined into Rlt100 (Figure 2.1). The entire region on pW3C1 was sequenced on both strands by a combination of sequencing the DNA flanking each of the transposon inserts as well as primer walking. The rhamnose region spans 10,959 bp. The DNA sequence revealed the presence of nine plausible open reading frames (ORFs) apparently arranged as at least two divergently transcribed units. The genes *rhaD* and *rhaI* are transcribed in one direction whereas the genes *rhaR*, *rhaS*, *rhaT*, *rhaP*, *rhaQ*, *rhaU* and *rhaK* are divergently transcribed (Figure 2.1).

The genes *rhaD* and *rhaI* code for a putative oxidoreductase and a putative sugar isomerase respectively. RhaD is predicted to be a 698 amino acid protein that has 87% identity with a similar hypothetical oxidoreductase in *Agrobacterium tumefaciens.* Conserved Domain Architecture Retrieval Tool (CDART) (Geer *et al.* 2002) analysis of the translated peptide revealed that this protein contains two domains, a short chain dehydrogenase domain as well as a type II aldolase domain. A putative NAD^+ binding site was also predicted to exist.

A gap of 193 bp was found between the *rhaD* and *rhaI* genes. The RhaI protein is predicted to be a 430 amino acid sugar isomerase showing 84% identity to the predicted product of a similar gene from *A. tumefaciens*. Transposon inserts yielding mutants unable to use L-rhamnose were not found beyond *rhaI* and sequencing beyond *rhaI* did not reveal any further open reading frames which might be associated with rhamnose utilization.

Genes *rhaRSTPQUK* were found to be oriented divergently from *rhaDI* (Figure 2.1). The *rhaR* gene encodes a 270 amino acid protein that displays 79% identity to a

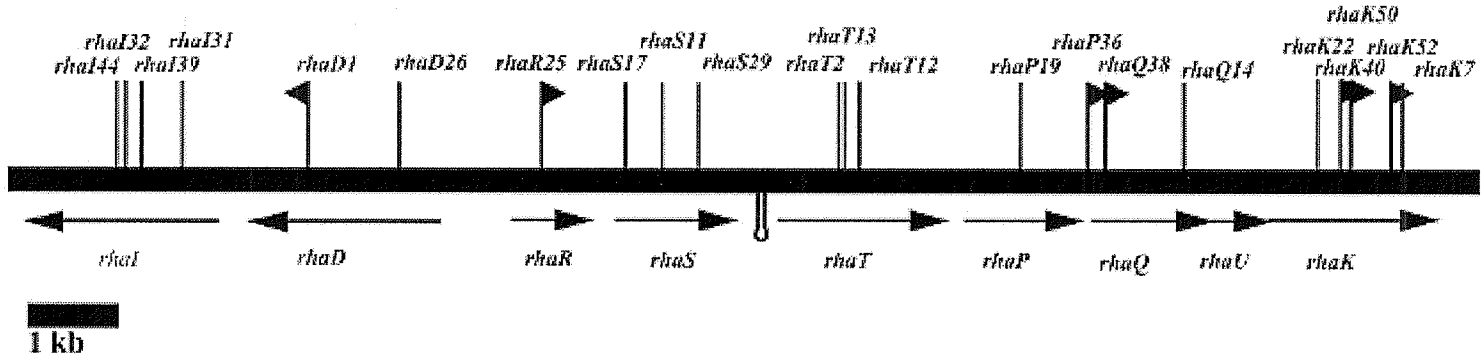


Figure 2.1. Rhamnose region of *R. leguminosarum* bv. *trifolii* strain Rlt100. Heavy solid horizontal line, Rlt100 genomic DNA; lines with arrows, ORFs; vertical lines, transposon inserts; vertical lines with flags, Tn5B20 inserts that yielded lacZ fusions; stem-loop, position of a transcriptional attenuator. Gene names are under each respective ORF. The sequence data for the region were submitted to the GenBank database and assigned accession number AF085687. Reproduced from the Journal of Bacteriology, volume 186, Richardson, J. S., M. F. Hynes, and I. J. Oresnik, A genetic locus necessary for rhamnose uptake and catabolism in *Rhizobium leguminosarum* bv. *trifolii*, pp. 8433-8442, Copyright 2004, with permission from the American Society for Microbiology.

DeoR-type negative regulator from *Agrobacterium tumefaciens*, and 68% identity to *Mesorhizobium loti*, and *Sinorhizobium meliloti* DeoR-type negative regulators. Following the regulator gene, separated by 46 bp, is *rhaS*. The gene *rhaS* encodes a potential periplasmic sugar binding protein. The genes *rhaTPQ* encode an ABC-type ATPase (*rhaT*) followed by two membrane spanning transmembrane proteins (*rhaPQ*) respectively. The genes *rhaS* and *rhaT* are separated by 73 bp, whereas the intergenic spaces between *rhaT*/*rhaP* and *rhaP*/*rhaQ* are 9 and 36 bp respectively. We note that the intergenic region between *rhaS* and *rhaT* appears to have a palindromic sequence that has the potential to form a hairpin loop consisting of a 10 bp GC-rich stem with a 4 bp loop. This structure has a predicted ΔG° value of -24.4 kcal/mol and a predicted Tm of 126°C.

The genes *rhaU* and *rhaK* follow *rhaQ*. The start site of *rhaU* is predicted to overlap the stop codon of *rhaQ* (ATGA). BLASTX analysis of RhaU shows that this protein is similar to hypothetically conserved proteins that are found in a wide variety of Eubacteria. These genes are all classified as being similar to *yulD* from *Bacillus subtilis*. Interestingly, the gene *yulD* is found in a cluster of genes that resemble the rhamnose catabolic operon in *E. coli*. There is no ascribed function to the gene product of *yulD* in any organism.

The final gene in this region is *rhaK*, predicted to encode a 460 amino acid protein that is 74% similar to a putative rhamnose kinase from *A. tumefaciens*. The *rhaK* predicted open reading frame overlaps the stop codon of *rhaU* (ATGA). No transposon inserts were isolated beyond *rhaK* that led to an inability of pW3C1 to complement Rlt101. BLASTX analysis of the nucleotide sequence following *rhaK* did not reveal any

significant homologies suggesting that *rhaK* was on a terminal end of the rhamnose complementing region.

It is of interest that there are clusters of genes in both *A. tumefaciens* (Wood *et al.* 2001) and *S. meliloti* (Galibert *et al.* 2001) that appear to correspond very closely to the rhamnose utilization gene cluster we have identified here. The order of the genes is precisely conserved, as is the presence of the stem-loop structure between *rhaS* and *rhaT*. However, the genetic context of the gene clusters is highly different in the three bacteria. In *A. tumefaciens* C58 they are located on the linear chromosome (Atu3484-Atu3492) and flanked by *panCB* and a predicted non-heme haloperoxidase, whereas in *S. meliloti* 1021 they are chromosomally encoded (Smc02321-Smc03003) and located immediately adjacent to the chemotaxis operon. This is in contrast to both strains of *R. leguminosarum,* Rlt100 and Rlv3841, in which the genes are located on a large plasmid (Oresnik *et al.* 1998). A comparison with the *R. leguminosarum* Rlv3841 genomic sequence (http://www.sanger.ac.uk/Projects/R_leguminosarum/), reveals a similar organization of the *rha* genes to that reported here, and the predicted proteins encoded by the strain Rlv3841 sequence are 88-96% identical to those from Rlt100. It is of note that the genes are found in an entirely different context in these three related members of the *Rhizobiaceae* family.

2.3.2 Rhamnose region contains two complementation groups

Sequence analysis of the rhamnose region revealed several intergenic gaps that could contain putative promoters. To address the possibility that multiple transcripts were produced in this region, representative insertion mutants for each gene (carried on either

pW3C1 or pW3A) were chosen and a complementation analysis was carried out (Table 2.2). The results clearly show that although an intergenic gap of 193 bp exists between *rhaD* and *rhaI* these two genes are in one transcriptional unit. Similarly the 46 bp and 76 bp integenic regions that exist between *rhaR*/*rhaS* and *rhaS*/*rhaT* respectively do not contain independent promoters and all the genes transcribed divergently from *rhaDI* appear to constitute a single transcript (Table 2.2).

2.3.3 Regulation of rhaRSTPQUK

Initial regulatory experiments have been previously carried out utilizing the *rhaD1* allele either on a plasmid or as a single copy chromosomal recombinant (Oresnik *et al.* 1998). Those data showed that *rhaD1* was induced by L-rhamnose and clover root exudate but repressed by glucose (Oresnik *et al.* 1998). Since the *rhaRSTPQUK* transcript contained a putative regulatory gene and components that could be used for the transport and catabolism of L-rhamnose, it was of interest to examine the expression of this region using representative Tn*5*B20 fusions that were spread out across the transcriptional unit (Table 2.3).

All the strains carrying fusions in this unit showed higher activity when grown in the presence of L-rhamnose than in the presence of glucose (Table 2.3). Strains Rlt128 and Rlt151 containing *rhaP36*::Tn*5*B20 and *rhaQ38*::Tn*5*B20 showed an induction of 28 and 13 fold respectively. The results suggested that transcription of the operon was induced by the presence of rhamnose and that *rhaPQ* were transcribed at greatly reduced levels under non-inducing conditions.

Table 2.2 Complementation of L-rhamnose mutants

Strain	Relevant genotype	Gene mutated	Growth with plasmid						
			pW3AR1[b] *rhaD1* dehydrogenase	pMR79 *rhaI31* isomerase	pMR73 *rhaR25* regulator	pMR65 *rhaS17* PBP	pW3AR2 *rhaT2* ABC	pMR67 *rhaP19* permease	pMR98 *rhaK50* kinase
Rlt100	wild-type	N/A[a]	+	+	+	+	+	+	+
Rlt105	*rhaD1*	dehydrogenase	-	-	+	+	+	+	+
Rlt124	*rhaI31*	isomerase	-	-	+	+	+	+	+
Rlt117	*rhaR25*	regulator	+	+	-	-	-	-	-
Rlt112	*rhaS17*	binding protein	+	+	-	-	-	-	-
Rlt106	*rhaT2*	ABC protein	+	+	-	-	-	-	-
Rlt115	*rhaP19*	permease	+	+	-	-	-	-	-
Rlt144	*rhaK50*	kinase	+	+	-	-	-	-	-

Designations are as follows: +, like wild-type growth; -, no growth.
[a] N/A= not applicable
[b] dehydrogenase = oxidoreductase (dehydrogenase/aldolase)
Reproduced from the Journal of Bacteriology, volume 186, Richardson, J. S., M. F. Hynes, and I. J. Oresnik, A genetic locus necessary for rhamnose uptake and catabolism in *Rhizobium leguminosarum* bv. *trifolii*, pp. 8433-8442, Copyright 2004, with permission from the American Society for Microbiology..

Table 2.3. Effect of L-rhamnose and glucose on β-galactosidase fusions to genes in the L-rhamnose transport operon.

Strain	Relevant Genotype	Glucose[a]	Rhamnose	Induction[b]
Rlt117	*rhaR25*::Tn*5*B20	2824 (185)[c]	10,451 (751)[d]	3.7
Rlt128	*rhaP36*::Tn*5*B20	24 (1)	660 (46)[e]	28
Rlt151	*rhaQ38*::Tn*5*B20	33 (4)	424 (41)	13
Rlt131	*rhaK40*::Tn*5*B20	2 (1)	5 (0)	2
Rlt144	*rhaK50*::Tn*5*B20	2 (1)	11 (2)[e]	5
Rlt146	*rhaK52*::Tn*5*B20	4 (2)	12 (1)[e]	3

[a] The values given represent β-galactosidase activity expressed in Miller units following overnight growth in defined media with either glucose/glycerol or rhamnose/glycerol as carbon sources. Values are the average of at least 4 independent replicates; values in parentheses represent standard error.
[b] Ratio of rhamnose values to the glucose values.
[c] Value represents seven independent replicates.
[d] Value represents six independent replicates.
[e] Value represents three independent replicates.
Reproduced from the Journal of Bacteriology, volume 186, Richardson, J. S., M. F. Hynes, and I. J. Oresnik, A genetic locus necessary for rhamnose uptake and catabolism in *Rhizobium leguminosarum* bv. *trifolii*, pp. 8433-8442, Copyright 2004, with permission from the American Society for Microbiology.

Fusions to *rhaK* showed low levels of β-galactosidase activity. Initially Rlt146 carrying *rhaK52*::Tn*5*B20 was assayed and showed a low induction. To ensure that induction was indeed occurring two additional *rhaK* alleles, *rhaK40*::Tn*5*B20 and *rhaK50*::Tn*5*B20 in strains Rlt131 and Rlt144 were also assayed. Consistent with the results using Rlt146, these fusions also showed reproducible low fusion activity that showed between a 2-5 fold increase of β-galactosidase activity when grown in the presence of rhamnose (Table 2.3). Low β-galactosidase activity was also observed when these alleles were present on multicopy plasmid pW3C1 in an Rlt100 background (data not shown).

In contrast to the activity found with fusions to either of the permease genes or the kinase gene, *rhaR25*::Tn*5*B20 in strain Rlt117 had very high activity under non-inducing conditions (Table 2.3). This fusion, however, also showed induction in the presence of rhamnose. The *rhaR25* fusion showed the same high activity when grown on glucose regardless of whether it was integrated into the chromosome or present on a multicopy plasmid (Table 2.3, Table 2.4).

2.3.4 RhaR is a negative regulator

Sequence analysis of *rhaR* suggested that this gene encodes a DeoR-type negative regulator (Mortensen *et al.* 1989). To show that RhaR acts as a negative regulator it was reasoned that operon basal activity should increase if a strain lacking *rhaR* was constructed. To test this hypothesis, pMR73 carrying *rhaR25*::Tn*5*B20 was mobilized into both the wild-type Rlt100 as well as Rlt101 which is missing the entire rhamnose region. The results show that in the absence of a functional copy of RhaR, fusion activity,

Table 2.4. Effect of RhaR on *rhaR* expression.

Strain	Relevant Genotype Chromosomal/plasmid	Glucose[a]	Rhamnose	Induction[b]
Rlt100	wild-type	17(2)	15(2)	1
Rlt101	*Δrha*/-	28(1)	21(1)	1
Rlt154[c]	wild-type/*rhaR25*	1,108(84)	2,408(210)	2.2
Rlt153[d]	Δrha/*rhaR25*	2,650(44)	7,029(288)	2.7
Rlt117	*rhaR25*/-	2,624(70)	7,780(255)	3.0
Rlt197[e]	*rhaR25*/*rhaR*$^+$	155(13)	600(32)	3.9

[a] The values given represent β-galactosidase activity expressed in Miller units following overnight growth in defined media with either glucose/glycerol or rhamnose/glycerol as carbon sources. Values are the average of at least four independent replicates except Rlt117 which is an average of three independent replicates; values in parenthesis represent standard error.

[b] Induction was calculated by dividing activity in the presence of rhamnose by the activity in the presence of glucose.

[c] Rlt154 = Rlt100(pMR73).

[d] Rlt153 = Rlt101(pMR73).

[e] Rlt197= Rlt117(pMR53).

as measured by *rhaR25* increased 2.4 fold when the strain was grown on defined media with glucose (Rlt153 vs. Rlt154, Table 2.4). This is comparable to the level of induction seen when the wild-type Rlt100 carrying the fusion *rhaR25* is grown in the presence of rhamnose (Table 2.4). The fusion activity of *rhaR25* in a strain devoid of *rhaR* is still inducible and in fact appears to induce to higher levels (Table 2.4). These data are also consistent with the levels observed for Rlt117 that is also lacking functional RhaR (Table 2.3, Table 2.4).

To provide additional evidence that RhaR is a negative regulator, it was reasoned that if RhaR were over-expressed in a *rhaR25* background, the overall transcriptional activity should be reduced. To test this, the entire coding region of *rhaR* was PCR amplified, and cloned into pRK7813 such that it would be transcribed constitutively from the p_{lac} present as part of the multiple cloning site. To ensure expression a Shine-Dalgarno sequence was also added. The introduction of this construct, pMR53, into Rlt117 reduced the expression of a chromosomal copy of *rhaR25* by approximately 17 fold when grown on glucose (Rlt117 vs. Rlt197, Table 2.4). However, it was observed that even with the over-expression of RhaR in Rlt197, the *rhaR25* fusion was still inducible by rhamnose. Together these data suggest that the induction of this operon is independent of the negative regulation provided by RhaR (Table 2.4).

2.3.5 Transcription extends past rhaR but not beyond rhaS under non-inducing conditions

Analysis of the transcriptional fusion data suggested that the transcriptional pattern of *rhaR* was very different from that of the rest of the transcript (Table 2.3).

However, mutations in *rhaR* were shown to be polar on the *rhaSTPQUK* (Table 2.2). This suggests that a level of control must exist to allow differential transcription of the promoter proximal genes relative to distal genes in this operon (Figure 2.1). It was hypothesized that the putative hairpin loop that had a similarity to a transcriptional terminator might play a role in the regulation of this operon by preventing high levels of transcription beyond *rhaS* when the bacteria are grown under non-inducing conditions.

To provide direct physical evidence to support this hypothesis, transcriptional analysis was carried out using RNA dot blot analysis. Individual RNA probes corresponding to *rhaR*, *rhaS*, *rhaT* and *rhaK* were generated and used to probe RNA isolated from both Rlt100 or Rlt100(pW3C1) grown on defined media with either glycerol/glucose or glycerol/rhamnose. The data from these experiments showed that the steady state mRNA ratios under non-inducing conditions when compared to full induction for *rhaR*, *rhaS*, *rhaT* and *rhaK* were 0.24±0.03, 0.13±0.02, 0.03±0.01, and 0.04±0.01 respectively. These data are consistent with the hypothesis that the hairpin loop found in the intergenic region between *rhaS* and *rhaT* plays a role in attenuating further transcription into the operon (Figure 2.2). Moreover, RNA isolated from Rlt117 (containing *rhaR25*) grown in either the presence of glucose or rhamnose and probed with either *rhaS*, *rhaT*, or *rhaK* did not show any hybridization (data not shown). This also provides further evidence that all genes in this transcript are transcribed from a single promoter.

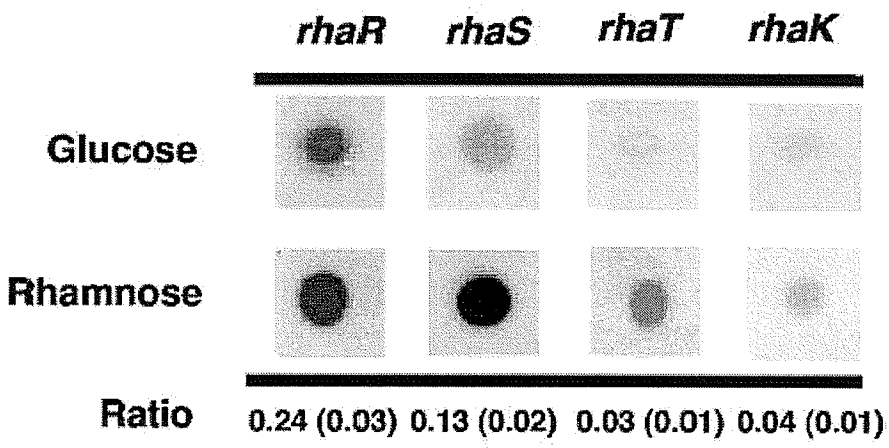


Figure 2.2. RNA Dot blot experiments. Equal volumes and concentrations of total RNA (6 μg) from Rlt100 cultures grown in defined media with either glycerol/glucose or glycerol/rhamnose as carbon sources were spotted onto nylon charged membranes and probed with *rhaR*, *rhaS*, *rhaT*, and *rhaK* specific probes as described in material and methods. The figure shows the results of one representative experiment. Ratios represent the quantitated glucose values divided by the quantitated rhamnose values derived from three independent experiments. Values in parenthesis represent the standard deviation. Reproduced from the Journal of Bacteriology, volume 186, Richardson, J. S., M. F. Hynes, and I. J. Oresnik, A genetic locus necessary for rhamnose uptake and catabolism in *Rhizobium leguminosarum* bv. *trifolii*, pp. 8433-8442, Copyright 2004, with permission from the American Society for Microbiology.

2.3.6 Transport of rhamnose is inducible and is carried out by the ABC transporter

Sequence analysis suggested that *rhaSTPQ* might encode an ABC-type transporter. To show functionality, transport assays utilizing tritiated rhamnose were carried out. Strain Rlt100, when grown in the presence of rhamnose, was able to take up tritiated rhamnose from the culture medium (Figure 2.3A). However, when Rlt100 was grown in the presence of glucose the amount of tritiated rhamnose uptake was not above background levels.

To further characterize the rhamnose transport system experiments were carried out to determine the K_M for rhamnose transport. The K_M for rhamnose uptake in Rlt100 was found to be approximately 13 ± 2 μM ($n = 4$, data not shown).

Naturally occurring sugars are generally found in a D-configuration. The exceptions to this are the naturally occurring sugars L-rhamnose, L-fucose (which is also a methyl-pentose), and L-arabinose. These sugars are also aldoses that are found as constituents of plant cell walls (McNeil *et al.* 1984). To determine specificity of the rhamnose transporter, a series of inhibition experiments was carried out using tritiated rhamnose and an excess of unlabelled rhamnose, fucose or arabinose. It was determined that the transport of tritiated rhamnose was not inhibited by either L-fucose or L-arabinose even if these were present at a 100 fold excess with respect to the tritiated rhamnose. The transport of tritiated rhamnose; however, was reduced to 12% with a 7.5 fold excess of unlabeled rhamnose and completely inhibited at 75 fold excess (data not shown).

To determine if the proteins encoded by *rhaSTPQ* were responsible for the ability of Rlt100 to transport rhamnose, strains Rlt112, Rlt106, Rlt128, and Rlt151,

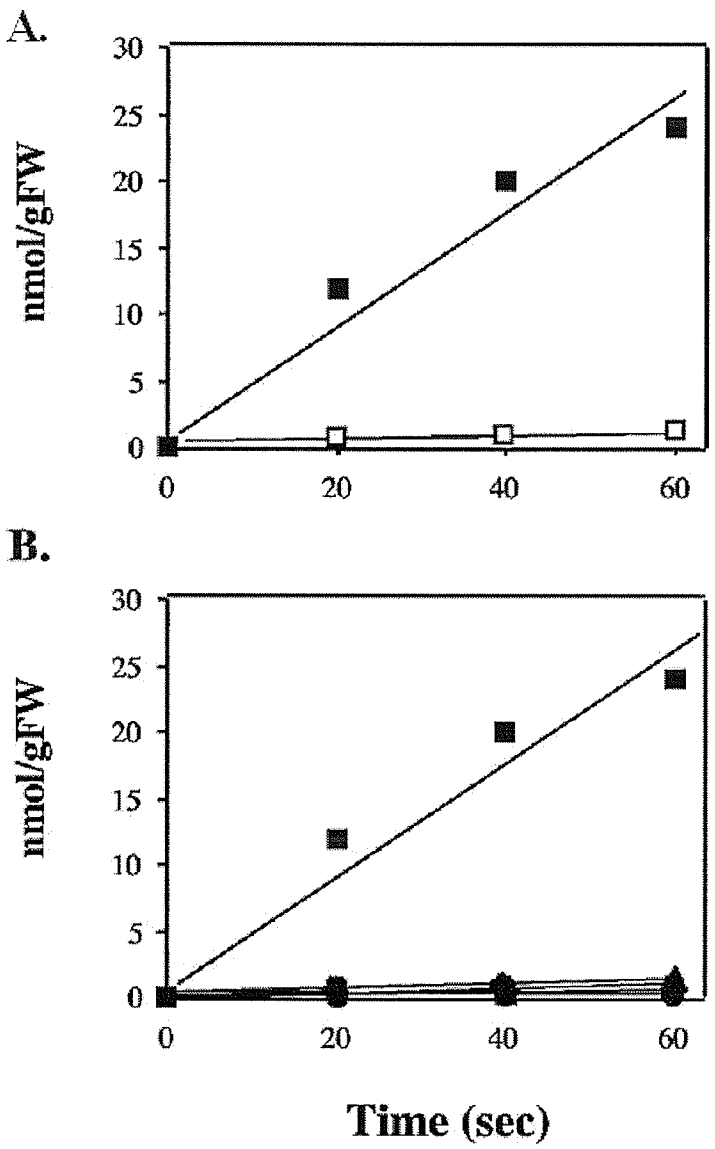


Figure 2.3. Rhamnose transport experiments. Strains were grown to mid-log phase in defined media containing either glucose/glycerol (open symbols) or rhamnose/glycerol (closed symbols) and initial uptake rates were determined using [^{3}H] rhamnose. Symbols: squares, Rlt100 (wild-type); diamonds, Rlt117 (*rhaR25*); triangles, Rlt112 (*rhaS17*); circles, Rlt106 (*rhaT2*); +, Rlt151 (*rhaQ38*); X, Rlt128 (*rhaP36*). **A.** Comparison of rhamnose uptake by Rlt100 grown on rhamnose/glycerol and glucose/glycerol. **B.** Comparison of uptake of induced wild-type and representative mutants. The data represents the average of three independent experiments. Error bars are smaller than the symbols in all cases. Reproduced from the Journal of Bacteriology, volume 186, Richardson, J. S., M. F. Hynes, and I. J. Oresnik, A genetic locus necessary for rhamnose uptake and catabolism in *Rhizobium leguminosarum* bv. *trifolii*, pp. 8433-8442, Copyright 2004, with permission from the American Society for Microbiology.

corresponding to mutations in *rhaS*, *rhaT*, *rhaP*, and *rhaQ* respectively, were tested for their ability to transport rhamnose. In all cases the mutants were unable to take up the labeled rhamnose from the medium (Figure 2.3B).

Additionally, Rlt117 as well as Rlt105 were also assayed for their ability to transport rhamnose. It was found that Rlt117, which has an insertion upstream of *rhaSTPQ*, was unable to transport rhamnose (Figure 2.3B) whereas Rlt105, which carries a mutation in the *rhaDI* transcript, had transport rates similar to the wild-type Rlt100 (data not shown).

2.3.7 Catabolism of L-rhamnose follows a novel pathway

The catabolism of L-rhamnose has been studied at both the biochemical and genetic level in *E. coli*. The pathway has been shown to proceed by an isomerization of rhamnose to yield rhamnulose, then a phosphorylation to give rhamnulose-1-phosphate, followed by an aldolase reaction which cleaves the rhamnulose-1-phosphate to give dihydroxy-acetone phosphate and lactaldehyde (Power 1967, Sawada and Takagi 1964, Takagi and Sawada 1964a, Takagi and Sawada 1964b).

In *E. coli* a biochemical lesion in a catabolic pathway following the synthesis of a phosphorylated sugar intermediate can give rise to a conditional lethal phenotype if the bacterium is grown on a medium that contains a non-inducing carbon source that can be catabolized. This particular phenomenon was first shown with respect to galactose utilization in *E. coli*. Galactose utilization is carried out by the subsequent actions of three enzymes: GalK (galactose kinase), GalT (galactose-1-phosphate uridyl transferase), and GalE (UDP-galactose 4-epimerase) that convert galactose first to galactose-1-phosphate,

then to UDP-galactose, and finally to glucose-1-phosphate. Whereas *galK* mutants are able to grow in the presence of glycerol and galactose, *galE* and *galT* mutants are sensitive to galactose in the presence of glycerol. This sensitivity is correlated with the accumulation of intermediates of this pathway: galactose-1-phosphate and/or UDP-galactose (Adhya and Shapiro 1969, Sundararajan *et al.* 1962). An analogous growth inhibition and accumulation of rhamnulose-1-phosphate has also been demonstrated to take place in the *E. coli* rhamnose catabolic pathway (Power 1967).

Since the gene order as well as the predicted functions of the open reading frames in Rlt100 differed from those of the *E. coli* rhamnose operon, we wished to determine whether the catabolism of this methyl-pentose followed the same biochemical pathway. To determine this, *R. leguminosarum rhaI*, *rhaD*, and *rhaK* mutants were tested for their ability to grow on rhamnose, rhamnose/glycerol and glycerol defined media. The results show, that as expected, *rhaK* mutants were able to grow on glycerol/rhamnose medium (Table 2.5). However, *rhaD* and *rhaI* mutants were unable to grow on glycerol/rhamnose (Table 2.5). Taken together, these data suggest that unlike *E. coli*, where the first step of the catabolic pathway is an isomerization, the first step in the catabolism of rhamnose in *R. leguminosarum* strain Rlt100 is a phosphorylation.

It was hypothesized that if RhaK was indeed the first step of rhamnose catabolism, then either a *rhaD/rhaK* or a *rhaI/rhaK* double mutant should have the same phenotype on glycerol/rhamnose media as a *rhaK* mutant. To test this, Rlt211 containing *rhaD1/rhaK58* and Rlt212 containing *rhaI31/rhaK58* were constructed. When these strains were tested on glycerol/rhamnose media, the results showed that both of these

Table 2.5. Growth of representative mutant *Rhizobium* strains on glycerol, rhamnose, and glycerol/rhamnose as the carbon sources.[a]

Strain	Relevant Genotype	Location of mutation(s)	Gly[bc]	Gly/Rha	Rha
Rlt100	wild-type	N/A	+	+	+
Rlt106	*rhaT*2	ABC	+	+	-
Rlt117	*rhaR*25	regulator	+	+	-
Rlt144	*rhaK*50	kinase	+	+	-
Rlt151	*rhaQ*38	permease	+	+	-
Rlt105	*rhaD*1	dehydrogenase [d]	+	-	-
Rlt124	*rhaI31*	isomerase	+	-	-
Rlt211	*rhaD1,rhaK58*	dehydrogenase, kinase	+	+	-
Rlt212	*rhaI31,rhaK58*	isomerase, kinase	+	+	-

[a] Designations are as follows: +, like wild-type growth; -, no growth.
[b] 15 mM of each carbon source added.
[c] Gly, glycerol; Gly/Rha, glycerol/rhamnose; Rha, rhamnose.
[d] dehydrogenase = oxidoreductase (dehydrogenase/aldolase)

strains were able to grow (Table 2.5). These data are consistent with the hypothesis that in *R. leguminosarum*, RhaK and not RhaI is the first biochemical step in the catabolism of L-rhamnose.

2.4 Discussion

Genes involved in the transport and catabolism of L-rhamnose in *R. leguminosarum* bv. *trifolii* strain Rlt100 are plasmid encoded and are necessary for the bacterium to be competitive for nodule occupancy (Oresnik *et al.* 1998). Sequence analysis of this region was consistent with the hypothesis that the locus contained two transcripts, *rhaDI* and *rhaRSTPQUK*. The analysis was ambiguous, however, because intergenic regions that could contain promoters existed within the putative *rhaRSTPQUK* transcript (Figure 2.1), and β-galactosidase fusion analysis showed differential expression (Table 2.3). It was therefore of interest to determine whether these genes were in one transcript. Based on the isolation of transposon inserts in a negative regulator RhaR (Figure 2.1), the inability of Rlt117 which carries the polar *rhaR25* allele to transport rhamnose (Figure 2.3), the complementation data (Table 2.2), and the lack of detectable mRNA for genes distal to *rhaR25*, we conclude that *rhaRSTPQUK* is transcribed as a single transcript.

The transcription of the *rha::lacZ* fusions in the *rhaRSTPQUK* operon can be roughly divided into three groups: the activity exhibited by *rhaR25*::Tn*5*B20 (Table 2.3, Table 2.4), the activity exhibited by *rhaP36*::Tn*5*B20 and *rhaQ38*::Tn*5*B20, and that of fusions in *rhaK* (Table 2.3). Stoichiometric balance between components of an ATP binding cassette transport system is often necessary for correct functioning. As such, it is

not unexpected to find ABC transporter operons organized such that components that need to be found in higher concentrations are encoded by genes which are found at the 5' end of an operon (Boos 1996, di Guan *et al.* 1988).

The *in vivo* activity of RhaR is consistent with its sequence analysis as a DeoR-type negative regulator (Table 2.4). Transcriptional activity of the *rhaR25* fusion responds to the absence or presence of RhaR with corresponding increases or decreases in assayed activity. It is of interest to note that although the basal activity of a strain carrying the *rhaR25* fusion grown on glucose can be manipulated to increase or decrease with respect to the absence or presence of RhaR, this fusion is still inducible by rhamnose (Table 2.4). The complete absence of RhaR leads to a large increase in the rate of transcription of *rhaR* as measured by the *rhaR25* fusion allele (Table 2.3, Table 2.4). This argues that the regulation of the *rhaR* transcript is under two levels of control, a negative regulator, RhaR, as well as another regulator that is not encoded on genes found within this region. It is also of note that induction in Rlt117 is occurring even though we are unable to detect the transport of L-rhamnose into the cell (Figure 2.3).

The presence of hairpin attenuators has been previously described in a number of operons (Bailey *et al.* 1997, Rutberg 1997). RNA dot blot experiments were carried out to address whether the hairpin-like sequence that is found in the intergenic region between *rhaS* and *rhaT* was correlated with the lack of activity of fusions found distal to *rhaS* when grown in the presence of glucose (Table 2.3). The data clearly show that transcription drops off dramatically between *rhaS* and *rhaT* (Figure 2.2). The mechanism(s) by which anti-termination might occur under inducing conditions in the *rhaRSTPQUK* operon are at present unknown.

The RNA dot blot data also show that the gene for the sugar binding component, *rhaS*, is also expressed when Rlt100 is grown in the presence of glucose. Assuming transcription is proportional to translation this would indicate that under non-inducing conditions RhaS is present in the periplasm. The presence of the periplasmic binding component under non-inducing conditions suggests that this protein may play a role in addition to transport. Sugar-binding proteins have been shown to participate in signal transduction cascades in chemotaxis (Gardina *et al.* 1992, Kemner *et al.* 1997, Mowbray and Koshland 1987, Van Bastelaere *et al.* 1999, Zhang *et al.* 1992) and the in activation of other gene regulation pathways (Peng *et al.* 1998) in addition to being substrate recognition proteins for ABC transporters. Tests using a previously described swarm plate assay (Yost *et al.* 1998) to evaluate whether strains carrying a *rhaS* mutation had altered chemotaxis toward rhamnose found no difference between these strains and the wild-type (data not shown).

The β-galactosidase fusion data from three independent *rhaK* alleles, either integrated into the chromosome (Table 2.3), or as plasmids in a wild-type background, yielded values that were consistently lower than any other alleles in the operon. Dot blot data, however, clearly show that induction of *rhaK* occurs when grown in the presence of rhamnose (Figure 2.2). The data are consistent with a lower rate of transcription through *rhaK*, which in turn may play a role in limiting the amount of RhaK that is produced. It is possible that RhaK is a very stable, or a very active protein, and therefore the gene does not need to be transcribed at the same level as the transporter. We have, however, shown that strains carrying *rhaD* and *rhaI* mutations have a sensitivity phenotype if they are grown on rhamnose/glycerol that is *rhaK* dependent (Table 2.5). The phenotype is

presumably due to the build up of a phosphorylated intermediate (Adhya and Shapiro 1969, Power 1967, Sundararajan *et al.* 1962). Since it is highly unlikely that the bacterium would ever encounter a situation where it would entirely depend upon L-rhamnose as its sole source of carbon in nature, RhaK may need to be limiting to ensure that the metabolic pool of the phosphorylated intermediate is kept at a sub-inhibitory concentration.

Transport experiments using tritiated rhamnose show that the wild-type strain Rlt100 does not take up L-rhamnose when grown in the presence of glucose and that transport is inducible by the presence of L-rhamnose. The inducible nature of the transporter is consistent with the regulatory studies carried out with the genetic fusions as well as the RNA dot blots. The inability of representative mutants in the regulator and the components of the ABC transporter to take up labeled L-rhamnose demonstrate that the operon containing the ABC transporter is necessary for the transport of rhamnose and that a secondary transporter for this sugar is not present or is not expressed under the growth conditions being utilized. Mutants carrying lesions in *rhaD* or *rhaI* were able to transport rhamnose although they were unable to grow on the sugar as a sole carbon source (data not shown).

The biochemical pathway for the catabolism of methyl-pentose sugars is relatively well understood in enteric bacteria and this model has been assumed to be true for all organisms. The simple phenotypic testing of representative mutants in Rlt100 unable to grow using L-rhamnose as a sole carbon compound suggests that the pathway order, and hence the order of the biochemical reactions, may not be conserved. The simplest interpretation of our data suggests that following transport of L-rhamnose into

the cell the first catabolic step is the phosphorylation of L-rhamnose and not an isomerization to yield a keto-sugar as seen in *E. coli*. Although the position of a gene does not imply the order of a biochemical pathway or its function, it may be a noteworthy observation that the kinase is found in the same transcript as the sugar transporter suggesting that the primary catabolic step may work in concert with the transport of the sugar.

Since the pathway appears to be initiated with the phosphorylation of L-rhamnose rather than rhamnulose, the presumed substrates of each of the subsequent enzymatic steps are currently unavailable commercially making it difficult to demonstrate biochemical activity for either the isomerase or the dehydrogenase/aldolase. Nevertheless, we attempted to develop a coupled *in vitro* assay to demonstrate dehydrogenase/aldolase activity utilizing a rhamnose induced cell free extract, ATP, rhamnose, and $NAD(P)^+$. The goal of such an assay would be to test strains carrying either a dehydrogenase/aldolase or an isomerase mutation and thus be able to order the pathway. We were unsuccessful in our attempts. Presumably either the quantity of the product from either the kinase and/or the isomerase reaction may have been insufficient to act as a substrate for the dehydrogenase reaction or a necessary co-factor is absent. Our work in this area is ongoing.

In this report we have characterized a locus necessary for L-rhamnose catabolism and determined that in *R. leguminosarum* bv. *trifolii* rhamnose transport is inducible and dependent on an ABC-type transporter that is regulated by a DeoR-type negative regulator. As well, a level of transcriptional regulation may be due to a hairpin loop located in the *rhaRSTPQUK* transcript that may aid in expression of genes needed for

uptake and catabolism at levels proportional to the amount of each individual protein required. Additionally, we have shown that the catabolism of L-rhamnose follows a novel pathway that to our knowledge has not been previously described. We are currently focusing our work on understanding the molecular aspects of L-rhamnose catabolism and regulation in *R. leguminosarum*. By understanding these fundamental aspects it is hoped that we can achieve a better understanding of how these mutants grow in the rhizosphere and why their inability to catabolize L-rhamnose leads to a non-competitive nodulation phenotype.

Chapter 3

Rhamnose Transport is Sugar-Kinase (RhaK) Dependent in *Rhizobium leguminosarum* bv. *trifolii*

3.1 Introduction

ATP binding cassette (ABC) transporters mediate the transport of molecules across cellular membranes using the free energy derived from the binding and hydrolysis of ATP they are ubiquitous within archaea, bacteria and eukaryotes (Dassa 2003, Higgins 1992). In humans, mutations within ABC transporter proteins lead to many genetic disorders and are also responsible for anticancer drug resistance (Bates 2003). In other organisms they are involved in cellular processes such as protection of cells from cytotoxins, signal transduction, protein secretion, antigen presentation, bacterial pathogenesis, sporulation, antibiotic, antifungal, and herbicide resistance, to name a few (Altenberg 2003, Hanrahan *et al.* 2003, Higgins and Linton 2004, Konig *et al.* 1999, Poelarends 2002, Rea 1998, Schneider and Hunke 1998).

Three classes of ABC systems are currently recognized. The first class are exporters, the second class have systems lacking transmembrane domains and are involved in non-transport cellular processes and antibiotic resistance, and the third class are importers (Allikmets *et al.* 1997, Dassa 2003, Dassa and Bouige 2001, Saurin *et al.* 1999).

ABC transporters involved in sugar uptake belong to the third class of the ABC family. These transporters are unique to prokaryotes, they are found in both archaea and bacteria and are capable of concentrating their substrates greater than 10,000-fold against a concentration gradient and generally have a high affinity for their substrate (Higgins 1992). The organization of ABC transporters involved in uptake is comprised of a cytoplasmic ABC ATPase protein, periplasmic transmembrane protein and a periplasmic binding protein (Holland and Blight 1999, Saurin *et al.* 1999). Although there are reports

of ABC transporters containing multiple periplasmic binding proteins, there have not been any reports of ABC transporters requiring any other accessory protein components (Biemans-Oldehinkel and Poolman. 2003, van der Heide and Poolman 2002).

It has been estimated that about 5% of the *E. coli* genome is dedicated to ABC transporter proteins, accounting for around 70 ABC transporters (Blattner *et al.* 1997, Linton and Higgins 1998). However, genome sequencing of several members of the *Rhizobiaceae* family like *Sinorhizobium meliloti*, *Agrobacterium tumefaciens, Mesorhizobium loti*, *Bradyrhizobium japonicum* and *Rhizobium leguminosarum* (http://www.sanger.ac.uk/Projects/R_leguminosarum/) have shown that these organisms contain a higher number of ABC transporters (Barnett *et al.* 2001, Finan *et al.* 2001, Galibert *et al.* 2001, Kaneko *et al.* 2000, Kaneko *et al.* 2002, Wood *et al.* 2001). Within the *S. meliloti* genome it has been noted that genes encoding transport systems constitute the largest single class of genes accounting for 12% of the total coding capability (Galibert *et al.* 2001). Of these, most belong to the ABC transporter class, and up to 430 ABC transport system genes have been detected (Finan *et al.* 2001, Galibert *et al.* 2001). It may be anticipated that with such a great number of transporters, the *Rhizobium* system may shed new light on the structure or functioning of ABC transporters. In fact, this has been previously shown to be true (Hosie and Poole 2001, Walshaw and Poole 1996).

Strains of *R. leguminosarum* unable to catabolize rhamnose, a methyl-pentose sugar, are compromised in their ability to compete for nodule occupancy versus wild-type strains (Oresnik *et al.* 1998). Characterization of an 11 kb region in wild-type strain Rlt100 of *R. leguminosarum* identified two transcripts that contained the genes necessary

for the catabolism and high affinity transport of L-rhamnose (Oresnik *et al.* 1998, Richardson *et al.* 2004).

We previously reported that the rhamnose catabolic pathway order and biochemical reactions are not conserved between *R. leguminosarum* and *E. coli.* In *E. coli* catabolism is initiated by an isomerization of rhamnose, whereas, in *R. leguminosarum* the first step is presumed to be carried out by the kinase protein encoded by *rhaK,* presumably generating a phosphorylated rhamnose intermediate. To support the genetic data, we have started characterizing RhaK. Mutagenesis of the wild-type rhamnose kinase gene, *rhaK* generated a mutant with a non-transport phenotype. Introduction of a wild-type copy of *rhaK* on a plasmid complemented the growth, biochemical, and transport phenotypes associated with strains carrying a *rhaK* mutation. This suggests that the rhamnose transporter activity in *R. leguminosarum* is dependent upon the presence of a functional sugar-kinase.

3.2 Materials and methods

3.2.1 Bacterial strains and culture conditions

Bacterial strains and plasmids used and generated in this work are listed in Table 3.1. *R. leguminosarum*, strain Rlt100 (original designation W14-2 (Baldani *et al.* 1992)) and strains derived form this wild-type were routinely grown at 30°C on TY as a complex medium (Beringer 1974) and VMM (Vincent 1970) as a defined medium. Carbon sources were filter-sterilized and added to defined media to a final concentration of 15 mM. When required, antibiotics were added to solid or liquid media at the following concentrations: tetracycline (Tc), either 10 μg ml^{-1} or 5 μg ml^{-1}; neomycin (Nm), 200

Table 3.1. Bacterial strains and plasmids

Strain or plasmid	Genotype or phenotype	Reference or source
Strains		
R. leguminosarum		
Rlt100	W14-2, *R. leguminosarum* Sm^r wild-type	(Baldani *et al.* 1992)
Rlt101	W14-2b, cured of plasmid a and d, deletion in plasmid c	(Moenne-Locoz 1995)
Rlt105	Rlt100 *rhaD1*::Tn*5*B20	(Oresnik *et al.* 1998)
Rlt106	Rlt100 *rhaT2*::Tn*5*B20	(Oresnik *et al.* 1998)
Rlt114	Rlt100 *rhaK22*::Tn*5*	(Richardson *et al.* 2004)
Rlt124	Rlt100 *rhaI31*::Tn*5*B20	(Richardson *et al.* 2004)
Rlt128	Rlt100 *rhaP36*::Tn*5*B20	(Richardson *et al.* 2004)
Rlt131	Rlt100 *rhaK40*::Tn*5*B20	(Richardson *et al.* 2004)
Rlt144	Rlt100 *rhaK50*::Tn*5*B20	(Richardson *et al.* 2004)
Rlt146	Rlt100 *rhaK52*::Tn*5*B20	(Richardson *et al.* 2004)
Rlt151	Rlt100 *rhaQ38*::Tn*5*B20	(Richardson *et al.* 2004)
Rlt198	Rlt131, pMR110	(Richardson and Oresnik 2007)
Rlt199	Rlt144, pMR110	(Richardson and Oresnik 2007)
Rlt200	Rlt146, pMR110	(Richardson and Oresnik 2007)
Rlt206	Rlt146, pMR125	(Richardson and Oresnik 2007)
Rlt207	Rlt146, pMR129	(Richardson and Oresnik 2007)
Rlt211	Rlt105, *rhaK58*::pKNOCK-Tc	(Richardson *et al.* 2004)
Rlt212	Rlt124, *rhaK58*::pKNOCK-Tc	(Richardson *et al.* 2004)
Rlt213	Rlt106, pMR110	(Richardson and Oresnik 2007)
Rlt228	Rlt101, pMR86	(Richardson and Oresnik 2007)
Rlt233	Rlt151, pW3C1	(Richardson and Oresnik 2007)
Rlt236	Rlt151, pMR125	(Richardson and Oresnik 2007)
Rlt237	Rlt151, pMR110	(Richardson and Oresnik 2007)
Rlt250	Rlt100, pMR86	(Richardson and Oresnik 2007)
Rlt253	Rlt114, pMR86	(Richardson and Oresnik 2007)
Rlt263	Rlt100, pMR178	(Richardson and Oresnik 2007)
Rlt264	Rlt100, pMR179	(Richardson and Oresnik 2007)
E. coli		
MT616	MT607, pRK600	(Finan 1986)
DH5α	*end*A *hsd*R17 *sup*E44 *thi*-1 *rec*A1 *gyr*A96 *rel*AΔ(*arg*F-*lac*ZYA)U169ϕ80d*lac*ZΔM15	B.R.L. Inc.
Plasmids		
pBluescript II SK	Cloning vector, ColE1 *oriV*, Ap^r	Stratagene
pRK600	pRK2013 *npt*::Tn*9* Cm^r Nm-Km^s	(Finan 1986)
pRK7813	RK2 *ori*, pUC9 polylinker, λ *cos* site, Tc^r	(Jones and Gutterson 1987)
pRSETA	His_6 N-terminal tag, T7 promoter, Ap^r	Invitrogen

pW3C1	pRK7813 cosmid from Rlt100 cosmid bank, containing entire rhamnose locus	(Oresnik *et al.* 1998)
pMR45	*rhaK* coding sequence in pRSETA	(Richardson and Oresnik 2007)
pMR86	pW3C1 *rhaQ38*::Tn*5B20*	(Richardson *et al.* 2004)
pMR106	*rhaK* coding sequence in pBluescript II SK	(Richardson and Oresnik 2007)
pMR110	*rhaK* coding sequence in pRK7813	(Richardson and Oresnik 2007)
pMR125	*rhaK* K16D in pRK7813	(Richardson and Oresnik 2007)
pMR129	*rhaK* K16M in pRK7813	(Richardson and Oresnik 2007)
pMR144	*rhaK* K16D in pRSETA	(Richardson and Oresnik 2007)
pMR149	*rhaK* K16M in pRSETA	(Richardson and Oresnik 2007)
pMR178	5' His_6 tagged *rhaK* coding sequence in pRK7813	(Richardson and Oresnik 2007)
pMR179	3' His_6 tagged *rhaK* coding sequence in pRK7813	(Richardson and Oresnik 2007)

μg ml^{-1}; streptomycin (Sm), 200 μg ml^{-1}; ampicillin (Ap), 100 μg ml^{-1}; kanamycin (Km); 50 μg ml^{-1}. Bacterial growth was monitored spectrophotometrically at 600 nm.

3.2.2 DNA manipulations, sequencing and sequence analysis

Standard techniques were used for plasmid isolation, restriction enzyme digestions, ligations, transformations, and agarose gel electrophoresis (Sambrook 1989). Sequencing reactions were done using dye terminators at the University of Calgary Core DNA Services using an ABI automated sequencer. Sequence data were analyzed using DNASIS (Hitachi Software Engineering Co., San Bruno, CA. U.S.A.). Database searches were done using the BLASTX program (Altschul *et al.* 1997).

3.2.3 Construction of wild-type kinase

Construction of pMR28 was carried out by amplifying the coding region of *rhaK* utilizing the primers kinaseB5' (5' ATA**GGATCC**ATGACCGCCAGTTCCTATC 3') and kinaseH3' (5' ATA**AAGCTT**CTATGCCATCGCCGCGTA 3') utilizing pW3C1 as a template. The amplification product was cloned as a *Bam*HI/*Hin*dIII fragment into pBluescript II SK by utilizing the restriction sites that were introduced into the primers. The construct was verified by nucleotide sequencing.

3.2.4 Over-expression and purification of RhaK

To construct an inframe His_6 tagged RhaK, pW3C1 was used as a template. The primers used in the PCR amplification were kinaseB5' (5' ATAA**GGATCC**ATGACCGCCAGTTCCTATC 3') and kinaseH3', (5'

ATA**AAGCTT**CTATGCCATCGCCGCGTA 3'). This fragment was ligated into pRSETA expression vector (Schoepfer 1993) using *Bam*HI and *Hind*III such that *rhaK* was under the control of the T7 promoter and inframe to an N-terminal His_6 tag. The resulting construct was confirmed by sequencing and named pMR45. pMR45 was transformed into *E. coli* strain BL21 (DE3) (Miroux and Walker 1996). Cultures were grown to mid-log phase (OD_{600} = 0.5) in LB ^{amp} were induced with IPTG (1 mM) at 37°C while shaking at 240 rpm, for 4 hours.

Two additional inframe 5' and 3' His_6 tagged RhaK constructs were generated using pRK7813 broad host range vector such that *rhaK* was expressed constitutively from the p_{lac} promoter from pRK7813, within the wild-type strain Rlt100. The template used for both constructs was pMR110, the primers used in the PCR amplification for the 5' His_6 tagged RhaK construct were 5'His6H (5' ATA**AAGCTT**AGGAGAAATAATTATGAGAGGATCGCATCATCATCATCATCATA CCGCCAGTTCCTATCGC 3') and 3'BAMHI/PSTI, (5' ATAT**GGATCCCCTGCAG**CTATGCCATCGCCGCGTA 3'). The primer used in the 3' His_6 tagged RhaK construct were 5'HINDIII/RBS (5' ATAT**AAGCTT**GGAGAACTGCAGATGACCGCCAGTTCCTATC 3') and 3'His6B (5'AT**GGATCC**CTAATGATGATGATGATGATGCGATCCTCTTGCCATCGCCGCG TACC 3') these fragments were ligated into pRK7813 using *Bam*HI and *Hind*III. The resulting construct was confirmed by sequencing using the primers pRK7813 forward primer (5' GTGCTGCAAGGCGATTAAGTT 3') and pRK7813 reverse primer (5' TGTGAGCGGATAACAATTTCAC 3'). The constructs were named pMR178 (5' His_6 tagged RhaK) pMR179 (3' His_6 tagged RhaK). These constructs were mated into wild-

type Rlt100 by selecting the antibiotic marker on the vector, they were called Rlt263 (Rlt100, pMR178) and Rlt264 (Rlt100, pMR179). Cultures of strains of His_6 tagged RhaK were grown to mid-log phase (OD_{600} = 0.5) in TY ^{TC} at 30°C while shaking at 240 rpm.

Cells were harvested by centrifugation (6000 x g for 10 minutes) and resuspended in 2 ml g^{-1} wet pellet in binding buffer (20 mM Tris pH 7.8, 300 mM NaCl, 10 mM imidazole, 5 mM β-mercaptoethanol). Cells were lysed with two passages through a French pressure cell (16000 psi) and cell extracts were prepared by centrifugation (6000 x g for 10 minutes). His_6 tagged RhaK was immobilized by passing cell free extract through a nickel affinity column, attached to FPLC apparatus (Bio-Rad biologic LP chromatography system). Washing was done with wash buffer (20 mM Tris pH 7.8, 300 mM NaCl, 20 mM imidazole, 5 mM β-mercaptoethanol) for 60 minutes at a 1 ml min^{-1} flow rate. The RhaK fusion was then eluted from the Nickel affinity column (Bio-Rad Ni-NTA superflow) using elution buffer (20 mM Tris pH 7.8, 300 mM NaCl, 250 mM imidazole, 5 mM β-mercaptoethanol). The purity of pMR45 was confirmed with SDS-PAGE gel electrophoresis. SDS gels of eluted fractions were transferred electrophoretically to a nitrocellulose membrane for western blot analysis. The presence of the His_6 tagged RhaK was confirmed colormetrically using goat-anti-mouse HRP antibody with the opti-4CN substrate detection kit as suggested by the manufacturer (Bio-Rad Laboratories Hercules Ca.).

3.2.5 Generation and over-expression of rhaK active site mutations.

The generation of *rhaK* ATPase active site base pair mutations was accomplished using the megaprimer PCR method (Sambrook and Russell 2001). The template DNA used in this amplification was from pMR106. Construction of pMR106 was carried out by amplification of *rhaK* using pW3C1 as template and the primers 5'HindIII/RBS (5' ATAT**AAGCTT**GGAGAACTGCAGATGACCGCCAGTTCCTATC 3') and 3'BamHI/PstI (5' ATAT**GGATCCCTGCAG**CTATGCCATCGCCGCGTA 3'). This fragment was ligated into pBluescript II SK and into the broad host range vector pRK7813 (Jones and Gutterson 1987), using *Bam*HI and *Hind*III restriction enzyme sites within the primers and called pMR106 and pMR110 respectively. The primers used for the megaprimer method were R1 external (5' GTAAAACGACGGCCA 3'), and F2 external primer (5' GGATCCCTGCAGCTATGCCATCGC 3'), mutagenic primers lysine to methionine M primer (3' TGTAGCCGTACTGGTT 5') or lysine to aspartate M primer (3' TGTAGCCGCTGTGGTT 5'). The amplification product was cloned as a *Bam*HI/*Hind*III fragment into pRK7813 using restriction sites introduced into the primers, subsequent clones were called pMR129 (K to M) and pMR125 (K to D), these constructs were verified by nucleotide sequencing. Protein over-expression analysis of the *rhaK* base pair mutants carried on pMR129 and pMR125, was accomplished using the same cloning strategy for over-expression of pMR45; the only deviation to the protocol was the template used, pMR125 was used to generate pMR144, aspartate active site his_6 tagged mutant while pMR129, methionine active site His_6 tagged mutant was used as template for pMR149. Western blot analysis was accomplished using Bio-Rad opti-4CN goat-anti-mouse HRP antibody detection kit, to identify the His_6 tagged *rhaK* active site mutants.

3.2.6 Transport assays

Uptake of rhamnose was carried out essentially as previously described (Richardson *et al.* 2004). Radioactive [^{3}H] rhamnose (5 Ci/mmol) was purchased from American Radiolabeled Chemicals Ltd. St. Louis, Missouri. Transport assays were initiated by the addition of tritiated rhamnose to a final concentration of 2 μM (125,000 dpm) 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 the defined salts medium and the radioactivity that remained on the filter was determined using a liquid scintillation spectrophotometer (Beckman LS6500). Samples were taken every 20 seconds and continued for up to two minutes. Initial rates were calculated from the linear portion of the data (generally the first 60 seconds).

Whereas our standard uptake protocol uses cells that have been induced overnight in defined medium containing glycerol and L-rhamnose (Richardson *et al.* 2004), strains carrying either *rhaDI* or *rhaI* mutations were unable to grow in containing in this medium for these extended time periods so optimal induction times for strain Rlt105 and Rlt124 were determined experimentally. It was found that 7 hours of exposure to L-rhamnose was sufficient to induce these genes without any detriment to the growth of these strains due to the sensitivity phenotype exhibited by isomerase or dehydrogenase/aldolase mutant strains (data not shown).

3.2.7 In vitro RhaK dependent rhamnose intermediate assays

Strains were grown on minimal medium to mid-log phase, harvested by centrifugation (6000 x g for 10 minutes) and resuspended in 2 ml g^{-1} wet pellet in lysis buffer. Resuspended cells were lysed by 2 passages through a French pressure cell (16000 psi) and cell debris was removed by centrifugation (10000 x g for 10 min at 5°C in a Sorval SLA-3000 rotor).

Whatman DE81 cellulose chromatography paper was first treated with 1 M formic acid, then extensively washed with ddH_2O rinse to remove any impurities and finally dried prior to use. Isolated, induced, cell free extracts were combined with 50 mM HEPES buffer pH 7.6, 5 mM $MgCl_2$, 10 mM ATP, and 2 mM tritiated rhamnose and incubated at 30°C. Assays were carried out in a final volume of 50 µl. 10 µl samples were taken at 1, 5, 10 and 30-minute time points and quenched by placing each sample in a boiling water bath for 3 minutes then into an ice bath until all samples were ready to be loaded. The 10 µl quenched aliquots were spotted and separated by descending paper chromatography using 50 mM formic acid pH 3.2. To determine the position of the label, the chromatography paper was cut into 1 inch square pieces (representing approximately 10-15% of the total paper length) and assayed using a liquid scintillation spectrophotometer (Beckman LS6500).

3.2.8 RhaK phylogenetic tree construction

Kinase sequence were selected based on homology of the RhaK amino acid sequence using BLASTP (Altschul *et al.* 1997). The sequences were aligned using CLUSTAL X (Thompson *et al.* 1997). Programs contained within PHYLIP Version 3.6a (Felsenstein 1985, Felsenstein 2002) were used for phylogenetic analysis. Divergence (or

distance) between two sequences was calculated by PROTDIST (setting: JTT). The resulting distance matrix was used to construct a phylogenetic tree with the NEIGHBOR program (NJ option), and the phylogenetic tree was evaluated using the bootstrap procedure (SEQBOOT, 1000 replicates). Bootstrap replicates were analyzed with CONSENSE program to generate the majority rule consensus tree. The phylogenetic tree were drawn using the TREEVIEW (Page 1996) program using the PHYLIP output file as an input file. *Pirellula* sp.1 was selected as the comparative out-group.

3.3 Results

3.3.1 Strains carrying a mutant rhaK allele do not transport rhamnose

We previously cloned, sequenced, and characterized a region from *Rhizobium leguminosarum* bv. *trifolii* necessary for the utilization of rhamnose (Oresnik *et al.* 1998, Richardson *et al.* 2004). Strains containing mutations within this locus were shown to be less competitive for nodule occupancy when co-inoculated with the wild-type strain (Oresnik *et al.* 1998). Nucleotide sequence analysis of this region suggests that it contains three genes necessary for the catabolism of rhamnose *rhaD*, *rhaI,* and *rhaK*, encoding a dehydrogenase/aldolase, an isomerase, and a kinase respectively. This locus also contains all the components necessary for a functional ABC importer; *rhaS, rhaT, rhaP, and rhaQ* encode a sugar binding protein, an ABC ATPase and two transmembrane permeases respectively, in addition to a negative regulator (*rhaR*) and a small gene that encodes a protein of unknown function (*rhaU*). These genes are organized in two divergently transcribed operons. One transcript consists of *rhaDI* whereas the other consists of *rhaRSTPQUK*. Expression and transport activity are induced by rhamnose, the

transporter has a high affinity for its substrate, and strains carrying mutations that affected the expression of the ABC transporter components are unable to transport rhamnose (Richardson *et al.* 2004).

While characterizing the rhamnose locus we used strain Rlt144, which carries the *rhaK50* allele generated by a *Tn*5B20 insertion, as a putative positive control for the rhamnose transport assays. This appeared to be a reasonable choice since a mutation in *rhaK* is not polar on any of the core components of the ABC transporter. Surprisingly, strain Rlt144 appeared to have a non-transport phenotype similar to that of the other components of the ABC transport complex (Figure 3.1A). To ensure that this was not a strain specific artifact, two other strains, Rlt131 and Rlt146, carrying *rhaK40* and *rhaK52* respectively, were also assayed for their ability to transport rhamnose. These strains were independently isolated and have different points of insertion in *rhaK* (Richardson *et al.* 2004). In both cases these strains did not show any transport activity (Figure 3.1).

We had previously shown that introduction of a *rhaK* mutant allele into a strain carrying a polar *rhaD* mutant (genotypically *rhaD⁻rhaT*) enables this strain to use glycerol as a carbon source in the presence of rhamnose; whereas the inability of strains carrying a polar *rhaD* mutant with a functional kinase presumably led to the accumulation of a toxic phosphorylated intermediate (Richardson *et al.* 2004). In light of our result with other *rhaK* alleles, we wished to ascertain if in fact the phenotype reversal was due to an inability to take up rhamnose. Strain Rlt105 (*rhaDT*) was clearly capable of transporting rhamnose (Figure 3.1B). We note that the rates appear to be somewhat lower than the rates measured for Rlt100; however, this may be attributed to the toxicity exhibited by this strain which required a shorter induction time such that sufficient cells

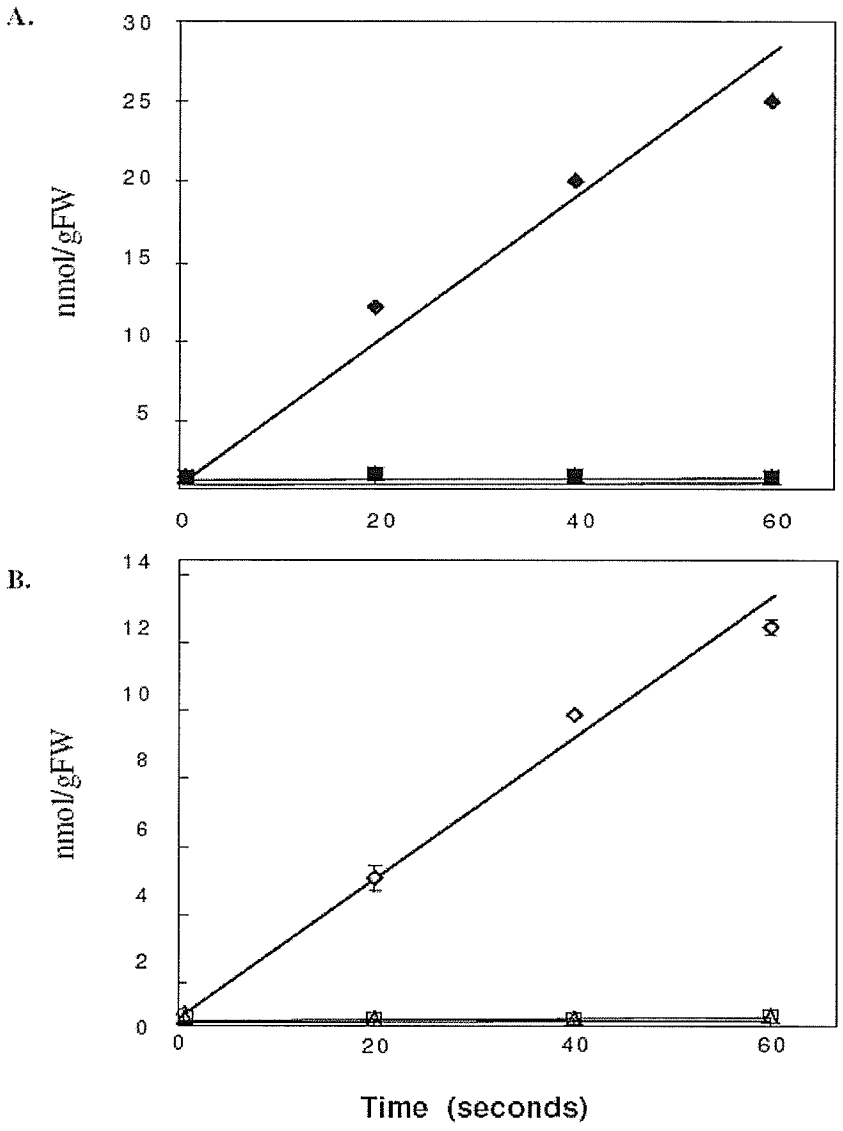


Figure 3.1. L-Rhamnose transport assays. Strains were grown to mid-log phase in defined medium containing rhamnose/glycerol and uptake was determined using tritiated rhamnose. **A.** contains the wild-type strain (Rlt100) and three independent *rhaK* mutants (Rlt131, Rlt144, and Rlt146). **B.** contains a *rhaD/I* mutant (Rlt105) as well as *rhaD/K* mutant (Rlt211) and *rhaI/K* (Rlt212) double mutants. Symbols are as follows: filled diamond, Rlt100; filled square, Rlt131; filled triangle, Rlt144; x, Rlt146; open diamond, Rlt105; open square, Rlt211; open triangle, Rlt212. The data represents the average of at least three independent experiments. Error bars were smaller than the symbols if absent. Note that where not visible, lines and symbols overlap.

could be grown in a glycerol/rhamnose medium to measure uptake rates. Strains Rlt210 (*rhaDΓ*, *rhaK*⁻), and Rlt211 (*rhaΓ*, *rhaK*⁻), like other strains containing *rhaK*⁻ alleles, were unable to transport labeled rhamnose (Figure 3.1 B). The reverse of the glycerol sensitivity phenotype in Rlt211 and Rlt212 is likely due to the absence of an unmetabolizable intermediate; these results suggest that inability to transport rhamnose rather than phosphorylation is responsible for the reversal of the sensitivity.

3.3.2 Rhamnose kinase mutants can be complemented with rhaK⁺ in trans

The inability of *rhaK* mutants to transport rhamnose could be due to the dependence of the transporter on the presence of RhaK. Alternately the insertion mutations could cause a transcript instability that led to an absence of the other components of the ABC transporter. To rule out the latter possibility, it was reasoned that if the rhamnose mutants were complemented with a wild-type copy of *rhaK*, and the ability to grow on rhamnose as a sole carbon source, as well as the ability to transport rhamnose were restored, then it could be concluded that the mutations were not affecting the expression of the other ABC transport components. *rhaK* was PCR amplified and cloned with a ribosome binding site into the broad host range plasmid pRK7813 such that transcription would occur constitutively from P_{lac} which is part of the plasmid multiple cloning site (Jones and Gutterson 1987). The resultant construct was verified by sequencing and conjugally introduced into strains Rlt131, Rlt144, and Rlt146 by selecting for the transfer of the plasmid encoded antibiotic determinant as previously described (Finan 1986). The transconjugant strains were designated Rlt198, Rlt199, and Rlt200 respectively. These were tested first for their ability to grow using rhamnose as a

sole carbon source, and subsequently for their ability to transport rhamnose. In all cases the introduction of *rhaK*$^{+}$ into strains carrying a *rhaK* mutation resulted in the strains being complemented both for their ability to grow on rhamnose as a sole carbon source as well as their ability to transport rhamnose at levels comparable to that of the wild-type Rlt100 (Table 3.2).

The results however do not eliminate the possibility that RhaK may have cryptic transport functions independent of the core ABC transport complex. To address this remote possibility strain Rlt213 was constructed. Strain Rlt213 was constructed by introducing a plasmid encoding a wild-type copy of *rhaK* into strain Rlt106 which carries the *rhaT2* allele. The *rhaT2* mutation is a lesion in the gene that encodes the ABC protein, a critical component of a functioning ABC transporter. Moreover the mutation is also polar on *rhaPQUK*, thus eliminating all components of the ABC transport complex except the periplasmic sugar binding protein encoded by *rhaS*. Neither Rlt106 nor Rlt213 could grow on minimal media supplemented with rhamnose, nor could they transport tritiated rhamnose from the growth medium (Table 3.2). Together these results suggest that transport of rhamnose in Rlt100 is dependent on both the components of the ABC transporter (RhaSTPQ) and the kinase (RhaK).

3.3.3 Strains with rhaK mutations retain transcription of ABC components

It is possible that the metabolic product of the RhaK or a subsequent catabolic step may alter the transcription of the ABC transporter genes and thus lead to an absence of rhamnose transport. To address this issue, a series of constructs were made that contained the *rhaQ38 lacZ* fusion allele either integrated into the genome, or present on a

Table 3.2. Transport of L-rhamnose is dependent on the presence of RhaK and an ABC transport complex

Strain	Relevant Genotype	Glucose [a]	Rhamnose [b]	Transport (nmol/gFW/min) [c]
Rlt100	wild-type	+	+	28 ± 4
Rlt131	*rhaK40*	+	-	< 0.05
Rlt144	*rhaK50*	+	-	< 0.05
Rlt146	*rhaK52*	+	-	< 0.05
Rlt198	*rhaK40; rhaK*$^+$	+	+	27 ± 3
Rlt199	*rhaK50; rhaK*$^+$	+	+	30 ± 1
Rlt200	*rhaK52; rhaK*$^+$	+	+	31 ± 1
Rlt213	*rhaT2; rhaK*$^+$	+	-	< 0.05
Rlt106	*rhaT2*	+	-	< 0.05

[a b] Strains carrying *rhaK* alleles, *rhaT* alleles or *rhaK, rhaT* alleles and a plasmid containing a wild-type copy of *rhaK* were tested for their ability to grow on minimal medium (VMM) supplemented with 15 mM rhamnose or glucose.
[c] Strains were grown as broth cultures in minimal medium (VMM) supplemented with 15 mM rhamnose/glycerol. Initial transport rates were determined using tritiated rhamnose.

plasmid, in the presence, or absence, of *rhaK*, expressed from either a plasmid or the genome (Table 3.3). The results clearly show that the fusion activity from *rhaQ38* is not abolished in the absence of *rhaK*. The data suggests that the transcription of *rhaSTPQ* does occur in the *rhaK* alleles that are transport deficient.

3.3.4 Detection of a kinase-dependent rhamnose intermediate

The inability of strains carrying *rhaK* mutations to take up rhamnose poses yet another question, does RhaK need to be present as structural entity, or in a biochemically active form for transport to occur? Genetic evidence is consistent with RhaK acting on rhamnose before either the *rhaD* or *rhaI* gene products (Richardson *et al.* 2004) (Figure 3.1, Table 3.2). Since rhamnose kinase activity has not been previously reported, a standard assay for RhaK does not exist. Therefore direct testing of the genetic data has not been possible. To answer these questions an *in vitro* RhaK assay was developed.

Initial assays relied on over-expressing wild-type *rhaK* with an N-terminal His_6 tag, eluting the bound RhaK from a nickel affinity column, dialyzing the protein, and incubating the protein in the presence of an appropriate buffer, ATP, and radiolabeled rhamnose. The reaction product(s) were then separated by paper chromatography using Whatman DE-81 paper and a formic acid mobile phase in a manner analogous to separating phosphorylated nucleosides from nucleoside sugars (Zadrazil 1973). Utilizing these conditions phosphorylated compounds are extremely retarded in their migration whereas non-phosphorylated sugars migrate at, or very near, the solvent front (Zadrazil 1973). The position of the label was then quantified using a scintillation counter.

Table 3.3. Expression of *rhaQ* was not abolished in *R. leguminosarum* strains containing *rhaK* mutations

Strain	Relevant Characteristics	*rhaK* Genotype	Glucose[a b]	Rhamnose	Induction[c]
Rlt100	wild-type	+	12	9	n/a
Rlt101	Δ *rha* region	-	9	7	n/a
Rlt114	*rhaK22*::Tn*5*	-	9	9	n/a
Rlt151	*rhaQ38*::Tn*5*B20; genomic fusion	-	75 (3)	406 (104)	5.4
Rlt100(pMR86)[d]	*rhaQ38*::Tn*5*B20; cosmid fusion	+; from genome	150 (8)	507 (48)	3.4
Rlt101(pMR86)[e]	*rhaQ38*::Tn*5*B20; cosmid fusion	-	230 (30)	677 (71)	2.9
Rlt114(pMR86)[f]	*rhaK22*::Tn*5*/*rhaQ38*::Tn*5*B20; cosmid fusion	-	158 (3)	517 (37)	3.3
Rlt151(pW3C1)[g]	*rhaQ38*::Tn5B20; genomic fusion	+; from cosmid	250 (24)	668 (71)	2.7
Rlt151(pMR110)[h]	*rhaQ38*::Tn*5*B20; genomic fusion	+; from plasmid	48 (8)	668 (71)	13.9
Rlt151(pMR125)[i]	*rhaQ38*::Tn*5*B20; genomic fusion	-; bp mutation on plasmid	83 (2)	238 (19)	2.9

[a] The values given represent β-galactosidase activity expressed in Miller units following growth in defined media with either glucose/glycerol or rhamnose/glycerol as carbon sources.
[b] Values are the average of at least three independent replicates; values in parenthesis represent standard error. Rlt100, Rlt101, and Rlt114 do not contain transcriptional fusions; values are representative of one data set and represent background readings in the β-galactosidase activity assay.
n/a = not applicable.
[c] Induction was calculated by dividing activity in the presence of rhamnose by the activity in the presence of glucose.
[d, e, f, g, h, i] Are referred to in the strain list as Rlt250, Rlt228, Rlt253, Rlt233, Rlt237, and Rlt236 respectively

Although we were capable of isolating a full-length protein product, the yields were low, and experiments utilizing labeled rhamnose as a substrate were never successful. Moreover, if transport activity was dependent upon RhaK, we were not convinced that RhaK activity might not be dependent on ABC transport components. Additional component(s) required for kinase activity might be present in a *R. leguminosarum* cell free extract whereas they may be missing from an over-expressed protein isolated from *E. coli*. We therefore abandoned the over-expression system and developed an assay using a cell-free lysate of Rlt100.

Assays utilizing either labeled ATP or measuring the amount of ATP hydrolyzed were not practical because the assays were carried out using a crude extract. For the assay to be specific we set up the following criteria; activity from the cell free extract would need to be dependent upon rhamnose, ATP, induced extract, the signal of the intermediate(s) would need to be well over the background, and accumulation of reaction products would need to be time dependent. A labeled peak accumulated with an Rf value of 0.24 was observed following a thirty minute assay (Figure 3.2). The biochemical conversion of substrate into a rhamnose species was time dependent over the course of the assay and the peak observed on the chromatography paper was consistent with the behavior of a phosphorylated intermediate (Figure 3.2). This peak was dependent on the presence of induced extract, ATP, and rhamnose (data not shown).

3.3.5 Accumulated peak is RhaK dependent

Although the peak that accumulated with an Rf value of 0.24 fulfilled the criteria for a RhaK assay, it was necessary to show that this peak was dependent upon *rhaK*. To

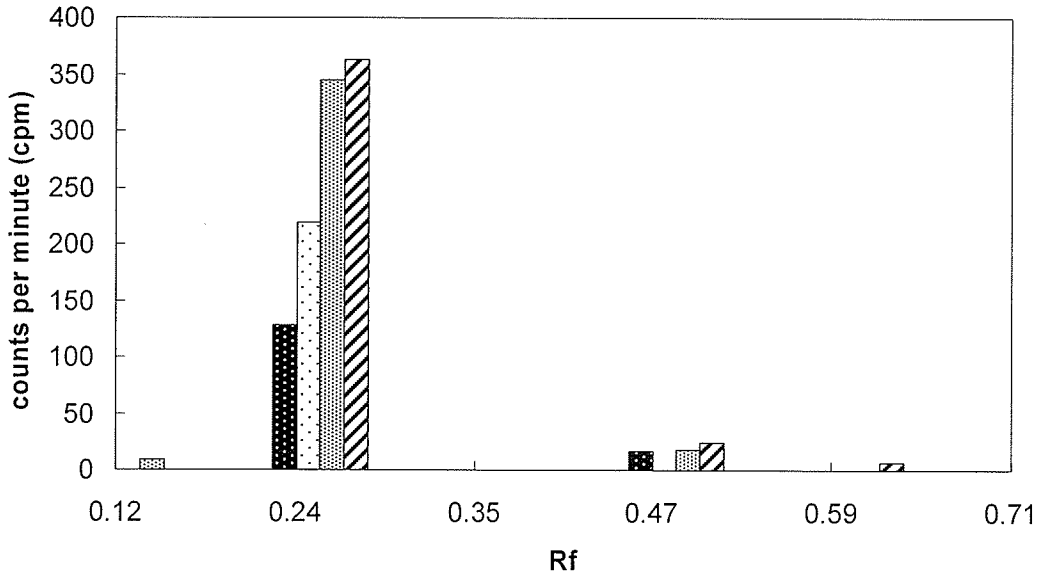


Figure 3.2. Accumulation of a time dependent intermediate using descending paper chromatography. Rlt100 wild-type rhamnose induced cell extract was isolated and assayed as described in materials and methods. Samples were taken and quenched at one, five, ten fifteen, and thirty minutes. Symbols are as follows: filled bar, one minute; filled bar white spots, five minutes; white bar few black spots, ten minutes; white bar dense black spots, fifteen minutes; white bar cross-hatched, thirty minutes. The tritiated rhamnose peak that runs at the solvent front (Rf >0.8) is not shown. One representative set of data is presented; experiments were replicated at least three times.

confirm, we used induced extracts from strains carrying various alleles that did not allow growth using rhamnose as the sole carbon source.

When rhamnose induced cell free extracts from strains and Rlt105 (*rhaD*1/*I*$^-$), Rlt146 (*rhaK*52), Rlt211 (*rhaD*1/*rhaK*58), and Rlt212 *(rhaI*31/*rhaK*58) were assayed, activity was present in Rlt105, whereas it was abolished Rlt144, Rlt211, and Rlt212 (Table 3.4). To determine if this activity could be restored with the addition of wild-type *rhaK*, strain Rlt200, was also assayed (Table 3.4). These data show that the activity observed in the RhaK assay is independent of *rhaD* and *rhaI*, its absence is correlated with the presence of a *rhaK* mutant allele, and complementation with a wild-type copy of *rhaK* restores activity. Together these data show that the activity we are assaying is due to RhaK.

3.3.6 Transport and biochemical activities are dependent on the presence of a functional kinase.

RhaK affects both the interconversion as well as a whole cell transport of rhamnose (Table 3.4, Table 3.2). From these data it is not clear if transport and interconversion are interdependent. To resolve this we wished to construct site-directed mutants of RhaK that should have compromised biochemical activity and determine whether these mutants could transport rhamnose.

Database searches using BLASTP (Altschul *et al.* 1997), as well as CDART (Geer *et al.* 2002), clearly recognize RhaK to be a sugar kinase (e value = 8×10^{-46}) predicted to belong to the sugar kinase-hsp70-actin ATPase protein superfamily (Figure 1.8, Figure 1.9) (Bork *et al.* 1992). This family of proteins contains a conserved β-loop

Table 3.4. Activity of phosphorylated intermediate is *rhaK* dependent

Strain	Relevant Genotype	Activity (μmol/min/mg)[a]
Rlt100	wild-type	136.8 ± 29.6
Rlt105	*rhaD1*	33.0 ± 10.4
Rlt124	*rhaI31*	101.0 ± 40.6
Rlt146	*rhaK52*	n.d.[b]
Rlt200	*rhaK52/rhaK*$^{+}$	32.9 ± 5.8
Rlt211	*rhaD1/rhaK58*	n.d.
Rlt212	*rhaI31/rhaK58*	n.d.

[a] Strains were grow in minimal medium (VMM) supplemented with 15 mM rhamnose/glycerol, induced cell extract was isolated and combined with 50 mM HEPES buffer pH 7.6, 5 mM MgCl2, 10 mM ATP and 2 mM tritiated rhamnose then incubated at 30°C. Samples were taken and quenched at one, five, ten fifteen, and thirty minutes. 10 μl of this mixture was then spotted and run on Whatman DE81 cellulose chromatography paper with 50 mM formic acid as a solvent. The filter paper was partitioned and sections corresponding to Rf values of approximately 0.24 were assayed for labeled rhamnose using liquid scintillation. Data represent a minimum of three independent replicates the standard error is shown.

[b] n.d. not detected.

motif that forms the pocket that binds the γ-phosphate of ATP, it consists of (Ile/Leu/Val)-X-(Ile/Leu/Val/Cys)-Asp-X-Gly-(Thr/Ser/Gly)-(Thr/Ser/Gly)-X-X-(Arg/Lys/Cys) and is relatively conserved amongst members of the FGGY family (Figure 1.8, Figure 1.11) (Flaherty *et al.* 1991). The invariant Asp interacts with the Mg^{2+} ion complexed with the ATP (Figure 1.10) (Pettigrew *et al.* 1998), the invariant Gly contacts the γ-phosphate of the ATP molecule (Figure 1.11) (Kabsch and Holmes 1995). The hydroxyl from the amide side chains of (Thr/Ser/Gly) and (Thr/Ser/Gly) residues bond the oxygen atoms from the γ-phosphate of ATP, while the (Arg/Lys/Cys) appears to balance the negatively charged γ-phosphate (Mao *et al.* 1999).

This conserved motif corresponds roughly to the RhaK sequence Ile9, Ala10, Val11, Leu12, Asp13, Ile14, Gly15, Lys16, Thr17, Asn18, Ala19, and Lys20 where the spacing at the beginning of the conserved region differs in addition to Lys16 in RhaK instead of (Thr/Ser/Gly). Since RhaK activity is dependent on ATP (data not shown) the β-loop motif of RhaK was targeted for mutagenesis. The large positively charged lysine16 residue within the β-loop motif was changed to be either an aspartate or a methionine. We hypothesized that both of these substitutions of Lys16 in RhaK within the β-hairpin likely involved in contacting the γ-phosphate of ATP, should inhibit binding of the γ-phosphate of ATP within the pocket, thereby preventing ATP dependent kinase activity (Figure 1.12). The mutated copies of the gene were cloned into the broad host range vector pRK7813, sequenced, and introduced into strain Rlt146, generating strains Rlt206 and Rlt207.

In both cases, the mutant strains Rlt206 and Rlt207 were unable to grow on minimal media supplemented with rhamnose as a sole carbon source (Table 3.5). These

Table 3.5. Biochemical activity of RhaK is necessary for growth on and transport of L-rhamnose

Strain	Relevant Genotype	Glc [a]	Rha [b]	Transport (nmol/gFW/min) [c]	Activity (μmol/min/mg) [d]
Rlt100	wild-type	+	+	30.7 ± 1	117.9 ± 29.3
Rlt146	*rhaK*52	+	-	< 0.05	n.d.[e]
Rlt200	*rhaK52/rhaK*$^+$	+	+	23.5 ± 1	48.1 ± 17.1
Rlt206	*rhaK*52/*rhaK* K16D	+	-	< 0.05	n.d.
Rlt207	*rhaK*52/*rhaK* K16M	+	-	< 0.05	n.d.

[a b] Strains carrying *rhaK* alleles, or *rhaK* alleles and a plasmid containing a wild-type copy of *rhaK* were tested for their ability to grow on minimal medium (VMM) supplemented with 15 mM Rha; rhamnose, or Glc; glucose.
[c] Strains were grown as broth cultures in minimal medium (VMM) supplemented with 15 mM rhamnose/glycerol. Initial transport rates were determined using tritiated rhamnose. Data represent a minimum of three replicates.
[d] Data represents three independent replicates.
[e] n.d. not detected.

strains were also unable to import labeled rhamnose in radioactive transport assays and both had transport rates of < 0.05 nmol/gFW/min (Table 3.5). Additionally, neither mutant had any observable kinase activity in the radioactive chromatography assay (Table 3.5).

To ensure that the γ-phosphate β-loop mutants were indeed being translated properly, and that our mutants were not simply unable to be properly translated, the wild-type, and both mutant kinase genes were PCR amplified, cloned into the pRSETA expression vector containing an N-terminal His_6 tag, plasmids generating plasmids pMR45, pMR149 and pMR144. The inserts of these constructs were completely sequenced, transformed into *E. coli* strain BL21 (DE3) and expressed. Western blot analysis of all three strains utilizing a goat-anti-mouse His_6 monoclonal antibody showed that each construct was capable of being translated to a peptide of the correct size in *E. coli* (data not shown). Together these data suggest that transport of rhamnose by the rhamnose ABC transporter is dependent upon biochemically active RhaK.

3.3.7 Phylogenetic and domain analysis of RhaK

Since the only motif that we were able to identify was from a family of proteins that contain several sugar kinases, it was thought that a phylogenetic analysis of kinases with high amino acid identities to RhaK might be useful in identifying other sugar kinases that may affect transport function. It was reasoned that this analysis might also identify regions that are conserved and thus be necessary for either transport or kinase activity.

Analysis of 35 related kinases is shown (Figure 3.3). Since the genome of Rlt100 has not been sequenced it was not possible to identify orthologues in related organisms using a reciprocal best match method (Parkhill 2002). However, searches for homologues of RhaK in the individual proteomes of *A. tumefaciens*, *M. loti*, and *S. meliloti* using BLASTP show that the proteins delineated by the shaded box in Figure 3.3 are the best matches in the respective organisms. Moreover, an examination of the flanking sequences to the respective genes shows an identical arrangement of genes and putative transcripts at this locus in each organism. In addition, insertion mutations isolated in the *S. meliloti* gene, Smc03003 (NP_384741.1), also have an inability to grow utilizing rhamnose as a sole carbon source (Poysti, Richardson, and Oresnik unpublished data). Using this protein sequence to carry out a reciprocal best hit analysis against the genomes of *M. loti* and *A. tumefaciens* showed that the delineated area represents proteins that are orthologous to RhaK and more than likely are involved in rhamnose catabolism in these organisms (Figure 3.3).

To facilitate identifying regions that may play an important role in the function of this protein the most closely related proteins were aligned using both CLUSTAL X (Thompson *et al.* 1997) and PRALINE (Simossis and Heringa 2005). Whereas the former program aligns on the basis of statistical probability using the entire protein, PRALINE utilizes a PSI-BLAST step that constructs an alignment by recognizing localized regions of identity (Simossis and Heringa 2005). Both alignments gave congruent results and the PRALINE data are presented (Figure 3.4). A comparison of the two analyses shows that among the most closely related proteins to RhaK, there are at least 65 highly conserved residues (score > 7). Surprisingly, of these, 14 of the conserved

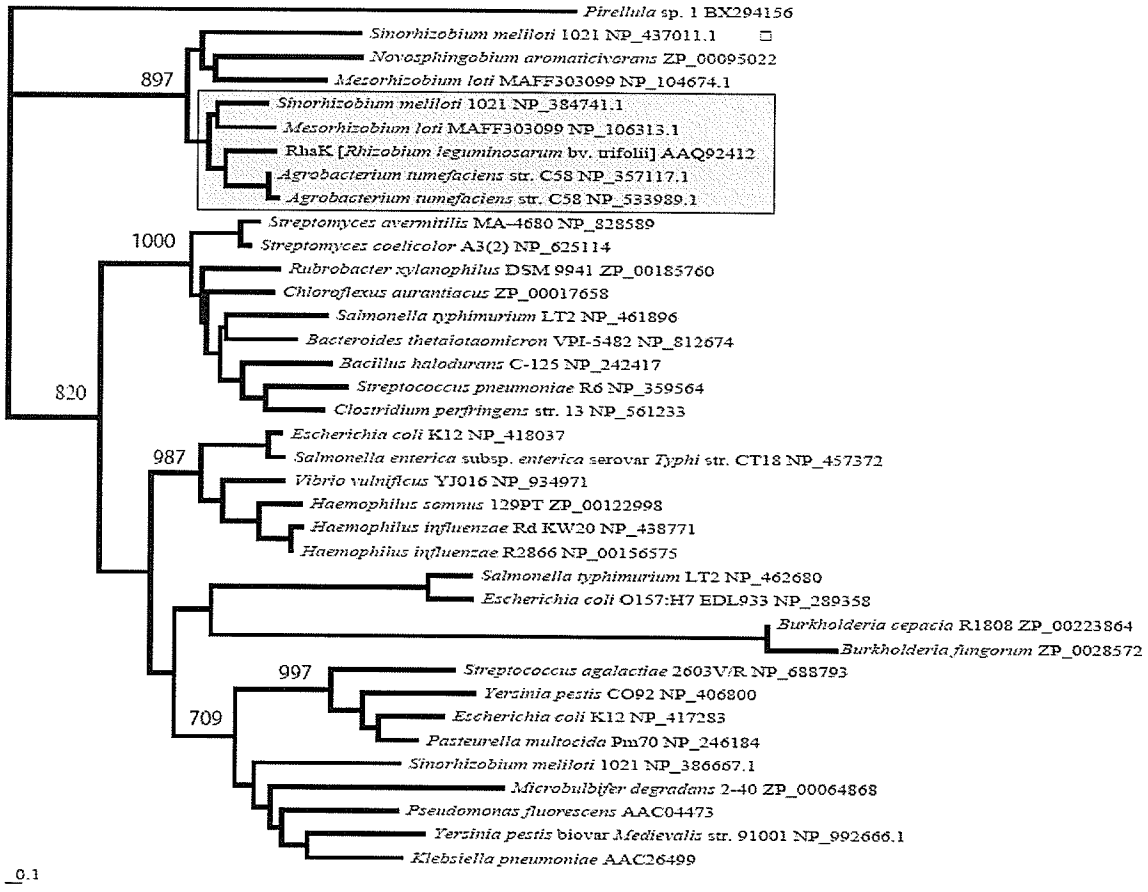


Figure 3.3. RhaK phylogenetic tree based on predicted kinase sequences. Related kinase sequences were identified using BLASTP and the tree was constructed as described in materials and methods. Sequence accession numbers for the predicted sequences are listed next to the corresponding strains. Bootstrap values less than 700 are not shown.

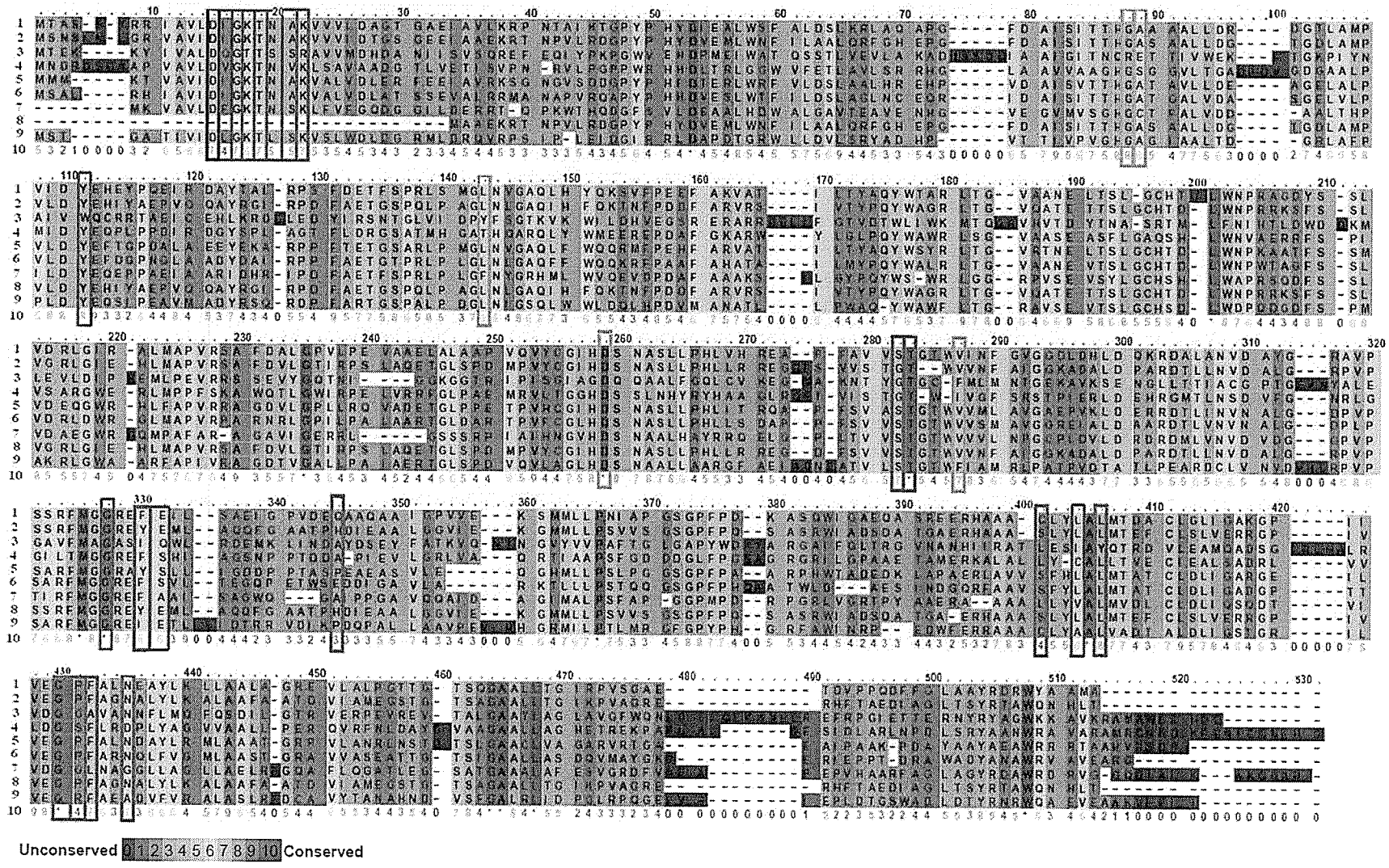

Unconserved 0 1 2 3 4 5 6 7 8 9 10 Conserved

Figure 3.4. PRALINE alignment of sugar kinases similar to RhaK. Alignments of RhaK-related kinases and glycerol kinase were performed using PRALINE. The amino acid consistency scoring ranges from 0 (least conserved) to 10 (invariant, denoted as *). The amino acid consistencies are color-coded, from least conserved: dark blue; light blue; dark green; light green; yellow; orange; dark orange; red, most conserved, and correspond to numbers below the alignment. The aligned RhaK and related kinase residues are highlighted using GlpK residues with known interactions as a model. Glycerol kinase ATP/ADP binding, Mg^{2+} binding, and substrate binding are highlighted by a red rectangle, blue rectangle and a green rectangle respectively. Labels are as follows:1, *R. leguminosarum,* strain Rlt100, RhaK (AAQ92412); 2, *A. tumefaciens*, strain C58, AGR_L_2672 (NP_357117); 3, *E. coli*, strain K12, GlpK (glycerol kinase) (NP_418361); 4, *S. meliloti*, strain 1021, (NP_437011); 5, *S. meliloti*, strain 1021, SMc03003 (NP_384741.1); 6, *M. loti*, strain MAFF303099, (NP_106313.1); 7, *M. loti*, strain MAFF303099, mll3599 (NP_104674.1); 8, *A. tumefaciens*, strain C58, (NP_533989.1); 9, *N. aromaticivorans*, (ZP_00305423); 10, consistency.

residues are either glycines or alanines, 19 are either isoleucines, leucines, or valines, and the remainder are charged residues (Figure 3.4). Overall the alignment of the data set shows that there are regions of conservation providing evidence for functional constraint among these proteins.

CDART analysis of RhaK clearly defines the RhaK as a sugar kinase that aligns with the FGGY-type sugar kinases. FGGY type kinases usually contain a FGGY N and FGGY C domain. These two domains have been shown to be able to interact with each other. RhaK contains only a FGGY N domain and is 99.2% aligned with it between the first amino acid of the protein and amino acid residue 244. One of the better characterized members of this family of proteins is *E. coli* glycerol kinase, GlpK. To better understand RhaK, the protein sequences of RhaK and GlpK were aligned (data not shown). Utilizing the GlpK crystal structure data we were able to putatively identify residues that were important for substrate and ATP binding as well as active site residues (Hurley *et al.* 1993, Pettigrew *et al.* 1998). Using GlpK as a model, the catalytic residue in the putative active site for RhaK, and the other aligned kinases, is the aspartate at position 259 in the alignment. Consistent with what is known about GlpK, where the active site aspartate is found in a hydrophobic pocket (Pettigrew *et al.* 1998), the RhaK Asp238, position 259 in the alignment, is also found in a highly hydrophobic region of the protein (data not shown). As well, the majority of the residues necessary for ATP/ADP binding by GlpK are also conserved in RhaK and its homologues (Figure 3.4).

3.4 Discussion

In this Chapter we present data that RhaK can act on rhamnose directly. These results also corroborate the genetic data that suggests *rhaK* encodes a determinant that carries out the first step in the catabolism of rhamnose in Rlt100 (Richardson *et al.* 2004). Our data show that rhamnose transport is dependent upon the presence of *rhaK* (Figure 3.1, Table 3.2) yet RhaK does not have transport activity in itself (Table 3.2). In addition, the wild-type strain also has a biochemical activity that is consistent with *rhaK* playing a direct catabolic role on the conversion of rhamnose (Figure 3.2, Table 3.4), and that RhaK mutants unable to carry out the biochemical activity are also unable to transport rhamnose (Table 3.5). The most straightforward explanation of the data is that RhaK plays a role in both the transport as well as the primary interconversion of the substrate. This role, however, may be either direct or indirect.

In *E. coli* the ABC protein of the maltose transport system (MalK) interacts with the transcriptional regulator MalT (Panagiotidis *et al.* 1998). This interaction leads to a repression of *mal* gene expression that will, presumably, affect transport activity over time due to a decrease of transporter components. Physiologically the effect is negated in the presence of maltose (Boos and Bohm 2000). In contrast, expression of the *rhaRSTPQUK* operon is not altered when it is measured in either a wild-type or *rhaK* mutant background (Table 3.3) (Richardson *et al.* 2004). Moreover induction of the *rhaR*, and presumably the rest of the operon, has also been shown to occur in a manner independent of *rhaK* (Richardson *et al.* 2004). This suggests that the mode of action of RhaK does not occur at the level of transcription, but must occur at either the translational or post-translational level. Although our data suggests a direct relationship between the biochemical activity of RhaK and transport of rhamnose, we cannot rule out

the possibility that RhaK is interacting with another protein component that is necessary to activate the ABC transporter. Far Western analysis and protein pull-down experiments using RhaK isolated from *E. coli* over-expressing *rhaK* have failed to identify any candidate proteins (data not shown).

Interestingly, it has been previously observed that the completed *S. meliloti* genome does not contain a recognizable phospho-*enol*-transferase (PTS) system, and that the chromosome and pSymB encode a large number of ABC transporters. The implication of these observations was that sugars were transported and subsequently phosphorylated by cytoplasmic sugar kinases (Finan *et al.* 2001, Galibert *et al.* 2001). If *R. leguminosarum* bv. *trifolii* is similar to *S. meliloti*; it is not unreasonable to assume that an analogous situation exists. However, in our case, it appears that transport and substrate modification are linked; not unlike the PTS system (Postma 1984). Although this particular phenomenon has not been previously reported for ABC transporters, it seems unlikely that this particular mechanism will be unique and it is likely that it will occur in other ABC transport systems as well. It is therefore important that the regions of the protein that predict this type of interaction can be identified.

The presence of a FGGY N motif is interesting in this respect. *E. coli* glycerol kinase has been shown to be a FGGY-type kinase and it has been crystallized in a complex with the IIA^{glc} component (Hurley *et al.* 1993). In this respect, a precedent for the interaction of an FGGY domain and a transporter complex has been reported. GlpK interacts with the PTS protein IIA^{glc} which leads to a stable complex that does not have glycerol kinase activity (Hurley *et al.* 1993). This interaction is carried out utilizing two distinct regions; a 10 amino acid region that facilitates protein-protein interaction, and a 3

amino acid region of GlpK that does not interact with IIA^{glc} directly but is at the boundary between the ATPase catalytic core and the IIA^{glc} binding domain (Pawlyk and Pettigrew 2002). Alignment of GlpK and RhaK clearly show that the region from RhaK that corresponds to the 10 amino acids (P^{495}GIETTERNY^{503} on alignment) as well as the three amino acid residues (G^{447}TR^{450} on alignment) are regions that show lower amino acid sequence conservation suggesting that these regions probably do not play similar roles in RhaK (Figure 3.4).

Interestingly GlpK has also been postulated to interact with the components of the glycerol facilitator (Jin *et al.* 1982). Further studies concluded that glycerol kinase activity was activated by interaction with the glycerol facilitator protein, GlpF, and that this may be analogous to the kinase-porin interactions found in mitochondria (Brdiczka 1990, Voegele *et al.* 1993). Heterologous expression of GlpF, clearly demonstrated glycerol uptake in the absence of GlpK (Maurel 1994). Although these examples show similarity to what we have described for RhaK, it is noteworthy to point out that unlike the glycerol transport system we do not see any transport of rhamnose in the absence of RhaK (Figure 3.1, Table 3.2, Table 3.5).

In an effort to try to identify other similar kinase proteins, we also carried out a phylogenetic analysis of RhaK (Figure 3.3). Phylogenetic analysis clearly shows that RhaK is not a unique protein and that it does cluster with other proteins that have been annotated as sugar kinases (Figure 3.3). Within this grouping, we note that the gene encoding the protein Smc03003 is also necessary for rhamnose utilization in *S. meliloti* (data not shown), and that the most closely related protein to RhaK is from *A. tumefaciens*. Utilizing a reciprocal best hit BLAST analysis we can conclude that the 4

proteins delineated in Figure 3.3 are orthologues and presumably will play a role in rhamnose utilization and perhaps affect transport activity.

The demonstration that RhaK plays a role in both catabolism and transport of rhamnose was unexpected. Whereas the initial mutational analysis targeting the γ-phosphate-binding β-loop and thus demonstrating that a functioning kinase was necessary for transport was straight forward, further characterization and delineation of the domains necessary and the mechanism of action will be more involved. Although we were able to predict the putative active site, as well as residues that may be involved in ATPase activity, we could not predict which regions may affect transport activity. To provide this, we are currently pursuing an in-depth characterization of this protein and its involvement in transport by using both a targeted and random mutagenesis of *rhaK*. Moreover with the availability of completed genome sequences of a number of closely related bacteria, it is possible to target orthologues to identify other kinases that function in catabolism as well as transport.

Chapter 4

RhaU of *Rhizobium leguminosarum* is an L-rhamnose mutarotase

4.1 Introduction

Rhizobium leguminosarum bv. *trifolii* is a gram-negative soil bacterium that can form symbiotic associations with various species of clover. The plant provides the bacteria with energy for growth and in return the *Rhizobium* provides the plant with fixed nitrogen from nitrogen fixing root nodules. Competition for nodule occupancy exists between strains of *Rhizobium* within the rhizosphere (Dowling and Broughton 1986, Triplett and Sadowsky 1992) and strains of *R. leguminosarum* unable to catabolize L-rhamnose are compromised in their ability to compete for nodule occupancy versus wild-type strains (Oresnik *et al.* 1998).

L-Rhamnose is a methyl-pentose monosaccharide found in the mucilage of a number of legume plants, and is a constituent of pectin in the form of rhamno-galacturonan within the cell walls of dicotyledonous plants (Knee *et al.* 2001, McNeil *et al.* 1984). The 11-kb L-rhamnose locus in *R. leguminosarum* encodes catabolic and L-rhamnose transport genes (Oresnik *et al.* 1998, Richardson *et al.* 2004). These genes are organized in two divergently transcribed operons regulated by a negative regulator, *rhaR* (Richardson *et al.* 2004). One transcript contains *rhaDI*, encoding a putative dehydrogenase/aldolase and an isomerase respectively, while the other consists of *rhaRSTPQUK* (Richardson *et al.* 2004). *rhaS, rhaT, rhaP, and rhaQ* encode the components of an ABC transporter (Ames 1986) including, respectively, a periplasmic sugar binding protein, an ABC ATPase and two transmembrane proteins. Within the L-rhamnose catabolic pathway, the enzymatic action of the kinase, encoded by *rhaK,* has been shown to precede the action of the dehydrogenase/aldolase and the isomerase encoded by *rhaD* and *rhaI* respectively (Richardson *et al.* 2004). Moreover, biochemical

activity of the kinase appears to be necessary for L-rhamnose transport (Richardson and Oresnik 2007).

Among the nine proteins encoded by the L-rhamnose operon only *rhaU* does not have an assigned role and data base searches revealed a number of similar ORFs, none of which had an assigned function. However, YiiL, originally annotated as a hypothetical protein from the *E. coli* genome, has recently been shown to be a rhamnose mutarotase, providing a strong clue to the identity of RhaU.

L-Rhamnose is a pyranose monosaccharide (Ryu *et al.* 2005). Pyranose and furanose monosaccharides assume a dynamic equilibrium between the α-or β- anomers in solution (Drew 1998). Whereas spontaneous α- or β-anomeric conversion can occur rapidly for furanose sugars, spontaneous α- to β-anomeric conversion of pyranose monosaccharides occurs more slowly (glucose 0.015 conversions per minute) (Levy and Cook 1954, Livingstone 1977, Swain 1952). As well, catabolic enzymes have been shown to have a preference for a specific anomeric configuration of the monosaccharide such that if the preferred anomeric form is not readily available, use of that particular monosaccharide may be less efficient at lower concentrations (Ryu *et al.* 2004, Sigrell *et al.* 1999).

Mutarotation is a catalyzed phenomenon (Kim *et al.* 2003, Ryu *et al.* 2004, Thoden *et al.* 2003). Several proteins have been identified having mutarotatase functions and three such examples include rhamnose mutarotase (YiiL in *E. coli*) (Ryu *et al.* 2004, Ryu *et al.* 2005), galactose mutarotase (GalM) (Beebe *et al.* 2003, Beebe and Frey 1998, Bettenbrock and Alpert 1998, Bouffard *et al.* 1994, Holden *et al.* 2003, Thoden and

Holden 2002, Thoden *et al.* 2002, Thoden *et al.* 2003, Timson and Reece 2003), and fucose mutarotase (FucU) (Kim *et al.* 2003, Ryu *et al.* 2004).

In this chapter we characterize *rhaU*. Utilizing genetic and physical analysis we provide data to determine the biochemical placement of RhaU and provide evidence that is consistent with RhaU being a mutarotase.

4.2 Material and methods

4.2.1 Bacterial strains and culture conditions

Bacterial strains and plasmids used and generated in this work are listed in Table 4.1. *R. leguminosarum*, strain Rlt100 (original designation W14-2 (Baldani *et al.* 1992)) and strains derived form this wild-type strain were routinely grown at 30°C on TY as a complex medium (Beringer 1974) and VMM (Vincent 1970), a defined medium as previously described (Richardson *et al.* 2004). Carbon sources were filter-sterilized and added to defined media to a final concentration of 15 mM. When required, antibiotics were added to solid or liquid media at the following concentrations: tetracycline (Tc), either 10 μg ml^{-1} or 5 μg ml^{-1}; neomycin (Nm), 200 μg ml^{-1}; streptomycin (Sm), 200 μg ml^{-1}; gentamicin (Gm), either 15 μg ml^{-1} or 30 μg ml^{-1}; ampicillin (Ap), 100 μg ml^{-1}; kanamycin (Km); 50 μg ml^{-1}. Bacterial growth was monitored spectrophotometrically at 600 nm.

4.2.2 DNA manipulations, sequencing and sequence analysis

Standard techniques were used for plasmid isolation, restriction enzyme digestions, ligations, transformations, and agarose gel electrophoresis

Table 4.1. Bacterial strains and plasmids

Strain or plasmid	Genotype or phenotype	Reference or source
Strains		
R. leguminosarum		
Rlt100	W14-2, *R. leguminosarum* Sm^r wild-type	(Baldani *et al.* 1992)
Rlt100/pMR110	wild-type/*rhaK* in pRK7813	(Richardson *et al.* 2007)
Rlt105	Rlt100 *rhaD1*::Tn*5*B20	(Oresnik *et al.* 1998)
Rlt144	Rlt100 *rhaK50*::Tn*5*B20	(Richardson *et al.* 2004)
Rlt144/pMR110	Rlt100 *rhaK50*::Tn*5*B20/*rhaK* in pRK7813	(Richardson *et al.* 2007)
Rlt146	Rlt100 *rhaK52*::Tn*5*B20	(Richardson *et al.* 2004)
Rlt211	Rlt105 *rhaK58*::pKNOCK-Tc	(Richardson *et al.* 2004)
Rlt243	Rlt100Δ *rhaU*	(Richardson *et al.* 2007)
Rlt243/pMR110	Rlt100Δ *rhaU, rhaK* in pRK7813	(Richardson *et al.* 2007)
Rlt243/pMR183	Rlt100Δ *rhaU, rhaU* in pRK7813	(Richardson *et al.* 2007)
Rlt243/pW3C1	Rlt100Δ *rhaU,* cosmid from Rlt100 wild-type cosmid containing entire *rha* locus	(Richardson *et al.* 2007)
E. coli		
MT616	MT607, pRK600	(Finan 1986)
DH5α	*end*A *hsd*R17 *sup*E44 *thi*-1 *rec*A1 *gyr*A96 *rel*AΔ(*arg*F-*lac*ZYA)U169ϕ80d*lac*ZΔM15	B.R.L. Inc.
Plasmids		
pBluescript II SK	Cloning vector, ColE1 *oriV*, Ap^r	Stratagene
pRK600	pRK2013 *npt*::Tn*9* Cm^r $Nm\text{-}Km^s$	(Finan 1986)
pRK7813	RK2 *ori*, pUC9 polylinker, λ *cos* site, Tc^r	(Jones and Gutterson 1987)
pRSETA	His_6 N-terminal tag, T7 promoter, Ap^r	Invitrogen
pW3C1	pRK7813 cosmid from Rlt100 cosmid bank, containing entire rhamnose locus	(Oresnik *et al.* 1998)
pJQ200 SK^+	Suicide vector for homogenotization; P15a *ori, mob sacB*, Gm^r	(Quandt and Hynes 1993)
pMR110	*rhaK* coding sequence in pRK7813	(Richardson and Oresnik 2007)
pMR113	*rhaU* coding sequence in pRK7813	(Richardson *et al.* 2007)
pMR133	*rhaU* coding sequence in pRSETA	(Richardson *et al.* 2007)
pMR174	Δ*rhaU* flanking sequence in pJQ200 SK^+	(Richardson *et al.* 2007)
pMR183	*rhaU* coding sequence in pRK7813	(Richardson *et al.* 2007)

(Sambrook and Russell 2001). Genomic DNA was isolated using a modified version of the protocol outlined by Meade *et al.* (Meade *et al.* 1982), as previously described (Oresnik *et al.* 1998). Sequencing reactions were accomplished using dye terminators at the University of Calgary Core DNA Services. Sequence data were analyzed using DNASIS (Hitachi Software Engineering Co., San Bruno, CA. U.S.A.). Database searches were done using the BLASTX program (Altschul *et al.* 1997).

4.2.3 Construction of plasmid containing wild-type rhaU

Construction of pMR183 was carried out by amplifying from the putative start (ATG) to stop (TGA) coding region of *rhaU* using the primers 5'HINDRBSRHAU (5' ATAT**AAGCTT**GGAGAGTCGACATGACATTGGAAAAACACGC 3') and 3' ECORI RHAU (5' ATAT**GAATTC**TCATGGCATATGGAAGAGG 3'), pW3C1 was used as a template. The amplification product was cloned as a *Hin*dIII/*Eco*R1 fragment into a broad host range plasmid pRK7813 (Jones and Gutterson 1987) by using the restriction sites that were introduced into the primers. The construct was verified by nucleotide sequencing using primers designed to flank the multiple cloning sites of pRK7813, PRK7813FORWARDPRIMER (5' GTGCTGCAAGGCGATTAAGTT 3') and PRK7813REVERSEPRIMER (5' TGTGAGCGGATAACAATTTCAC 3').

4.2.4 Over-expression and purification of RhaU

To construct an inframe His_6 tagged RhaU, the coding region was amplified and cloned inframe into pRSETA (Schoepfer 1993) using *Bam*HI and *Eco*R1 restriction sites within the primers, such that *rhaU* was under the control of the T7 promoter and inframe

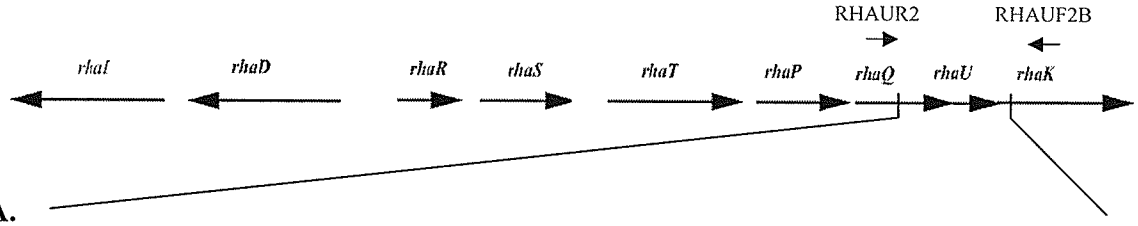
with the N-terminal His_6 tag. The primers used in the PCR amplification were 5' BAMH1RHAU (5' ATAT**GGATCC**GGAGATATGACATTGGAAAAACACGC 3') and 3' ECORI RHAU (5' ATAT**GAATTC**TCATGGCATATGGAAGAGG 3'). The sequence of the resulting construct, pMR133, was confirmed by sequencing using primers designed to flank the multiple cloning site of pRSETA, T7PROMOTER (5' AATACGACTCACTATAGGG 3') and PRSETAREVPRIMER (5' CTAGTTATTGCTCAGCGGTGG 3'). For protein expression, pMR133 was transformed into *E. coli* strain BL21 (DE3) (Miroux and Walker 1996). Cultures were grown to mid-log phase (OD_{600} = 0.5) (240 rpm), in LB ^{amp} and induced using IPTG (1 mM) at 37°C for 4 hours. Cells were harvested by centrifugation (6,000 x g for 10 minutes) and resuspended in 2 ml g^{-1} wet pellet in binding buffer (20 mM Tris pH 7.8, 300 mM NaCl, 10 mM imidazole, 5 mM β-mercaptoethanol). Cells were lysed with two passages through a French pressure cell (16,000 psi) and cell-free extracts were prepared by subsequent centrifugation of the lysate (6,000 x g for 10 minutes). His_6 tagged RhaU was immobilized by passing cell-free extract through a nickel affinity column, attached to FPLC apparatus (Bio-Rad Biologic LP chromatography system). Washing was done with wash buffer (20 mM Tris pH 7.8, 300 mM NaCl, 20 mM imidazole, 5 mM β-mercaptoethanol) for 60 minutes at a flow rate of 1 ml min^{-1}. The RhaU fusion protein was then eluted from the nickel affinity column (Bio-Rad Ni-NTA superflow) using elution buffer (20 mM Tris pH 7.8, 300 mM NaCl, 250 mM imidazole, 5 mM β-mercaptoethanol). The purity of RhaU was confirmed with SDS-PAGE gel electrophoresis (Laemmli 1970). SDS gels of eluted fractions were transferred electrophoretically to a nitrocellulose membrane for western blot analysis. The presence

of the His_6 tagged RhaU was confirmed using a His_6 monoclonal antibody as a primary antibody and a goat-anti-mouse secondary antibody conjugated with HRP. HRP was detected colormetrically with an opti-4CN substrate detection kit as suggested by the manufacturer (Bio-Rad Laboratories Hercules Ca.).

4.2.5 Generation of ΔrhaU strain

The generation of Δ*rhaU* strain was accomplished using the site-specific mutagenesis by overlap extension PCR method (Ho *et al.* 1989, Sambrook and Russell 2001), with pW3C1 as a template. The primers used for site-specific mutagenesis by overlap extension were R2 primer, RHAUR2 (5' AA**GGATCC**GGTCAGGGCTATGTCGTC 3'), F2 primer, RHAUF2B (5' AA**CTGCAG**GCAGCCGAGAGAGGTCAA 3') flanking primers. The mutagenic primers were FM primer, RHAUFM (5' TGGCGGTCATCGGGAGCTCATGAGCTTG 3'), and RM primer RHAURM (5' TGAGCTCCCGATGACCGCCAGTTCCTATC 3'). Using restriction sites introduced into the flanking primers the amplification product was cloned as a *Bam*HI/*Pst*1 fragment into pJQ200SK (Quandt and Hynes 1993). The subsequent clone was called pMR174. Gene replacement of Rlt100 *rhaU* with pMR174 (Δ*rhaU*), was carried out as previously described (Quandt and Hynes 1993). The resulting strain was called Rlt243. Correct replacement of the wild-type gene with deleted *rhaU* in strain Rlt243 was verified by nucleotide sequencing using Rlt243 genomic DNA and RHAUR2 and RHAUF2B flanking primers (Figure 4.1).

4.2.6 Transport assays



A.

rhaQ...***ATCAAGCTCATGAGCTCCCGATGA***CATTGGAAAAACACGCCTTCAAGATGCAGCT
CAATCCCGGCATGGAAGCCGAATACCGCAAGCGGCATGACGAGATCTGGCCGGAGCTGG
TCGATCTGCTGCACCAGTCGGGCGCCAGCGACTATTCGATCCATCTCGACCGCGAGACC
AACACGCTGTTCGGCGTGCTGACGCGGCCGAAGGACCACACGATGGCGAGCCTGCCGGA
CCATCCCGTCATGAAGAAATGGTGGGCGCACATGGCCGACATCATGGCGACGAACCCCG
ATAATTCGCCCGTCCAGAGCGACCTCGTCACCCTCTTCCATATGCC***ATGACCGCCAGTT***
CCTATCGCCG...***rhaK***

B.

Stop

rhaQ...***ATCAAGCTCAAGAGCTCCCGATGACCGCCAGTTCCTATCGCCG***...***rhaK***

Start

Figure 4.1. Construction of *rhaU* deletion. Sequence of **A.** wild-type (Rlt100) *rhaQUK* and **B.** *rhaU* deletion strain (Rlt243) generated by site-specific mutagenesis by overlap extension PCR. The sequence corresponding to *rhaQ* and *rhaK* are bold and italicized. *rhaQ* termination site and *rhaK* initiation site that overlap *rhaU* initiation and termination sites respectively, are underlined in both **A.**, and **B.** Schematic representation of the rhamnose catabolic operon is drawn above with corresponding flanking primers, RHAUR2 and RHAUF2B used for sequencing.

Uptake of rhamnose was carried out essentially as previously described (Richardson *et al.* 2004). Radioactive [^{3}H] rhamnose (5 Ci/mmol) was purchased from American Radiolabeled Chemicals Ltd. St. Louis, Missouri. Transport assays were initiated by the addition of tritiated rhamnose to a final concentration of 2 μM (125,000 dpm) 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 the defined salts medium and the radioactivity that remained on the filter was determined using a liquid scintillation spectrophotometer (Beckman LS6500). Samples were taken every 15 or 20 seconds and continued for up to one minute. Initial rates were calculated from the linear portion of the data (seconds).

4.2.7 In vitro RhaK assays

Cultures were grown on minimal medium to mid-log phase, harvested by centrifugation (6,000 x g for 10 minutes) and resuspended in 2 ml g^{-1} wet pellet in lysis buffer. Resuspended cells were lysed by 2 passages through a French pressure cell (16,000 psi) and cell debris was removed by centrifugation (10,000 x g for 10 min at 5°C in a Sorval SLA-3000 rotor). Induced cell-free extracts were combined with 50 mM HEPES buffer pH 7.6, 5 mM $MgCl_2$, 10 mM ATP, and 2 mM tritiated rhamnose and incubated at 30°C. Assays were carried out in a final volume of 100 μl. 10 μl samples were taken at 5, 10, 20, 40-minute time points and quenched by placing each sample in a boiling water bath for 10 minutes, then into an ice bath until all samples were ready to be loaded. Reactions were separated on Whatman DE81 cellulose chromatography paper

that had previously treated with 1 M formic acid, then extensively washed with ddH_2O to remove any impurities and finally dried prior to use. 10 µl quenched aliquots were spotted and separated by descending paper chromatography using 50 mM formic acid pH 3.2. To determine the position of the label, the chromatography paper was cut into 1 inch square pieces, and assayed using a liquid scintillation spectrophotometer (Beckman LS6500).

4.2.8 Generating a multiple sequence alignment using PRALINE

Amino acid sequence of RhaU identified by BLASTP identity scores (Altschul *et al.* 1997), and those deemed to be the most closely related to RhaU, by phylogenetic analysis were selected and compared to YiiL of *E. coli* (Ryu *et al.* 2004) using PRALINE multiple sequence alignment program in the default settings accessed at http://ibi.vu.nl/programs/pralinewww/help.php (Heringa 1999, Heringa 2000, Heringa 2002, Simossis and Heringa 2003, Simossis and Heringa 2004, Simossis and Heringa 2005).

4.2.9 RhaU phylogenetic tree construction

rhaU-like sequences were selected based on homology to the RhaU amino acid sequence using BLASTP (Altschul *et al.* 1997). The sequences were aligned using CLUSTAL X (Thompson *et al.* 1997). Programs contained within PHYLIP Version 3.6a (Felsenstein 1985, Felsenstein 2002) were used for phylogenetic analysis. Divergence (or distance) between two sequences was calculated by PROTDIST (setting: JTT). The resulting distance matrix was used to construct a phylogenetic tree with the NEIGHBOR

program (NJ option), and the phylogenetic tree was evaluated using the bootstrap procedure (SEQBOOT, 100 replicates). Bootstrap replicates were analyzed with CONSENSE program to generate the majority rule consensus tree. The phylogenetic tree was drawn using the TREEVIEW (Page 1996) program using the PHYLIP output file as an input file. *Streptomyces coelicolor* A3 was selected as the comparative out-group.

4.3 RESULTS

4.3.1 Generation of Rlt243 ΔrhaU strain

Previous random transposon mutagenesis of *R. leguminosarum* rhamnose locus yielded a total of 56 independent mutants, none of which provided a lesion within *rhaU* (Richardson *et al.* 2004). In order to determine if the product of *rhaU* was required for growth on or the transport of L-rhamnose, site-specific mutagenesis by overlap extension PCR (Higuchi *et al.* 1988, Sambrook and Russell 2001) was used to generate the *rhaU* deletion strain, Rlt243. Nucleotide sequencing of the *rhaQ/rhaK* region in Rlt243 confirmed that no inadvertent mutations had been introduced and that the predicted upstream *rhaQ* termination site and predicted downstream *rhaK* start site were not interrupted (Figure 4.1).

4.3.2 Rlt243 exhibits a slow growth phenotype on L-rhamnose

Sequence similarity suggested that RhaU was related to a group of hypothetical proteins that included YiiL of *E. coli* which was recently characterized as a rhamnose mutarotase, catalyzing the α- to β-anomeric conversion of L-rhamnose (Ryu *et al.* 2004,

Ryu *et al.* 2005). Due to the amino acid similarity within the predicted catalytic core we wished to determine if RhaU had a physiological function similar to YiiL.

We have previously shown that all of the other predicted genes in L-rhamnose catabolic locus were necessary for growth on L-rhamnose. Based on these observations we predicted that strains lacking *rhaU* would also be unable to grow using L-rhamnose as a sole carbon source. Following three days of growth on VMM L-rhamnose agar, Rlt243 appeared to grow slowly compared to the wild-type strain (Table 4.2). The slow growth phenotype for Rlt243 was different than the observations reported for *yiiL* mutant strain in *E. coli*. For strain Rlt243 the slow growth phenotype on L-rhamnose was observed under typical growth condition (0.14% L-rhamnose), whereas this phenotype was only observed at low L-rhamnose concentration (0.03%) for *yiiL* mutant strains (Ryu *et al.* 2005). Interestingly, after four days growth strains Rlt243 and Rlt100 were indiscernible on plates (Table 4.2).

To ensure that the growth phenotype for Rlt243 was not attributed to a pleiotropic effect within this strain, Rlt243 was also grown on complex media (TY) and VMM glucose. Strains Rlt243 and Rlt100 had no discernable differences in growth, suggesting the slow growth phenotype on VMM L-rhamnose was specific to L-rhamnose and not due to an overall reduced growth potential of the strain.

To quantitate the slow growth phenotype observed for Rlt243, a growth curve for strains Rlt100 and Rlt243 was generated (Figure 4.2). The mean growth rates for Rlt243 and Rlt100 were 0.0563 and 0.109 generations per hour, respectively (Figure 4.2). To substantiate that this was not due to a delay of growth, the strains were subcultured numerous times, yielding identical growth rates. Cultures from several trials of the

Table 4.2. Growth of *Rhizobium* strains with a representative of each on glucose, rhamnose, and glycerol-rhamnose as carbon sources [a b]

Strain	Relative Genotype	VMM glycerol	VMM glycerol/rhamnose	VMM rhamnose
Rlt100	wild-type	+	+	+
Rlt105	*rhaDI*$^-$	+	-	-
Rlt144	*rhaK*$^-$	+	+	-
Rlt144/pMR110	*rhaK*$^-$/*rhaK*$^+$	+	+	+
Rlt211	*rhaDI*$^-$/*rhaK*$^-$	+	+	-
Rlt243	Δ*rhaU*	+	+	±
Rlt243/pMR110	Δ*rhaU*/*rhaK*$^+$	+	+	+
Rlt243/pW3C1	Δ*rhaU*/*rha* locus$^+$	+	+	+
Rlt243/pMR183	Δ*rhaU*/*rhaU*$^+$	+	+	±

[a] Designations are as follows: +, like wild-type growth; -, no growth; ±, Denotes slower growth vs. Rlt100

[b] 15 mM concentrations of each carbon source was added to defined media

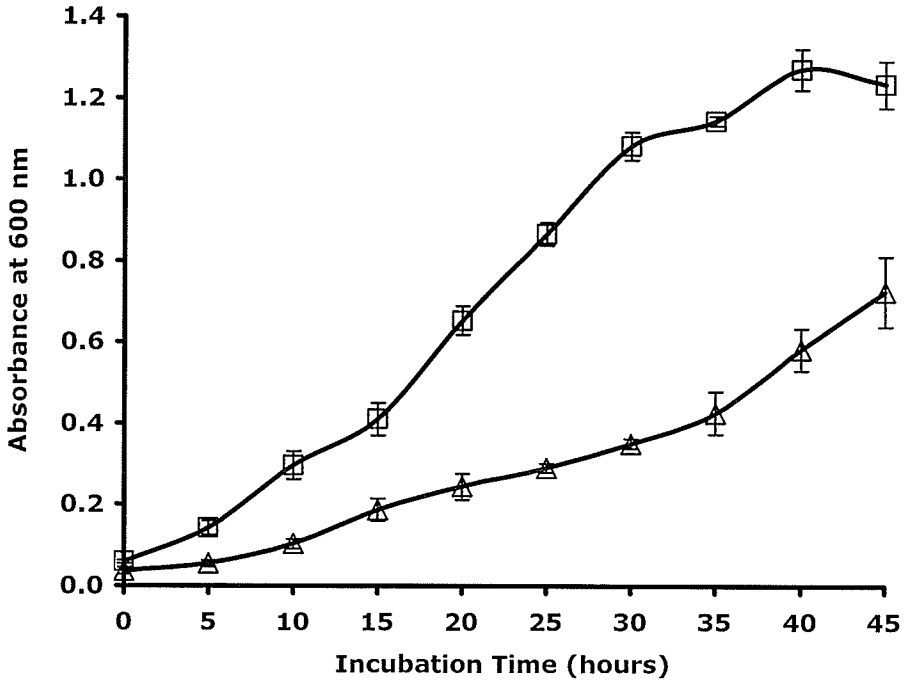


Figure 4.2. Growth curve of *R. leguminosarum* wild-type and *rhaU* deletion strains. Growth curve of wild-type (Rlt100) and *rhaU* deletion (Rlt243) *R. leguminosarum* strains as measured by absorbance at wavelength 600 nm. Strains were grown on minimal media supplemented with 15 mM rhamnose as a sole carbon source. Symbols are as follows: square, Rlt100; and triangle, Rlt243.

growth curve were also allowed to continue growing for greater than 96 hours until a maximum absorbance was reached. Both Rlt100 and Rlt243 reached near identical final optical densities suggesting that carbon utilization was the same in both the wild-type and Rlt243 strains, confirming that there was no lack of growth potential associated with strain Rlt243. Furthermore, an identical slow growth phenotype was observed after every round of repeated streaking on VMM L-rhamnose medium, confirming that the phenotype was not a result of up-regulation of some uncharacterized rhamnose catabolic genes, or the presence of a second site mutation. Plasmid pW3C1 encoding wild type *rhaU* complemented the slow growth phenotype confirming that the slow growth phenotype in Rlt243 was a result of Δ*rhaU* (Table 4.2)

4.3.3 Placement of RhaU in a biochemical pathway

RhaU would most reasonably be situated in a metabolic pathway to catalyze anomerization after transport into the cell prior to phosphorylation. A biochemical lesion in a catabolic pathway following the synthesis of a phosphorylated sugar intermediate can give rise to a conditional phenotype if the bacterium is grown on a medium that contains a non-inducing carbon source that can be catabolized (Adhya and Shapiro 1969). Using this logic it has been shown genetically that the action of RhaK occurs prior to RhaD and RhaI in *R. leguminosarum* the conditional phenotype of the *rhaD/I*-containing Rlt105 was relieved by introduction of *rhaK* in Rlt211 (Richardson *et al.* 2004). The conditional phenotype observed on glycerol-rhamnose plates for the strain Rlt105 was due to accumulation of a phosphorylated intermediate generated by the product of *rhaK* (Table 4.2). Whereas a *rhaD/I* mutant was unable to grow on glycerol-rhamnose, the

addition of a *rhaK* mutation into a *rhaD/I* strain generating *rhaDI/rhaK* double mutant strain Rlt211 was able to grow on glycerol-rhamnose plates eliminating the conditional phenotype.

We wished to address whether the action of RhaU occurred before or after the action of RhaK. Strain Rlt243 grew at wild-type levels on glycerol-rhamnose plates, implying that RhaU enzymatic action might occur before rhamnose is phosphorylated (Table 4.2). However, a *rhaU*-containing strain may simply have been less sensitive to the accumulation of phosphorylated rhamnose intermediates, and to test this possibility, RhaK levels were increased by introducing a *rhaK*-encoding plasmid into the *rhaU* deletion strain. The fact that even this RhaK over producing strain, Rlt243/pMR110, did not exhibit a conditional phenotype is consistent with RhaU working prior to RhaK in the catabolic pathway. Interestingly, the over-expressed RhaK had the unexpected effect of eliminating the slow growth phenotype on VMM rhamnose (Table 4.2).

To demonstrate that the inability of Rlt243 to grow at wild-type rates on L-rhamnose was due to the deletion of *rhaU*, pMR183 and pW3C1 were introduced into Rlt243. Introduction of pMR183 did not result in the strain being complemented for its ability to grow on rhamnose as a sole carbon source at levels comparable to that of the wild-type Rlt100. Rlt243/pMR183 appeared to have the same delayed growth phenotype as Rlt243 (Table 4.2). The introduction of pW3C1 however, did complement Rlt243 (Table 4.2).

To address the inability of pMR183 to complement the delayed growth phenotype of Rlt243, pMR183 was re-sequenced. The resultant sequence data showed that *rhaU* from pMR183 was identical to wild-type Rlt100 *rhaU* and that the spacing between the

ribosomal binding site and the start codon were as designated. It is unclear why pMR183 did not complement Rlt243.

4.3.4 rhaU is not necessary for the uptake of L rhamnose

Since the slow growth phenotype of the *rhaU* deletion strain was complemented by plasmid encoded wild type *rhaK*, it was of interest to determine if RhaK dependent L-rhamnose phosphorylation and transport were occurring or were altered in Rlt243 or derivatives carrying *rhaU*, *rhaK*, or the entire rhamnose locus. The transport rates for Rlt100, Rlt243, Rlt243/pMR183 and Rlt243/pWC3C1 appear to be comparable to wild-type transport rates. This suggests that uptake of Lrhamnose was normal in the absence of RhaU and that extra copies of RhaU did not enhance uptake. By contrast, transport in the *rhaK* mutant strain, Rlt144, was undetectable, but could be rescued by the plasmid encoded *rhaK* in Rlt144/pMR110. Strain Rlt243/pMR183 exhibited no additional ability to transport L-rhamnose in the presence of additional copies of *rhaU* on a plasmid suggesting transport rates were not increased due to extra copies of *rhaU* (Table 4.3).

The dependence on RhaK for L-rhamnose transport prompted us to add pMR110 to strains Rlt100 and Rlt243 to determine if extra copies of *rhaK* could have a cumulative effect on the rate of L-rhamnose imported. This did not increase the rate of transport for either Rlt100/pMR110 or Rlt243/pMR110 (Table 4.3). The data do however show that the construction of the *rhaU* deletion in Rlt243 using site-specific mutagenesis by overlap extension PCR had no ill effects of transcription and translation of *rhaK* as Rlt243 transport rates are similar to wild-type. If *rhaK* was abolished by this construction the transport rates would resemble Rlt146 (Table 4.3).

Table 4.3. *rhaU* is not necessary for the transport of tritiated L-rhamnose

Strain	Relevant Genotype	Transport (nmol/gFW/min)[a]
Rlt100	wild-type	30.7 ± 1
Rlt100/pMR110	wild-type/*rhaK*$^+$	24.7 ± 6.6
Rlt144	*rhaK*$^-$	< 0.05
Rlt243	Δ*rhaU*	36.9 ± 6.4
Rlt243/pMR183	Δ*rhaU*/*rhaU*$^+$	29.8 ± 4.0
Rlt243/pMR110	Δ*rhaU*/*rhaK*$^+$	28.3 ± 6.7
Rlt243/pW3C1	Δ*rhaU*/*rha*$^+$	28.5 ± 4.5

[a] Strains were grown as broth cultures in minimal medium (VMM) supplemented with 15 mM rhamnose/glycerol. Initial transport rates were determined using tritiated rhamnose. The data is presented as the mean ± standard deviation (n = 3).

4.3.5 Kinase activity is not affected in Rlt243

Similarly, we characterized kinase activity in the absence of RhaU (Table 4.4). It was determined that Rlt100 had an activity level of 426 μmol/min/mg, whereas a *rhaK* mutant, Rlt146 had background levels of activity. Kinase activity of Rlt243 was comparable to wild-type, the addition of pMR110 or pW3C1 resulted in kinase activities that were also comparable to wild-type activity levels. Taken together the *in vitro* data show that Rlt243 has the ability to phosphorylate rhamnose *in vivo*.

4.3.6 RhaU over-expression and purification

To further investigate the role that RhaU has in L-rhamnose catabolism in *R. leguminosarum*, a biochemical approach was initiated. Since the nearest homologue to *rhaU* was annotated as a hypothetical ORF, it was reasoned a logical starting point would be to demonstrate that *rhaU* indeed had the capability of encoding a protein. An inframe His_6 tagged version of RhaU under the control of a T7 promoter was constructed and verified. Induced cell-free extract of BL21 (DE3) containing pMR133 was passed through a nickel affinity column and RhaU was eluted as described in materials and methods. Analysis of the eluted fractions showed the presence of a doublet, the lower two bands migrated with an apparent mass of 12,000 and 14,000 Da, while the smaller upper band appeared to be a dimer with a mass of approximately 28,000 Da (Figure 4.3A lanes 2 and 3). Because these sizes were smaller than the predicted 16,595 Da, the protein was analyzed by mass spectrometry confirming a deconvoluted monomer size of 16,593 and providing no evidence of a second protein (P. Loewen personal communication). The purity of the His_6 tagged RhaU protein was confirmed with SDS-PAGE gel

Table 4.4. Kinase activity is not RhaU dependant

Strain	Relevant Genotype	Activity (μmol/min/mg)[a]
Rlt100	wild-type	426 [b]
Rlt100/pMR110	wild-type/*rhaK*$^+$	300
Rlt146	*rhaK*52	n.d.[c]
Rlt243	Δ*rhaU*	382
Rlt243/pMR110	Δ*rhaU*/*rhaK*$^+$	461
Rlt243/pW3C1	Δ*rhaU*/*rha*$^+$	479
Rlt243/pMR183	Δ*rhaU*/*rhaU*$^+$	339 [b]

[a] Strains were grown in minimal medium (VMM) supplemented with 15 mM rhamnose/glycerol, induced cell extract was isolated and processed as described in materials and methods. Values presented as one representative set of data.
[b] Data represents average of two independent replicates.
[c] n.d. not detected.

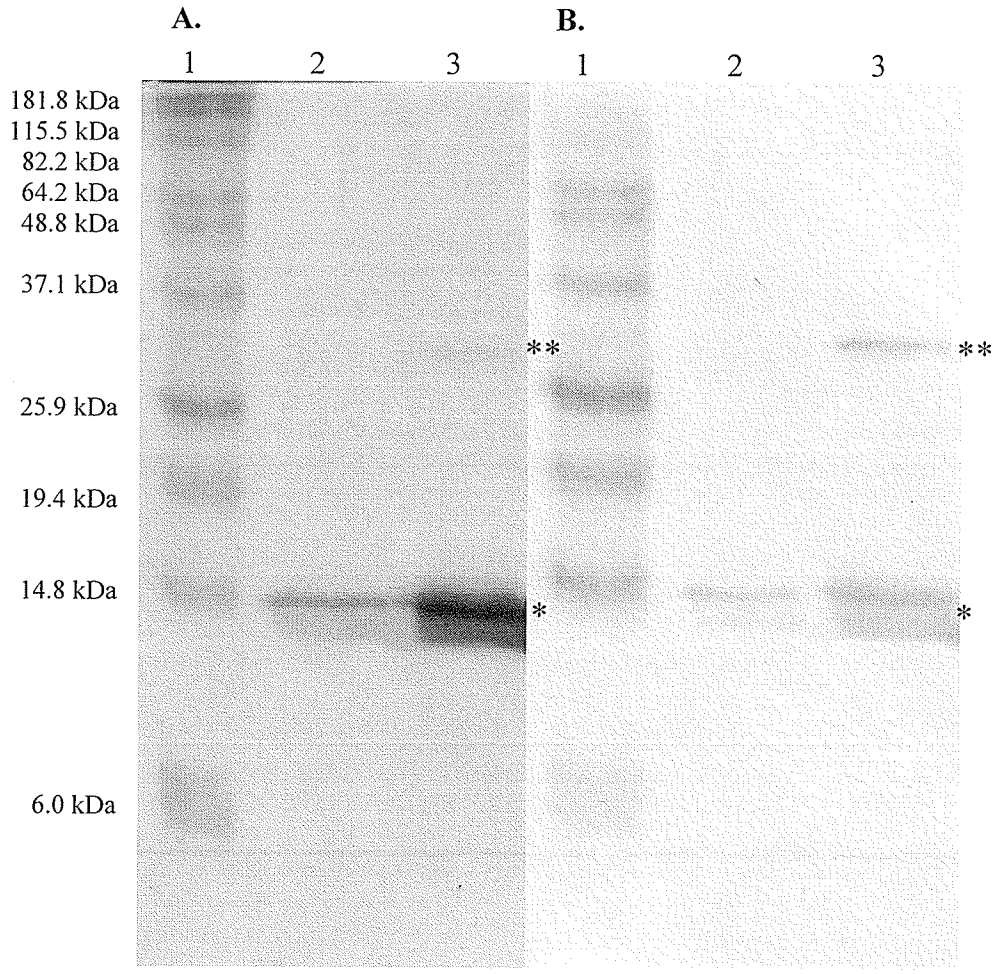


Figure 4.3. SDS polyacrylamide gel and associated western blot of RhaU. **A.** 15% SDS polyacrylamide gel. **B.** Western blot. Lane1 of both gel and blot loaded with Invitrogen benchmark prestained ladder, lanes 2 and 3 are loaded with successive RhaU elution fractions collected from a nickel affinity column. Monomeric RhaU is designated by asterisk (*), RhaU dimer designated by double asterisk (**).

electrophoresis (Figure 4.3). It should be noted that even under denaturing conditions a band of approximately 30 KDa corresponding to what would be expected for a RhaU dimer was also observed. To confirm this hypothesis, the elution fractions were western blotted (Figure 4.3B). The results support the hypothesis that RhaU is a protein and it exists as a dimer. Moreover, this data is corroborated by mass spec data and crystal structure that also show RhaU exists as a dimer and mass spectrum of RhaU also revealed that RhaU existed predominantly as a dimer and increasing ionization voltage caused its dissociation to monomer (Figure A.1.11, Figure A.1.12) (P. Loewen personal communication). This is similar to the dimer structures reported for YiiL (Ryu *et al.* 2005). We do not have an explanation for the existence of the two smaller bands of protein on the SDS gel.

4.3.7 Generation of phylogenetic tree and alignment of significant amino acid residues

Phylogenetic analysis of amino acid sequences identified by BLASTP as the most closely related to RhaU was accomplished as described in materials and methods. The resulting neighbor-joining tree with associated bootstrap values and divergence was generated (Figure 4.4). Predicted proteins related to RhaU were found within several diverse species. Several observations can be made based on this tree; there was a clade in which putative proteins were from several species that are closely related to *R. leguminosarum* including *R. etli, A. tumefaciens, M. loti* and, *S. meliloti* in addition to others that are more distantly related, these likely represent rhamnose mutarotases in the organisms or possibly mutarotases that act on other sugars. There was an *Agrobacterium* species and two *Brucella* species grouped within a different clade possibly representing a

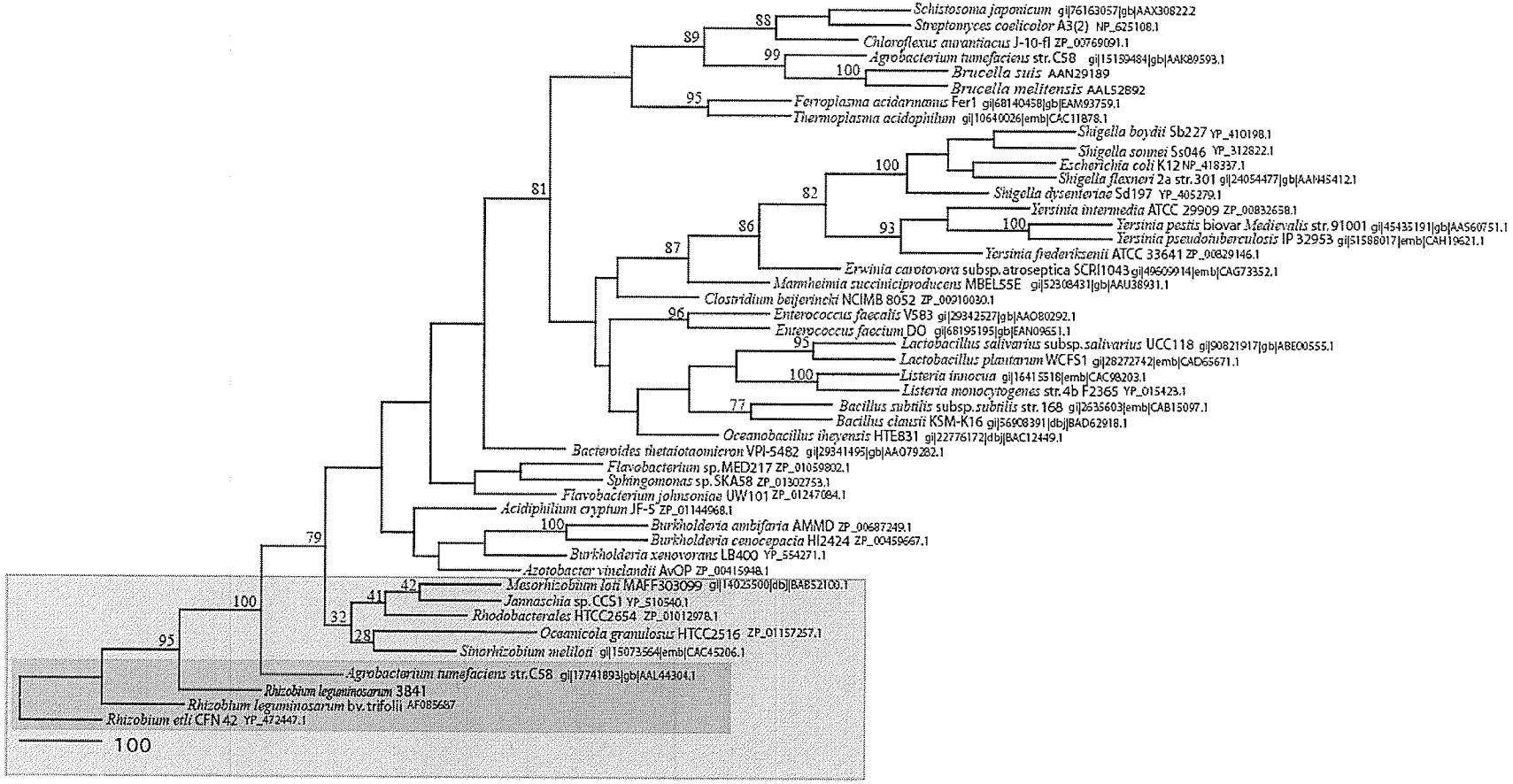


Figure 4.4. RhaU phylogenetic tree based on predicted *rhaU*-like sequences. Related sequences were identified using BLASTP and the tree was constructed as described in materials and methods. Sequence accession numbers for the predicted sequences are listed next to the corresponding strains. Bootstrap values less than 70 are not shown except in the shaded clade containing *rhaU* and predicted closest neighbors.

mutarotase of another sugar other L-than rhamnose. The *Brucella* species are however not found within the RhaU clade. Many enteric species including YiiL of *E. coli* group together, and appear to diverge from the RhaU clade.

Based on this divergence we generated a PRALINE amino acid alignment (Simossis *et al.* 2005) using the species from the RhaU clade in addition to YiiL of *E. coli* to determine what degree of amino acid residue conservation was maintained (Figure 4.5). While RhaU, YiiL, and the other putative proteins did not share complete residue conservation the aligned proteins were over 40% similar amino acid conservation over the entire length, most of the residues identified as critical for enzymatic action (Ryu *et al.* 2005) were conserved, suggesting the role of these proteins may also be conserved.

4.4 Discussion

The rationale for generating a Δ*rhaU* strain was to determine if RhaU was involved in rhamnose catabolism and if so, insight into the role of RhaU in L-rhamnose catabolism in *R. leguminosarum*. Strains lacking *rhaU* exhibit slower growth on L-rhamnose compared to wild-type strains. This growth phenotype is evident on plates as well as in broth, the phenotype is consistent with the rhamnose mutarotase activity ascribed to RhaU which would be expected to speed up L-rhamnose catabolism but not be essential. In contrast to *E. coli* where the slow growth phenotype of a *yiiL*-containing mutant strain was observed only at low L-rhamnose concentrations (Ryu *et al.* 2005) the phenotype was evident with typical growth conditions in *R. leguminosarum*. Elevated RhaK levels mask the slow growth phenotype of the *R. leguminosarum rhaU* mutant at higher rhamnose concentrations this suggests that the difference in the phenotypes in *E.*

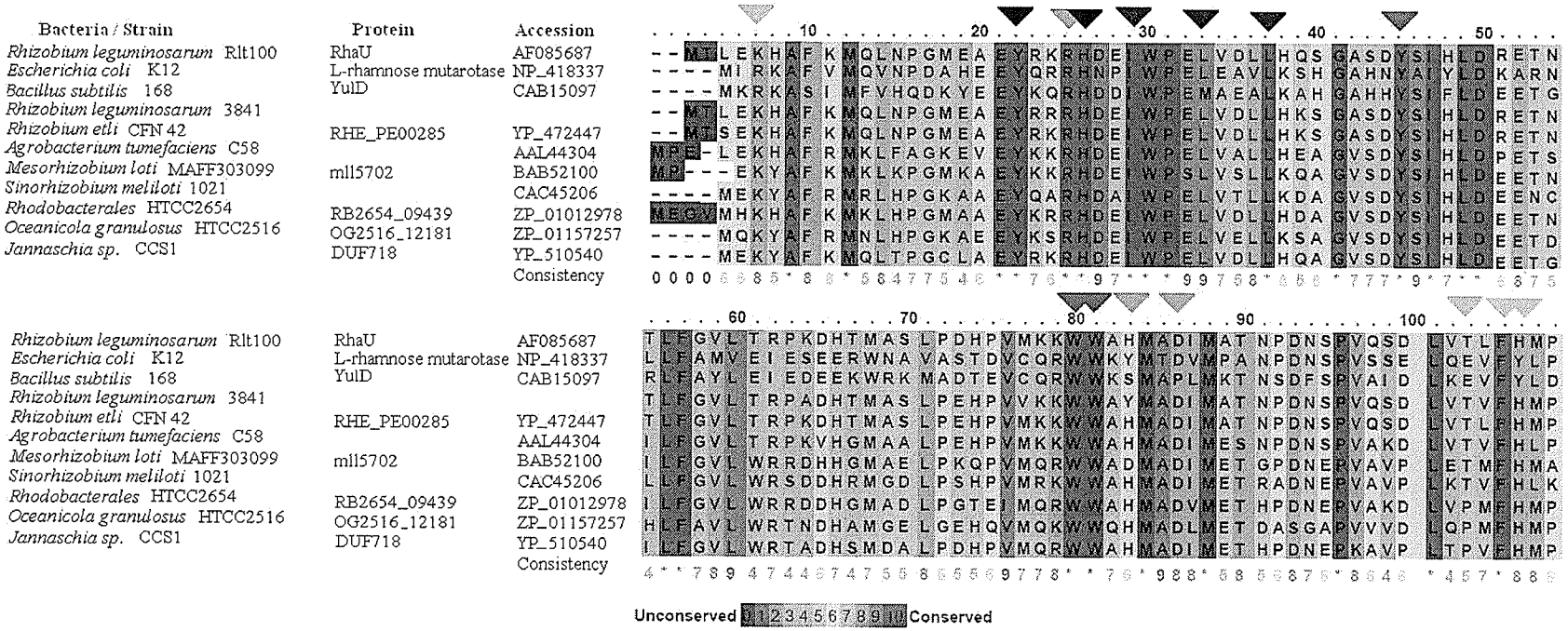


Figure 4.5. Predicted amino acid conservation alignment of RhaU. RhaU-related mutarotases performed using PRALINE. The bacteria, strain, the protein or predicted protein if available and the accession number was as described. The amino acid consistency scoring ranges from 0 (least conserved) to 10 (invariant, denoted as *). The amino acid consistencies are color-coded and correspond to numbers below the alignment. Arrow heads represent YiiL (Ryu *et al.* 2005) amino acid residue involved in mutarotation, the colors are as follows: catalytic residues (black), residues in catalytic core contacting rhamnose (red), residues within core (dark blue), triad one (green), triad two (light blue), contact with methyl group of L-rhamnose (purple).

coli yiiL mutant might be due to lower RhaK levels in *R. leguminosarum* compared to *E. coli*. Thus, the uncatalyzed α- to β-anomeric conversion is sufficient to support a low rate of growth in *R. leguminosarum* and the role of RhaU is to facilitate the anomerization thereby allowing a faster rate of growth. The equilibrium ratio of α and β anomers in L-rhamnose, determined by NMR to be 56:44 (Ryu *et al.* 2005), suggests a relatively low energy barrier for α to β anomerization, and displacement of the α to β anomerization equilibrium by phosphorylation of the β-anomer or a subsequent metabolite is not unexpected.

Prior to the classification of YiiL as a mutarotase, we observed that several organisms closely related to *R. leguminosarum*, contained predicted ORFs of similar size to *rhaU* and were identified as hypothetically conserved proteins with no ascribed function. *Rhizobium elti*, *Agrobacterium tumefaciens*, *Mesorhizobium loti* and *Sinorhizobium meliloti* ORFs potentially encoded proteins that shared 97%, 76%, 69%, and 64% amino acid identity to RhaU, respectively. Initial BLAST searches did not identify *yiiL* as a significant match to *rhaU*, but a CLUSTAL W sequence alignment suggested a 48% similarity at the nucleotide level. The overall similarity at the amino acid level between *rhaU* and *yiiL* is low. A PRALINE alignment that aligns on the basis of conservation of functional domains showed that the amino acid alignment was 40% conserved between these proteins (Figure 4.5). There is however conservation of residues predicted to form the catalytic core of YiiL and RhaU (Figure 4.5) (Ryu *et al.* 2005). Results of the analysis of RhaU crystals (Table A.1.10) soaked with L-rhamnose was communicated personally by P. Loewen, it was predicted that RhaU initially binds the α-anomer of L-rhamnose with contact points made at OH-2, OH-3, OH-4, and O5 (Figure

A.1.13). Here α-OH-1 is placed in an unfavorable position with OH-1 just 3.0 Å from the δ-CH_3 of Ile45. This interaction forces α-OH-1 to move towards Tyr20. Hydrogen bonding followed by protonation of O5 of L-rhamnose by His24 leads to sugar ring opening. Protonation of the His24 imidazole is favored by the second imidazole hydrogen bond with the main chain C=O of Tyr20. This reaction is stabilized by a hydrogen bond between O5 and the indole of Trp78. In the electron density map the bound β-anomer rhamnose OH-1 is about 3 Å from Tyr20 with good geometry for hydrogen bond formation (Appendix A.1.10, A.1.11). The proposed mechanism of anomerization in YiiL differs slightly from the mechanism proposed for RhaU. YiiL mechanism is dependent upon the ability of His-22 (corresponding to RhaU His-24) to donate and accept protons from L-rhamnose O-5 oxygen, and C-1 carbon atom hydroxyl moiety, in concert with a stabilization effect by Tyr-18 (corresponding to RhaU Tyr-20) hydroxyl on the N-1 imidazole ring of His-22 via a hydrogen bond (Ryu *et al.* 2005). This interaction may aid in stabilizing the short α-helix in addition to sustaining the α- and β-OH-1 orientations during catalysis (Serrano *et al.* 1992). Phe101 that was proposed to be involved in interacting with the methyl moiety of L-rhamnose in YiiL (Ryu *et al.* 2005). In RhaU the L-rhamnose methyl group fits in hydrophobic pocket composed of Ile27, Leu35, Leu31 and Phe103 (P. Loewen personal communication). Overall, most of the residues believed to be essential in enzyme function in YiiL are conserved within RhaU and all of the putative proteins suggested within the phylogenetic tree (Figure 4.4, Figure 4.5). This would suggest that even though the phylogenic analysis implies divergence of the proteins has occurred, conservation of essential residues enabling the catalytic role of α- to β-anomeric conversion of L-rhamnose in RhaU and YiiL are conserved.

In attempting to predict the order of enzymatic action within the L-rhamnose catabolic pathway three distinct models can be considered. The first, RhaU is a periplasmic protein and functions prior to transport (Figure 4.6). The second, RhaU activity occurs prior to RhaK dependent phosphorylation (Figure 4.7). The third, RhaU occurs following RhaK dependent phosphorylation (Figure 4.8).

Although periplasmic mutarotases like GalM, a galactose mutarotase of *Acinetobacter calcoaceticus* have been proposed (Bouffard *et al.* 1994, Gatz *et al.* 1986, Gatz and Hillen 1986) this is unlikely for RhaU for several reasons. Firstly, proteins whose fate is translocation into the periplasm often have a leader sequence that targets that protein for translocation, RhaU unlike GalM, does not contain a leader targeting sequence. Secondly, RhaU is not related to mutarotases that are periplasmically localized. Additionally, if RhaU occurred prior to transport, transport rates would be expected to be lower than wild-type, this however, is not observed (Table 4.3).

The placement of RhaU activity either before or after RhaK catalysis is complicated by the fact that RhaK appears to be necessary for transport and phosphorylation. If RhaU functions before phosphorylation as in the *E. coli* L-rhamnose catabolism pathway, RhaK will have to function in two steps separated by the RhaU mutarotation step. Conversely, placement of RhaU following phosphorylation would allow RhaK to carry out transport and phosphorylation as part of the same step, but require RhaU to isomerize a phosphorylated sugar.

We observed wild-type transport rates and accumulation of phosphorylated rhamnose in the absence of RhaU (Table 4.3, Table 4.4). This data appears to support the theory that RhaU mutarotation activity occurs after transport and phosphorylation.

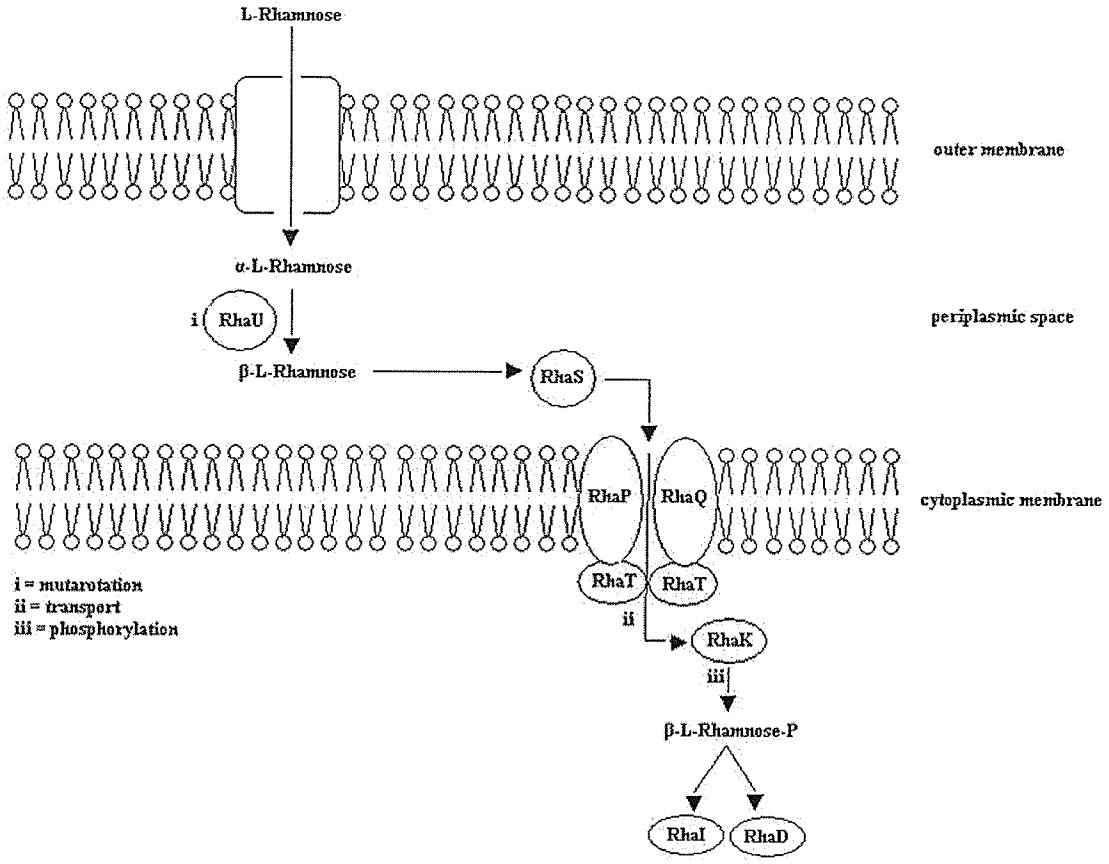


Figure 4.6. Schematic model for biological placement of RhaU in the periplasmic space.

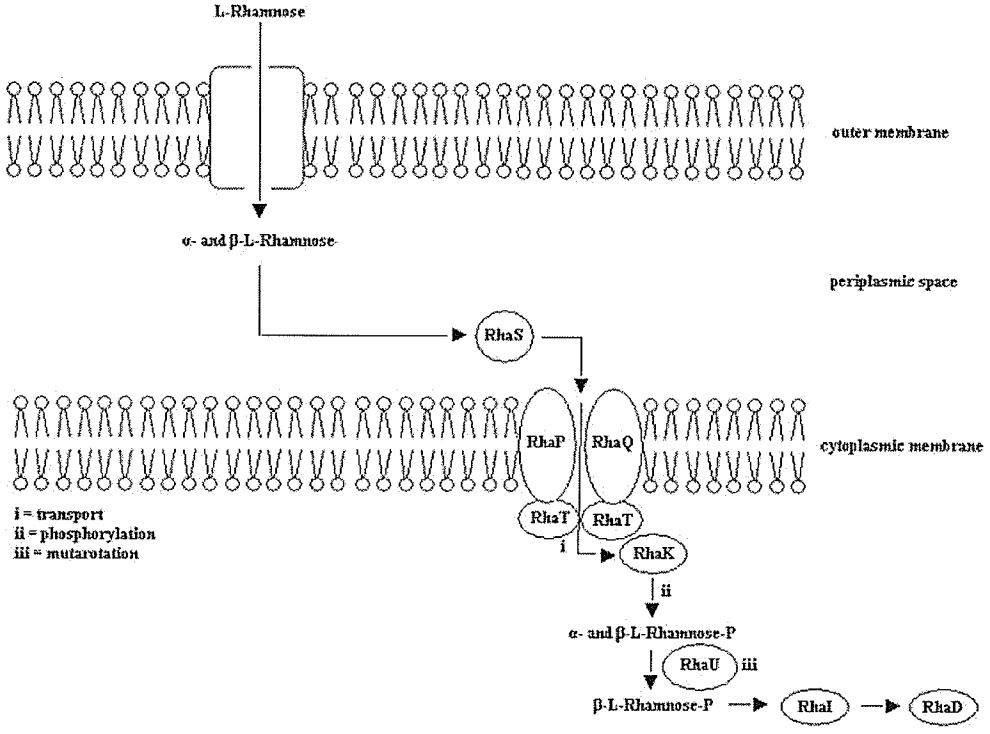


Figure 4.7. Schematic model for biological placement of RhaU occurring post-phosphorylation.

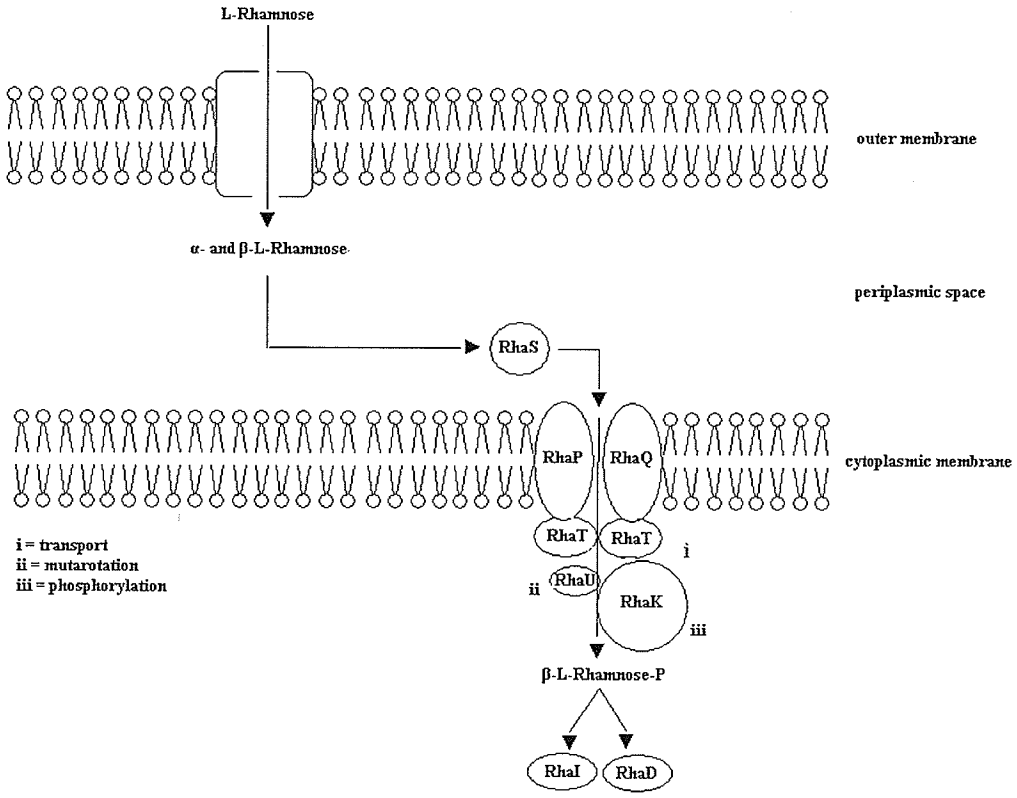


Figure 4.8. Schematic model for biological placement of RhaU occurring in concert with transport and phosphorylation.

However, several other lines of evidence argue in favor of mutarotation activity occurring before phosphorylation, for instance *rhaU* mutants grow normally on L-rhamnose-glycerol and that unphosphorylated L-rhamnose binds to RhaU (Figure A.1.12, Figure A.1.13) (P. Loewen personal communication) suggest strongly that RhaU functions prior to phosphorylation. Additionally, elevated levels of RhaK mask the slow growth phenotype of *rhaU* mutants consistent with the displacement of the $\alpha \approx \beta$ equilibrium towards the β-anomer by the phosphorylation step removing the β-anomer from the equilibrium. Although it is possible that RhaU activity occurs following RhaK phosphorylation, the most straight forward explanation of the experimental data favors its placement prior to RhaK phosphorylation.

In summary, this work has shown *rhaU* to encode a protein involved in L-rhamnose catabolism. Moreover, the bioinformatics analysis shows that it is most likely involved in mutarotation. Although the genetic and biochemical analysis can be interpreted as ambiguous, the crystal structure strongly favors placement of RhaU prior to RhaK (P. Loewen personal communication). The details of RhaU's placement in the *R. leguminosarum* L-rhamnose catabolic pathway as well as the possibility that RhaU might be physically associated with RhaK and/or the ABC transporter has not been determined and this requires further investigation.

Chapter 5

Thesis General Conclusions

5.1 Thesis Conclusions and Observations

5.1.1 Project history

When I began working on this project relatively little was known about L-rhamnose catabolism in *R. leguminosarum*. Previous work by Baldani (Baldani *et al.* 1992), and Oresnik (Oresnik *et al.* 1998) had generated several interesting questions regarding the fitness of a strain of *R. leguminosarum* in the rhizosphere and its ability to form a symbiotic association with a plant host. Focus was placed on characterization of genes located on the large plasmids of *R. leguminosarum* that account for a sizeable percentage of the genome in *R. leguminosarum*. Initial characterization of these plasmids revealed a locus responsible for the catabolism of L-rhamnose located on plasmid pRtrW14-2c (Oresnik *et al.* 1998). It was determined that strains of Rlt100 containing mutations within the genes encoding L-rhamnose catabolism were less effective versus wild-type strains for nodule occupancy (Oresnik *et al.* 1998). This observation prompted further characterization leading to the identification of genes in the L-rhamnose catabolic locus that are induced by L-rhamnose and plant root exudate (Oresnik *et al.* 1998). My goal was to characterize the rhamnose locus in order to increase our understanding of the mechanisms by which rhamnose catabolism contributes to effective nodulation and the overall fitness of strain Rlt100. Herein I discuss the significant findings of my project.

5.1.2 Characterization of a locus of genes responsible for L-rhamnose catabolism

The rhamnose locus was sequenced on both strands; it spans 10,959 bp and contains a total of nine putative genes (Figure 2.1). The gene *rhaD* codes for an oxidoreductase, while *rhaI* encodes a sugar isomerase, together they constitute one

transcript. A second transcript is divergently transcribed; it consists of *rhaRSTPQUK* encoding respectively a DeoR-type negative regulator, periplasmic sugar binding protein, ABC-type ATPase protein, two transmembrane proteins, a mutarotase, and a sugar kinase from the sugar kinase-hsp70-actin ATPase protein superfamily (Figure 2.1, Table 2.2).

5.1.3 L-Rhamnose ABC transporter

The uptake of L-rhamnose in *R. leguminosarum* is mediated by the genes *rhaSTPQ* encoding proteins that collectively constitute a typical core ABC importer of gram-negative bacteria. Transport assays using tritiated rhamnose suggested that uptake of rhamnose in strain Rlt100 was induced when grown in media containing L-rhamnose and repressed when grown in media containing glucose (Figure 2.3A); was dependent on the ABC transporter components (Figure 2.3B); and the L-rhamnose transporter is specific for the uptake of L-rhamnose and is not inhibited by other similar sugars L-fucose or L-arabinose.

5.1.4 Differential expression of the rhaRSTPQUK Transcript

The expression of *rhaRSTPQUK* using representative Tn*5*B20 fusions and β-galactosidase fusion analysis suggests all the strains carrying fusions in *rhaRSTPQUK* have higher activity when grown in the presence of L-rhamnose than in the presence of glucose (Table 2.3). The transcription level of the *rha::lacZ* fusions in the *rhaRSTPQUK* operon suggests that differential expression of the transcript occurs: 1) The expression levels exhibited by *rhaR25*::Tn*5*B20 (Table 2.3, Table 2.4); 2) The expression levels exhibited by *rhaP36*::Tn*5*B20 and *rhaQ38*::Tn*5*B20; 3) The expression levels exhibited

by fusions in *rhaK* (Table 2.3). Possible explanations for the observed differential expression of this transcript are discussed below.

5.1.5 Negative regulator RhaR

RhaR acts as a negative regulator since operon basal activity increased in strains lacking *rhaR* (Table 2.4). RhaR over-expressed such that it would be transcribed constitutively in a *rhaR25* background reduced the overall transcriptional activity, while still being inducible by L-rhamnose. Strain Rlt117 (*rhaR25::lacZ* fusion) had very high activity under non-inducing conditions in addition to an induction in the presence of L-rhamnose (Table 2.3). These increased induction levels are not unexpected as RhaR must be synthesized to achieve its function of preventing the catabolic genes in the L-rhamnose operon from being transcribed and subsequently translated thereby conserving cellular energy in the presence of a preferred carbon source. In the absence of a functional copy of RhaR, fusion activity of strain Rlt153 (measured by *rhaR25*) was elevated when the strain was grown on defined media with glucose while fusion activity induced to even higher levels on L-rhamnose (Table 2.4). This suggests that the regulation of this transcript had two levels of control, the negative regulator RhaR as well as induction of this operon independent of the negative regulation provided by RhaR (Table 2.4). Presently the inducing factor is not known. Interestingly, both strain Rlt117 containing a regulator fusion that is polar on downstream elements and strain Rlt153 lacking the entire rhamnose locus were induced with L-rhamnose. It is not clear how L-rhamnose enter the cell in the absence of a transporter complex; this does however raise other important questions. What is the inducer of transcription? Does L-rhamnose bind

RhaR to enable transcriptional activation or is it an L-rhamnose intermediate or a protein or protein complex?

5.1.6 Regulatory role of hairpin loop

The putative hairpin loop between *rhaS* and *rhaT* might regulate the operon by attenuating further transcription into the operon. This attenuation prevents high levels of transcription beyond *rhaS* when the bacteria are grown under non-inducing conditions, allowing differential transcription of the promoter proximal genes relative to distal genes in this operon (Figure 2.1).

The RNA dot blot data show that the gene for the sugar binding component, *rhaS*, is also expressed when Rlt100 is grown in the presence of glucose (Figure 2.2). This differential induction might account for the excess periplasmic binding protein mRNA relative to the rest of the transporter, contributing to high affinity L-rhamnose transport. The elevated expression of many periplasmic sugar binding proteins has been reported. This ensures that if the carbon source is present it can be efficiently captured and transported. In addition, some periplasmic sugar binding proteins are involved in chemotaxis response towards the respective sugar source. While preliminary experiments using strains containing *rhaS* mutations on motility agar did not reveal chemotaxis deficient phenotypes, it is still possible and I predict that RhaS has a role in chemotaxis. The use of more sensitive assays might be useful to determine if this is the case.

It appears that strain Rlt100 demonstrates several regulatory mechanisms of the L-rhamnose catabolic operon: general glucose repression; transcriptional regulation via RhaR; and differential transcriptional regulation via a terminator-like hairpin loop. It

would be interesting to remove the hairpin loop using crossover PCR, I predict this would enable the transcription expression to increase in genes distal to the loop and results similar to the L-rhamnose induced RNA dot blots (Figure 2.2) would be observed. There is also a possibility that the expression might be greatly elevated in the absence of the hairpin loop; if more ABC transport complexes are present elevated rates of L-rhamnose transport are also likely.

5.1.7 Additional regulatory analysis is required

While we have described some regulatory mechanisms of *rhaRSTPQUK*, *rhaDI* regulation has not been as straight forward (Table A.1.14, Table A.1.15). We have assigned both genes to a single transcriptional unit based on complementation data (Table 2.2), and physical searches of the nucleotide sequence revealed no additional transcriptional start sites between the genes. There is a 193 bp space between *rhaD* and *rhaI*, there does not appear to be any obvious function to the space, its role remains unclear. It is possible, using cross-over PCR to eliminate this 193 bp gap in an attempt to derive a functional role. In addition, real-time PCR of this transcript might provide useful information about how the two transcripts are independently regulated and expressed.

Perhaps the lactose negative regulation model might provide insight as to how the L-rhamnose negative regulation works. Lactose is a sugar that is also regulated by a negative regulator LacI. Under non-inducing conditions this protein binds the operator LacO, physically prevents transcription of the catabolic genes LacZ, LacY, and LacA from the upstream promoter LacP. LacI repressor protein manipulates the DNA of the lactose operon via four LacO monomers interacting together binding another operator site

82 base pairs before the transcriptional start site, this interaction physically bends the DNA at the -35 and -10 thereby preventing RNA polymerase from initiating transcription. Allolactose binds LacI under inducing conditions, preventing LacI from binding LacO, enabling transcription of the catabolic genes.

Crystallization analysis of RhaR regulator might reveal clues about potential molecule(s) that might bind RhaR, as allolactose binds LacI, thereby enabling transcription. Over-expression and purification of RhaR can be used in promoter studies to determine the location of the promoter(s), as it is unclear how *rhaRSTPQUK* and *rhaDI* are co-regulated by a common regulator. This might also reveal if RhaR can bind multiple locations on the DNA, as is observed in Lactose regulation. The possibility exists that the 193 base pair gap between *rhaD* and *rhaI* might have a role in regulation by folding the DNA as is observed in lactose regulation within the promoter region.

Another question also remains unanswered; why is rhamnose important for strains of *R. leguminosarum* to competitively nodulate a host plant? The rhamnose locus is up-regulated in the presence of root exudate and strains containing mutations in the genes encoding L-rhamnose catabolism are not as efficient at gaining nodule occupancy versus wild-type strains, however it is unclear why. The soil environment is generally nutrient limiting, so efficient up-regulation of genes involved in capturing available nutrients would be beneficial. Glucose catabolite repression ensures that if glucose is present operons of other catabolic carbon pathways like L-rhamnose are kept in check. Legume root exudate contains many sugars, among these are glucose and L-rhamnose (Bringhurst *et al.* 2001, Knee *et al.* 2001, McNeil *et al.* 1984). We have shown that transcriptional induction levels of *rhaRSTPQUK* are not expressed at high levels (Table 2.3, Table 2.4,

Figure 2.2) as well as reduced transport rates (Figure 2.3) in the presence of glucose, while they are induced with L-rhamnose. It is interesting that the L-rhamnose is up-regulated in the presence of root exudate (Oresnik *et al.* 1998) even though root exudate is known to contain glucose (Bringhurst *et al.* 2001). It is possible that a regulatory pathway exists that is similar to glucose catabolite repression, but functions to induce genes within the rhizosphere and/or upon Rhizobia contact with the root hair and/or possibly within the infection thread that might provide induction to override the existing regulation occurring within the operons. This might begin to explain the observation that induction of *rhaRSTPQUK* was observed independently of the negative regulation provided by RhaR (Table 2.4); however these mechanisms remain unknown.

5.1.8 Novel pathway

We reported that in *R. leguminosarum*, the gene order as well as the predicted functions of some of the genes differs from those reported in the *E. coli* rhamnose operon. In *E. coli* a biochemical lesion in a catabolic pathway following the synthesis of a phosphorylated sugar intermediate can give rise to a conditional lethal phenotype if the bacterium is grown on a medium that contains a non-inducing carbon source that can be catabolized (Adhya and Shapiro 1969). In *R. leguminosarum* strains containing *rhaD* and/or *rhaI* mutations were unable to grow on glycerol/rhamnose (Table 2.5). These results suggest that unlike *E. coli*, where the first step of the catabolic pathway is an isomerization, one of the first steps in the catabolism of L-rhamnose in *R. leguminosarum* strain Rlt100 is a phosphorylation. This may have implications extending beyond *R. leguminosarum*. If this pathway is different in *R. leguminosarum,* it suggests that the

pathway might also differ for other organisms related to *R. leguminosarum.* There is evidence for this being the case, inspection of similar probable L-rhamnose catabolic loci in *R. etli* and *S. meliloti* reveal similar gene organization. It is not unreasonable to suggest that these organisms have similar pathway order to *R. leguminosarum.*

5.1.9 R. leguminosarum operon contains a unique protein

R. leguminosarum operon contains a putative oxidoreductase gene that is not present in the *E. coli* L-rhamnose catabolic pathway. Analysis of the oxidoreductase translated peptide using Conserved Domain Architecture Retrieval Tool (CDART) (Geer *et al.* 2002) predicted that this protein contains a putative NAD^+ binding site, a short chain dehydrogenase domain as well as a type II aldolase domain. Currently the complete L-rhamnose pathway order in *R. leguminosarum* remains unclear, as does the oxidoreductase mechanism of catalysis. Crystallization analysis of this unique protein might enable us to assign function to it. It would also be interesting to see if the function of this protein could be uncoupled by truncating RhaD. We attempted an assay to observe an NAD^+ to NADH oxidation reaction, however we had limited success. Using a more sensitive assay it may be possible to observe truncated protein activity. Characterization of this unique protein might also provide information about the L-rhamnose intermediates generated within this pathway.

5.1.10 RhaK is necessary for L-rhamnose transport

The most significant contribution of my research was the identification of an ABC transporter that requires a kinase for it to function. We determined that strains of *R.*

leguminosarum containing mutations within *rhaK* were unable to transport L-rhamnose in tritiated L-rhamnose transport assays (Figure 3.1). This was unlike any other reported observation for an ABC transporter. While many ABC transporters for import have been described, and several among those were sugar importers, to our knowledge no reports of kinase dependent transport were reported. In order to confirm our initial results of RhaK-dependent transport several experiments were completed.

The introduction of *rhaK*$^{+}$ on a plasmid, into strains carrying a *rhaK* mutation resulted in the strains being complemented both for their ability to grow on L-rhamnose as a sole carbon source as well as their ability to transport L-rhamnose at levels comparable to that of the wild-type Rlt100 (Table 3.2). This suggested the *rhaK* mutations were not affecting the expression of the other ABC transport components.

To eliminate the possibility that RhaK is solely responsible for L-rhamnose transport, strains containing RhaK in the absence of a complete ABC transporter were tested and the results showed no grow on minimal media supplemented with rhamnose and lacked the ability to transport tritiated rhamnose from the growth medium (Table 3.2). Together these results suggest that transport of rhamnose in Rlt100 is dependent on both the components of the ABC transporter and RhaK.

To confirm that strains carrying mutations in *rhaK* had no affect the transcription of the ABC transporter genes, constructs were made that contained the *rhaQ38 lacZ* fusion allele either integrated into the genome, or present on a plasmid, in the presence, or absence, of *rhaK*, expressed from either a plasmid or the genome (Table 3.3). The results clearly show that the fusion activity from *rhaQ38* is not abolished in the absence

of *rhaK*. We concluded that transcription of *rhaSTPQ* occurs in strains containing *rhaK* mutant alleles that are transport deficient.

We developed an assay to detect RhaK dependent kinase activity using L-rhamnose induced cell-free extract, tritiated rhamnose, and ATP-Mg^{2+} using descending paper chromatography. The signal of the intermediate(s) detected were well over the background, and accumulation of reaction products was time dependent and the Rf value the intermediates ran at on the chromatography paper were consistent with values reported for other phosphorylated sugars (Zadrazil 1973). We concluded that RhaK affects both the interconversion as well as a whole cell transport of L-rhamnose (Table 3.2, Table 3.4).

The primary role of sugar kinases reported in the literature was phosphorylation. As RhaK was observed to be involved in the transport process as well as phosphorylation, we wished to determine if transport and phosphorylating activity were dependent on or independent of each other. To resolve this we constructed site-directed mutants of RhaK that have compromised biochemical activity, unable to phosphorylate and determine whether these mutants could transport rhamnose. The RhaK binding pocket was predicted based on the characterized glycerol kinase binding pocket necessary for γ-phosphate contact of ATP (Kabsch and Holmes 1995, Mao *et al.* 1999). In RhaK the conserved motif consists of Gly15, Lys16, Thr17, and Asn18, where the hydroxyl from the amide side chains of these residues are predicted to bond the oxygen atoms from the ATP γ-phosphates as was shown for the corresponding residues in GlpK (Mao *et al.* 1999). We hypothesized that a substitution of Lys16 in RhaK within the β-hairpin contacting the γ-phosphate of ATP should inhibit γ-phosphate of ATP from binding

within the pocket thereby preventing ATP dependent kinase activity. The resultant active site mutant strains generated Rlt206 and Rlt207 while shown to be translated into full length proteins were unable to grow on minimal media supplemented with L-rhamnose as a sole carbon source (Table 3.5). These strains were also unable to import labeled L-rhamnose in radioactive transport assays and both had negligible transport rates (Table 3.5). Additionally, neither mutant had any observable kinase activity in the radioactive chromatography assay (Table 3.5).

The most straightforward explanation of the data is that RhaK plays a role in both the transport as well as the primary interconversion of the substrate. This role, however, may be either direct or indirect. We speculated that either this new type of kinase-dependent ABC transporter is only found in *R. leguminosarum* or alternatively will likely be found in other similar transporters. We used phylogenetic analysis to predict similar kinase proteins (Figure 3.3, Figure 3.4). If a kinase-dependent ABC transporter is to be found within other organisms the kinase proteins we identified are the most likely. A future project might involve mutagenesis of these related kinases to determine if they are necessary for transport. The amino acid alignments we constructed identified several conserved residues. Targeted mutation of these residues might reveal interesting functions of RhaK, this in collaboration with crystallization analysis that is already underway in P. Loewen's lab. Past attempts by us to generate RhaK protein:protein interactions have failed; however it is still likely that RhaK interacts with another protein at some point in the pathway. It would not be unreasonable to predict that RhaK might interact with the transporter components in a similar manner to ABC proteins interact with the transmembrane proteins.

It is interesting to note that *rhaQUK* are transcribed together, and the stop codon of *rhaQ* and the start site of *rhaU* as well as the stop codon for *rhaU* and the start site for *rhaK* overlap (Figure 2.1, Figure 4.1). It was proposed that this might be translated as a multimeric fusion protein that would incorporate itself as a functional unit together with the other ABC transporter components (D. Court personal communication). In this scenario RhaQ, RhaU and RhaK would be in close proximity to each other and also to the other transport proteins. It was also proposed that the lack of complementation by strain Rlt243/ pMR183 containing a wild-type copy of *rhaU* on a plasmid (pMR183) in a Δ*rha*U background (Figure 4.2), might be due to the absence of the multimeric fusion protein. However, the ability to complement a *rhaK* mutant strain with a plasmid copy of *rhaK*, suggests that at least for RhaK, the multimeric fusion protein is not necessary. Overall, we do not have any data that suggests the formation of such a protein fusion; however the theory is intriguing. N-terminally histidine tagged fusion RhaQ over-expressed, purified, run on a gel followed by western blot analysis might resolve this, as a much larger protein would be observed if it existed as a fusion protein.

5.1.11 Determination of RhaU as a mutarotase

BLASTX analysis of RhaU showed that this protein was similar to hypothetically conserved proteins that are found in a wide variety of bacteria none of which had an assigned function. However, YiiL, originally annotated as a hypothetical protein from the *E. coli* genome, was recently been shown to be a rhamnose mutarotase facilitating the interconversion of α- and β-anomers (Ryu *et al.* 2004, Ryu *et al.* 2005). We compared the amino acid sequence of RhaU to both the specific amino acid residues involved in

rhamnose mutarotation in YiiL (Ryu *et al.* 2004, Ryu *et al.* 2005) and the amino acids of several putative proteins suggested to be closely related to RhaU (based on phylogenetic analysis) with a PRALINE alignment. The results suggested that almost all of the residues suggested to be involved in the α- to β-anomeric conversion of L-rhamnose in YiiL were conserved in RhaU and many of the RhaU related proteins, including catalytic residues, residues in catalytic core contacting rhamnose, residues within core, and contacting the methyl group of L-rhamnose (Figure 4.4, Figure 4.5) (Ryu *et al.* 2005). This suggested that RhaU likely functions in a similar manner as YiiL mutarotase. Based on crystallization of RhaU we confirm RhaU catalyses the α- to β-anomerization of L-rhamnose (P. Loewen personal communication).

Briefly as communicated by P. Loewen, summarized from Richardson *et al.* (Richardson *et al.* 2007) the findings of RhaU crystal analysis. Crystals of RhaU were soaked with L-rhamnose. As RhaU initially binds the α-anomer of L-rhamnose contact points are made at OH-2, OH-3, OH-4, and O5 with RhaU. Here α-OH-1 is placed in an unfavorable position with OH-1 just 3.0 Å from the δ-CH_3 of Ile45. This interaction forces α-OH-1 to move towards Tyr20. Hydrogen bonding followed by protonation of O5 of L-rhamnose by His24 leads to sugar ring opening. Protonation of the His24 imidazole is favored by the second imidazole hydrogen bond with the main chain C=O of Tyr20. This reaction is stabilized by a hydrogen bond between O5 and the indole of Trp78. In the electron density map the bound β-anomer rhamnose OH-1 is about 3 Å from Tyr20 with good geometry for hydrogen bond formation. This proposed catalytic mechanism of RhaU is slightly different than the proposed mechanism of YiiL (Richardson *et al.* 2007, Ryu *et al.* 2005). It should be noted that Ryu *et al.* were only able to obtain crystals of

YiiL when co-crystallized with L-rhamnose, whereas RhaU was crystallized and then soaked with L-rhamnose enabling us to observe β-L-rhamnose within the active site, strengthening our proposed catalytic mechanism (Richardson *et al.* 2007).

As with RhaK, phylogenetic analysis of RhaU identified several putative proteins that might act as an L-rhamnose mutarotase (Figure 4.4, Figure 4.5). Analysis of the organization of the genes in putative L-rhamnose coding regions of several related species like *S. meliloti*, are very similar to the organization of the L-rhamnose operon in *R. leguminosarum.* Additionally, a very similar putative protein exists in the catabolic pathway for L-fucose, another methyl pentose sugar in *S. meliloti*; however the genetic organization of genes and the genes involved in this pathway are much different than the L-rhamnose catabolic pathway. Interestingly, both the L-rhamnose and the L-fucose pathways in *E. coli* are similar to each other, and as we now know based on my research that the L-rhamnose pathway in *R. leguminosarum* differs, while L-fucose remains uncharacterized.

5.1.12 R. leguminosarum and E. coli mutarotase mutants exhibit different phenotypes

The phenotype exhibited by Δ*rhaU* strain is slow growth rate on media containing L-rhamnose compared to the wild-type strain. Consistent with this phenotype, RhaU is not essential for growth but the presence of mutarotase speeds up L-rhamnose catabolism (Table 4.2, Figure 4.2). The uncatalyzed L-rhamnose α- to β-anomeric conversion is sufficient to support a low rate of growth in *R. leguminosarum* and the role of the mutarotase is to facilitate the anomerization thereby allowing a faster rate of growth. Phenotypically a strain containing a mutarotase mutation in *E. coli* differs from the

Δ*rhaU* strain in *R. leguminosarum*. It was reported that a strain containing mutation in *yiiL* exhibits a slow growth rate compared to wild-type *E. coli* strain only at low concentrations of L-rhamnose (Ryu *et al.* 2005). In *R. leguminosarum* Δ*rhaU* strain, the slow growth rate is observed when L-rhamnose concentrations are consistent with typical growth conditions. The *rhaU* phenotype was masked by elevated levels of RhaK (Table 4.2), suggesting that the discrepancy with *E. coli yiiL* mutant may lie in lower RhaK levels (Table 2.3, Figure 2.2) in *R. leguminosarum* compared to *E. coli* (Richardson *et al.* 2007). The equilibrium ratio of α and β anomers of L-rhamnose in solution is 56:44 (Ryu *et al.* 2005), excess kinase resulted in greater conversion rates of available β anomers of L-rhamnose. This shift results in a displacement of the α ≈ β anomerization equilibrium by phosphorylation of the β-anomer or a subsequent metabolite.

5.1.13 Placement of RhaU in the Pathway

A biochemical lesion in a catabolic pathway following the synthesis of a phosphorylated sugar intermediate can give rise to a conditional phenotype if the bacterium is grown on a medium that contains a non-inducing carbon source that can be catabolized (Adhya and Shapiro 1969). A strain containing a Δ*rhaU,* grew normally on glycerol-rhamnose, implying that RhaU enzymatic action likely occur before L-rhamnose is phosphorylated (Table 4.2). This observation was supported by data suggesting RhaU binds unphosphorylated L-rhamnose (P. Loewen personal communication). Interestingly, Δ*rhaU* strain had an initial transport rate similar to wild-type rate suggesting that uptake of rhamnose still occurs in the absence of RhaU (Table 4.3). Δ*rhaU* strain was also assayed for RhaK dependent activity; in the absence of RhaU the *in vitro* data show that

Δ*rhaU* strain has the ability to phosphorylate rhamnose *in vivo* (Table 4.4). Placement of RhaU before phosphorylation implies that RhaK has two distinct roles, transport and phosphorylation separated by the RhaU mutarotation step.

5.1.14 Additional outstanding questions and future work

I believe that my research has generated two significant observations: the pathway for L-rhamnose catabolism in *R. leguminosarum* is different from the pathway for the model organism *E. coli*; transport of L-rhamnose in *R. leguminosarum* is dependent upon RhaK. However, there are several outstanding questions raised in the process of characterizing this catabolic locus.

We have shown that the catabolic pathway is novel however we do not know what happens to the L-rhamnose after it is phosphorylated. Does the oxidoreductase or the isomerase act on the rhamnose intermediate after the phosphorylation of the kinase? While we can conclude that the L-rhamnose is phosphorylated by RhaK, we do not know what carbon is the phosphate moiety associated with? It is not understood what catabolic intermediate(s) are formed along the pathway or what intermediate(s) eventually enter other general catabolic pathways?

While we concluded that RhaK has two distinct roles, it is not clear how the entire transport complex functions. RhaK must interact in some way with the ABC transporter to activate transport of rhamnose; however we do not understand how this interaction occurs. We do not know if RhaU is physically associated with RhaK and the ABC transporter. Site directed mutagenesis to conserved amino acid residues on RhaK might help to resolve this (Figure 3.4).

Overall my work summarized in my thesis has led to characterization of a previously uncharacterized locus. In the process of this characterization our work has been cited in a 2006 PNAS paper, entitled 'Mapping the *Sinorhizobium meliloti* 1021 solute-binding protein-dependent transportome' (Mauchline *et al.* 2006). This paper set out to provides a global expression map of one of the largest transporter families (transportome) and an invaluable tool to both understand their solute specificity and the relationships between members of large paralogous families. Within this publication 200 putative ABC transporter in the strain 1021 were identified. They cloned the putative promoter regions upstream of reporter genes using a multicopy plasmid and an integrating vector, subsequent fusions were tested with up to 174 inducing conditions (Mauchline *et al.* 2006). In this work they referred to our work as a model system, and they concluded their findings, in both the plasmid and integrated systems, to be valid based on comparisons to our published characterization of the L-rhamnose catabolic loci. This is evidence that our work is helping to advance research in studies within and beyond *R. leguminosarum.*

References

Abendroth, J., M. Bagdasarian, M. Sandkvist, and W. G. Hol. 2004. The structure of the cytoplasmic domain of EpsL, an inner membrane component of the type II secretion system of *Vibrio cholerae*: an unusual member of the actin-like ATPase superfamily. J Mol Biol **344:**619-33.

Adhya, S. L., and J. A. Shapiro. 1969. The galactose operon of *E. coli* K-12. I. Structural and pleiotropic mutations of the operon. Genetics **62:**231-47.

Alexeyev, M. F. 1999. The pKNOCK series of broad-host-range mobilizable suicide vectors for gene knockout and targeted DNA insertion into the chromosome of gram-negative bacteria. Biotechniques **26:**824-6, 828.

Allikmets, R., N. Singh, H. Sun, N. F. Shroyer, A. Hutchinson, A. Chidambaram, B. Gerrard, L. Baird, D. Stauffer, A. Peiffer, A. Rattner, P. Smallwood, Y. Li, K. L. Anderson, R. A. Lewis, J. Nathans, M. Leppert, M. Dean, and J. R. Lupski. 1997. A photoreceptor cell-specific ATP-binding transporter gene (ABCR) is mutated in recessive Stargardt macular dystrophy. Nat Genet **15:**236-246.

Altenberg, G. A. 2003. The engine of ABC proteins. News Physiol Sci. **18:**191-195.

Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res **25:**3389-402.

Ames, G. F. 1986. Bacterial periplasmic transport systems: structure, mechanism, and evolution. Annu Rev Biochem **55:**397-425.

Ames, G. F., C. E. Liu, A. K. Joshi, and K. Nikaido. 1996. Liganded and unliganded receptors interact with equal affinity with the membrane complex of periplasmic permeases, a subfamily of traffic ATPases. J Biol Chem **271:**14264-70.

Ames, G. F., C. S. Mimura, S. R. Holbrook, and V. Shyamala. 1992. Traffic ATPases: a superfamily of transport proteins operating from *Escherichia coli* to humans. Adv Enzymol Relat Areas Mol Biol **65:**1-47.

Ames, G. F., C. S. Mimura, and V. Shyamala. 1990. Bacterial periplasmic permeases belong to a family of transport proteins operating from *Escherichia coli* to human: Traffic ATPases. FEMS Microbiol Rev **6:**429-46.

Anderson, C. M., R. E. Stenkamp, R. C. McDonald, and T. A. Steitz. 1978. A refined model of the sugar binding site of yeast hexokinase B. J Mol Biol **123:**207-19.

Aravind, L., D. R. Walker, and E. V. Koonin. 1999. Conserved domains in DNA repair proteins and evolution of repair systems. Nucleic Acids Res **27:**1223-42.

Armstrong, S., Tabernero, L., Zhang, H., Hermodsen, M., and Stauffacher, C. 1998. Powering the ABC transporters: the 2.5 angstrom crystal structure of the ABC domain of RbsA. Pediatr. Pulmonl. **17:**91-92.

Ausubel, F. M. R., R. E. Brent, D. D. Kingston, J. G. Moore, J. G. Seidman, J. A. Smith, and K. Struhl. (ed.). 1994. Current protocols in molecular biology. J. Wiley and Sons, New York, N.Y.

Bailey, M. J., C. Hughes, and V. Koronakis. 1997. RfaH and the ops element, components of a novel system controlling bacterial transcription elongation. Mol Microbiol **26:**845-51.

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:**2308-2314.

Ballard, F. J. 1966. Kinetic studies with liver galactokinase. Biochem J **101:**70-5.

Banfalvi, Z., V. Sakanyan, C. Koncz, A. Kiss, I. Dusha, and A. Kondorosi. 1981. Location of nodulation and nitrogen fixation genes on a high molecular weight plasmid of *R. meliloti*. Mol Gen Genet **184:**318-25.

Barnett, M. J., R. F. Fisher, T. Jones, C. Komp, A. P. Abola, F. Barloy-Hubler, L. Bowser, D. Capela, F. Galibert, J. Gouzy, M. Gurjal, A. Hong, L. Huizar, R. W. Hyman, D. Kahn, M. L. Kahn, S. Kalman, D. H. Keating, C. Palm, M. C. Peck, R. Surzycki, D. H. Wells, K. C. Yeh, R. W. Davis, N. A. Federspiel, and L. S. R. 2001. Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti* pSymA megaplasmid. Proc. Natl. Acad. Sci. USA **98:**9883-9888.

Baron, C., and P. C. Zambryski. 1995. The plant response in pathogenesis, symbiosis, and wounding: variations on a common theme? Annu Rev Genet **29:**107-29.

Bates, S. E. 2003. Solving the problem of multidrug resistance: ABC transporters in clinical oncology, p. 359-392. *In* J. L. Linton, Cole, S. P. C., Kuchler, K., and Higgins, C. F. (ed.), ABC proteins from bacteria to man. Academic Press, London.

Bauer, B. E., H. Wolfger, and K. Kuchler. 1999. Inventory and function of yeast ABC proteins: about sex, stress, pleiotropic drug and heavy metal resistance. Biochim Biophys Acta **1461:**217-36.

Beebe, J. A., A. Arabshahi, J. G. Clifton, D. Ringe, G. A. Petsko, and P. A. Frey. 2003. Galactose mutarotase: pH dependence of enzymatic mutarotation. Biochemistry **42:**4414-20.

Beebe, J. A., and P. A. Frey. 1998. Galactose mutarotase: purification, characterization, and investigations of two important histidine residues. Biochemistry **37:**14989-97.

Bennett, W. S., Jr., and T. A. Steitz. 1978. Glucose-induced conformational change in yeast hexokinase. Proc Natl Acad Sci U S A **75:**4848-52.

Bennett, W. S., Jr., and T. A. Steitz. 1980. Structure of a complex between yeast hexokinase A and glucose. I. Structure determination and refinement at 3.5 A resolution. J Mol Biol **140:**183-209.

Benson, D. R., and W. B. Silvester. 1993. Biology of *Frankia* strains, actinomycete symbionts of actinorhizal plants. Microbiol Rev **57:**293-319.

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:**7747-55.

Bergmeyer, H.-U., Holz, G., Kauder, E.M., Möllering, H. and Wieland, O. . 1961. Crystallization of Glycerokinase in *Candida mycoderma*. Biochem. Z **333:**471-480.

Beringer, J. E. 1974. R factor transfer in *Rhizobium leguminosarum*. J Gen Microbiol **84:**188-98.

Beringer, J. E., J. L. Beynon, A. V. Buchanan-Wollason, and A. W. B. Johnston. 1978. Transfer of the drug resistance transposon *Tn*5 to *Rhizobium*. Nature **276:**633–634.

Bettenbrock, K., and C. A. Alpert. 1998. The *gal* genes for the Leloir pathway of *Lactobacillus casei* 64H. Appl Environ Microbiol **64:**2013-9.

Bhuvaneswari, T. V., A. A. Bhagwat, and W. D. Bauer. 1981. Transient susceptibility of root cells in four common legumes to nodulation by rhizobia. Plant Physiol **68:**1144-1149.

Bianchet, M. A., Y. H. Ko, L. M. Amzel, and P. L. Pedersen. 1997. Modeling of nucleotide binding domains of ABC transporter proteins based on a F1-ATPase *rec*A topology: structural model of the nucleotide binding domains of the cystic fibrosis transmembrane conductance regulator (CFTR). J Bioenerg Biomembr **29:**503-24.

Bibikova, T. N., E. B. Blancaflor, and S. Gilroy. 1999. Microtubules regulate tip growth and orientation in root hairs of *Arabidopsis thaliana*. Plant J **17:**657-65.

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:**5983-5993.

Binnie, R. A., H. Zhang, S. Mowbray, and M. A. Hermodson. 1992. Functional mapping of the surface of *Escherichia coli* ribose-binding protein: mutations that affect chemotaxis and transport. Protein Sci **1:**1642-51.

Bishop, L., R. Agbayani, Jr., S. V. Ambudkar, P. C. Maloney, and G. F. Ames. 1989. Reconstitution of a bacterial periplasmic permease in proteoliposomes and demonstration of ATP hydrolysis concomitant with transport. Proc Natl Acad Sci U S A **86:**6953-7.

Bjorkman, A. J., and S. L. Mowbray. 1998. Multiple open forms of ribose-binding protein trace the path of its conformational change. J Mol Biol **279:**651-64.

Blattler, W. A., and J. R. Knowles. 1979. Stereochemical course of phosphokinases. The use of adenosine [gamma-(S)-16O,17O,18O] triphosphate and the mechanistic consequences for the reactions catalyzed by glycerol kinase, hexokinase, pyruvate kinase, and acetate kinase. Biochemistry **18:**3927-33.

Blattner, F. R., G. Plunkett, 3rd, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao. 1997. The complete genome sequence of *Escherichia coli* K-12. Science **277:**1453-74.

Boos, W., and A. Bohm. 2000. Learning new tricks from an old dog: MalT of the *Escherichia coli* maltose system is part of a complex regulatory network. Trends Genet **16:**404-9.

Boos, W., Lucht, J. M. 1996. Periplasmic binding protein-dependent ABC transporters, p. 1175–1209. *In* F. C. Neidhardt (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology. ASM Press, Washington, D.C.

Boos, W., and H. Shuman. 1998. Maltose/maltodextrin system of *Escherichia coli*: transport, metabolism, and regulation. Microbiol Mol Biol Rev **62:**204-29.

Bork, P., C. Sander, and A. Valencia. 1992. An ATPase domain common to prokaryotic cell cycle proteins, sugar kinases, actin, and Hsp70 heat shock proteins. Proc Natl Acad Sci U S A **89:**7290-4.

Bork, P., C. Sander, and A. Valencia. 1993. Convergent evolution of similar enzymatic function on different protein folds: the hexokinase, ribokinase, and galactokinase families of sugar kinases. Protein Sci **2:**31-40.

Borst, P., and R. O. Elferink. 2002. Mammalian ABC transporters in health and disease. Annu Rev Biochem **71:**537-92.

Borths, E. L., K. P. Locher, A. T. Lee, and D. C. Rees. 2002. The structure of *Escherichia coli* BtuF and binding to its cognate ATP binding cassette transporter. Proc Natl Acad Sci U S A **99:**16642-7.

Bouffard, G. G., K. E. Rudd, and S. L. Adhya. 1994. Dependence of lactose metabolism upon mutarotase encoded in the *gal* operon in *Escherichia coli*. J Mol Biol **244:**269-78.

Bowen, G. D., and A. D. Rovira. 1976. Microbial colonization of plant roots. Annu. Rev. Phytopathol. **14:**121–144.

Brdiczka, D. 1990. Interaction of mitochondrial porin with cytosolic proteins. Experientia **46:**161-7.

Bringhurst, R. M., Z. G. Cardon, and D. J. Gage. 2001. Galactosides in the rhizosphere: utilization by Sinorhizobium meliloti and development of a biosensor. Proc Natl Acad Sci U S A **98:**4540-5.

Brink, B. A., J. Miller, R. W. Carlson, and K. D. Noel. 1990. Expression of *Rhizobium leguminosarum* CFN42 genes for lipopolysaccharide in strains derived from different *R. leguminosarum* soil isolates. J Bacteriol **172:**548-55.

Brom, S., A. Garcia de los Santos, T. Stepkowsky, M. Flores, G. Davila, D. Romero, and R. Palacios. 1992. Different plasmids of *Rhizobium leguminosarum* bv. *phaseoli* are required for optimal symbiotic performance. J Bacteriol **174:**5183-9.

Brönsted, J. N., and Guggenheim, E. A. 1927. Contribution to the theory of acid and base catalysis: the mutarotation of glucose. J. Amer. Chem. Soc. **49:**2554.

Bublitz, C., and E. P. Kennedy. 1954. Synthesis of phosphatides in isolated mitochondria. III. The enzymatic phosphorylation of glycerol. J Biol Chem **211:**951-61.

Bukau, B., and A. L. Horwich. 1998. The Hsp70 and Hsp60 chaperone machines. Cell **92:**351-66.

Bunting, K. D. 2002. ABC transporters as phenotypic markers and functional regulators of stem cells. Stem Cells **20:**11-20.

Busch, W., and M. H. Saier, Jr. 2002. The transporter classification (TC) system, 2002. Crit Rev Biochem Mol Biol **37:**287-337.

Caetano-Anolles, G., and P. M. Gresshoff. 1991a. Alfalfa controls nodulation during the onset of *Rhizobium*-induced cortical cell division. Plant Physiol **95:**366-373.

Caetano-Anolles, G., and P. M. Gresshoff. 1991b. Plant genetic control of nodulation. Annu Rev Microbiol **45:**345-82.

Caetano-Anolles, G., P. A. Johshi, and P. M. Gresshoff. 1991. Spontaneous nodules induce feedback suppression of nodulation in alfalfa. Planta **188:**77-92.

Callaham, D. A., and J. G. Torrey. 1981. The structural basis for infection of root hairs of *Trifolium repens* by *Rhizobium*. Can. J. Bot. **59:**1647–1664.

Capela, D., F. Barloy-Hubler, J. Gouzy, G. Bothe, F. Ampe, J. Batut, P. Boistard, A. Becker, M. Boutry, E. Cadieu, S. Dreano, S. Gloux, T. Godrie, A. Goffeau, D.

Kahn, E. Kiss, V. Lelaure, D. Masuy, T. Pohl, D. Portetelle, A. Pühler, B. Purnelle, U. Ramsperger, C. Renard, P. Thebault, M. Vandenbol, S. Weidner, and F. Galibert. 2001. Analysis of the chromosome sequence of the legume symbiont *Sinorhizobium meliloti* strain 1021. Proc Natl Acad Sci U S A **98:**9877-82.

Cardenas, L., L. Vidali, J. Domnguez, H. Prez, F. Snchez, P. K. Hepler, and C. Quinto. 1998. Rearrangement of actin microfilaments in plant root hairs responding to *Rhizobium etli* nodulation signals. Plant Physiol **116:**871-7.

Carol, R. J., and L. Dolan. 2002. Building a hair: tip growth in *Arabidopsis thaliana* root hairs. Philos Trans R Soc Lond B Biol Sci **357:**815-21.

Cassab, G. I. 1998. Plant cell wall proteins. Annu Rev Plant Physiol Plant Mol Biol **49:**281-309.

Chang, G., and C. B. Roth. 2001. Structure of MsbA from *E. coli*: a homolog of the multidrug resistance ATP binding cassette (ABC) transporters. Science **293:**1793-800.

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

Charles, T. C., Singh, R. S., and Finan, T. M. 1990. Lactose utilization and enzymes encoded by megaplasmids in *Rhizobium meliloti* SU47: Implications for population studies. J. Gen. Microbiol. **136:**2497-2502.

Chen, J., G. Lu, J. Lin, A. L. Davidson, and F. A. Quiocho. 2003. A tweezers-like motion of the ATP-binding cassette dimer in an ABC transport cycle. Mol Cell **12:**651-61.

Chen, J., S. Sharma, F. A. Quiocho, and A. L. Davidson. 2001. Trapping the transition state of an ATP-binding cassette transporter: evidence for a concerted mechanism of maltose transport. Proc Natl Acad Sci U S A **98:**1525-30.

Chen, Y. M., Y. Zhu, and E. C. Lin. 1987. The organization of the *fuc* regulon specifying L-fucose dissimilation in *Escherichia coli* K12 as determined by gene cloning. Mol Gen Genet **210:**331-7.

Clark, S. R., I. J. Oresnik, and M. F. Hynes. 2001. RpoN of *Rhizobium leguminosarum* bv. *viciae* strain VF39SM plays a central role in FnrN-dependent microaerobic regulation of genes involved in nitrogen fixation. Mol Gen Genet **264:**623-33.

Cosgrove, D. J., P. Bedinger, and D. M. Durachko. 1997. Group I allergens of grass pollen as cell wall-loosening agents. Proc Natl Acad Sci U S A **94:**6559-64.

Cosgrove, D. J., L. C. Li, H. T. Cho, S. Hoffmann-Benning, R. C. Moore, and D. Blecker. 2002. The growing world of expansins. Plant Cell Physiol **43:**1436-44.

Cotten, J. F., L. S. Ostedgaard, M. R. Carson, and M. J. Welsh. 1996. Effect of cystic fibrosis-associated mutations in the fourth intracellular loop of cystic fibrosis transmembrane conductance regulator. J Biol Chem **271:**21279-84.

Craig, E. A. 1989. Essential roles of 70kDa heat inducible proteins. Bioessays **11:**48-52.

Currier, S. J., S. E. Kane, M. C. Willingham, C. O. Cardarelli, I. Pastan, and M. M. Gottesman. 1992. Identification of residues in the first cytoplasmic loop of P-glycoprotein involved in the function of chimeric human MDR1-MDR2 transporters. J Biol Chem **267:**25153-9.

Dahl, S. G., I. Sylte, and A. W. Ravna. 2004. Structures and models of transporter proteins. J Pharmacol Exp Ther **309:**853-60.

Dassa, E. 2003. Phylogenetic and functional classification of ABC (ATP-binding cassette) systems., p. 3-35. *In* J. L. Linton, Cole, S. P. C., Kuchler, K., and Higgins, C. F. (ed.), ABC Proteins From Bacteria to Man. Academic Press, London.

Dassa, E., and P. Bouige. 2001. The ABC of ABCS: a phylogenetic and functional classification of ABC systems in living organisms. Res Microbiol **152:**211-29.

Dassa, E., and M. Hofnung. 1985a. Homologies between integral proteins of the inner membrane of binding protein transport systems in enterobacteria. Ann Inst Pasteur Microbiol **136A:**281-8.

Dassa, E., and M. Hofnung. 1985b. Sequence of gene *malG* in *E. coli* K12: homologies between integral membrane components from binding protein-dependent transport systems. EMBO J **4:**2287-93.

Davidson, A. L. 2002. Mechanism of coupling of transport to hydrolysis in bacterial ATP-binding cassette transporters. J Bacteriol **184:**1225-33.

Davidson, A. L., H. A. Shuman, and H. Nikaido. 1992. Mechanism of maltose transport in *Escherichia coli*: transmembrane signaling by periplasmic binding proteins. Proc Natl Acad Sci U S A **89:**2360-4.

de Maagd, R., R. de Rijk, I. H. Mulders, and B. J. Lugtenberg. 1989. Immunological characterization of *Rhizobium leguminosarum* outer membrane antigens by use of polyclonal and monoclonal antibodies. J Bacteriol **171:**1136-42.

de Maagd, R. A., I. H. Mulders, H. C. Canter Cremers, and B. J. Lugtenberg. 1992. Cloning, nucleotide sequencing, and expression in *Escherichia coli* of a *Rhizobium leguminosarum* gene encoding a symbiotically repressed outer membrane protein. J Bacteriol **174:**214-21.

de Ruijter, N. C. A., M. B. Rook, T. Bisseling, and A. M. C. Emons. 1998. Lipochito-oligosaccharides re-initiate root hair tip growth in *Vicia sativa* with high calcium and spectrin-like antigen at the tip. Plant J. **13:**341-350.

de Ruijter, N. C. A., T. Bisseling, and A. M. C. Emons. 1999. *Rhizobium* Nod factors induce an increase in sub-apical fine bundles of actin filaments in *Vicia sativa* root hairs within minutes. Mol. Plant-Microbe Interact. **12:**829–832.

Dean, M., and R. Allikmets. 2001. Complete characterization of the human ABC gene family. J Bioenerg Biomembr **33:**475-9.

Denarie, J., F. Debelle, and J. C. Prome. 1996. *Rhizobium* lipo-chitooligosaccharide nodulation factors: signaling molecules mediating recognition and morphogenesis. Annu Rev Biochem **65:**503-35.

di Guan, C., P. Li, P. D. Riggs, and H. Inouye. 1988. Vectors that facilitate the expression and purification of foreign peptides in *Escherichia coli* by fusion to maltose-binding protein. Gene **67:**21-30.

Diaz, C. L., T. Logman, H. C. Stam, and J. W. Kijne. 1995. Sugar-binding activity of pea lectin expressed in white clover hairy roots. Plant Physiol **109:**1167-1177.

Diaz, C. L., P. C. Van Spronsen, R. Bakhuizen, G. J. J. Logman, E. J. J. Lugtenberg, and J. W. Kijne. 1986. Correlation between infection by *Rhizobium leguminosarum* and lectin on the surface of *Pisum sativum* roots. Planta **168:**350-359.

Diederichs, K., J. Diez, G. Greller, C. Muller, J. Breed, C. Schnell, C. Vonrhein, W. Boos, and W. Welte. 2000. Crystal structure of MalK, the ATPase subunit of the trehalose/maltose ABC transporter of the archaeon *Thermococcus litoralis*. EMBO J **19:**5951-61.

Dowling, D. N., and W. J. Broughton. 1986. Competition for nodulation of legumes. Annu Rev Microbiol **40:**131-57.

Downie, J. A. 1994. Signaling strategies for nodulation of legumes by rhizobia. Trends Microbiol **2:**318-24.

Drew, K. N., Zajicek, J., Bondo, G., Bose, B., and Serianni, A. S. 1998. Carbohydr. Res. **307:**199–209.

Dudley, M. E., T. W. Jacobs, and S. R. Long. 1987. Microscopic studies of cell divisions induced in alfalfa roots by *Rhizobium meliloti*. Planta **171:**289-301.

Eckhardt, T. 1978. A rapid method for the identification of plasmid deoxyribonucleic acid in bacteria. Plasmid **1:**584-8.

Elferink, M. G., S. V. Albers, W. N. Konings, and A. J. Driessen. 2001. Sugar transport in *Sulfolobus solfataricus* is mediated by two families of binding protein-dependent ABC transporters. Mol Microbiol **39:**1494-503.

Erlandson, K. A., S. C. Delamarre, and C. A. Batt. 2001. Genetic evidence for a defective xylan degradation pathway in *Lactococcus lactis*. Appl Environ Microbiol **67:**1445-52.

Esseling, J. J., F. G. Lhuissier, and A. M. Emons. 2003. Nod factor-induced root hair curling: continuous polar growth towards the point of Nod factor application. Plant Physiol **132:**1982-8.

Eym, Y., Y. Park, and C. Park. 1996. Genetically probing the regions of ribose-binding protein involved in permease interaction. Mol Microbiol **21:**695-702.

Fahraeus, G. 1957. The infection of clover root hairs by nodule bacteria studied by a simple glass slide technique. J Gen Microbiol **16:**374-81.

Fath, M. J., Kolter, R. 1993. ABC transporters: bacterial exporters. Microbiol Rev **57:**995-1017.

Feese, M., Pettigrew, D. W., Meadow, N. D., Roseman, S., Remington, S. J. 1994. Cation-promoted association of a regulatory and target protein is controlled by protein phosphorylation. Proc Natl Acad Sci U S A **91:**3544-8.

Feese, M. D., Faber, H. R., Bystrom, C. E., Pettigrew, D. W., Remington, S. J. 1998. Glycerol kinase from *Escherichia coli* and an Ala65-->Thr mutant: the crystal structures reveal conformational changes with implications for allosteric regulation. Structure **6:**1407-18.

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39:**783-789.

Felsenstein, J. 2002. PHYLIP (Phylogeny Inference Package) version 3.6a, 3.6a ed. distributed by the author, Dept. of Genetics, University of Washington, Seattle.

Fetsch, E. E., Davidson, A. L. 2002. Vanadate-catalyzed photocleavage of the signature motif of an ATP-binding cassette (ABC) transporter. Proc Natl Acad Sci U S A **99:**9685-90.

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

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

Finan, T. M., S. Weidner, K. Wong, J. Buhrmester, P. Chain, F. J. Vorholter, I. Hernandez-Lucas, A. Becker, A. Cowie, J. Gouzy, B. Golding, and A. Puhler. 2001. The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endosymbiont *Sinorhizobium meliloti*. Proc Natl Acad Sci U S A **98:**9889-94.

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

Flaherty, K. M., C. DeLuca-Flaherty, and D. B. McKay. 1990. Three-dimensional structure of the ATPase fragment of a 70K heat-shock cognate protein. Nature **346:**623-8.

Flaherty, K. M., D. B. McKay, W. Kabsch, and K. C. Holmes. 1991. Similarity of the three-dimensional structures of actin and the ATPase fragment of a 70-kDa heat shock cognate protein. Proc Natl Acad Sci U S A **88:**5041-5.

Flocco, M. M., and S. L. Mowbray. 1994. The 1.9 A x-ray structure of a closed unliganded form of the periplasmic glucose/galactose receptor from *Salmonella typhimurium*. J Biol Chem **269:**8931-6.

Freiberg, C., R. Fellay, A. Bairoch, W. J. Broughton, A. Rosenthal, and X. Perret. 1997. Molecular basis of symbiosis between *Rhizobium* and legumes. Nature **387:**394-401.

Frey, P. A. 1996. The Leloir pathway: a mechanistic imperative for three enzymes to change the stereochemical configuration of a single carbon in galactose. Faseb J **10:**461-70.

Gage, D. J. 2002. Analysis of infection thread development using Gfp- and DsRed-expressing *Sinorhizobium meliloti*. J Bacteriol **184:**7042-6.

Gage, D. J. 2004. Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. Microbiol Mol Biol Rev **68:**280-300.

Gage, D. J., T. Bobo, and S. R. Long. 1996. Use of green fluorescent protein to visualize the early events of symbiosis between *Rhizobium meliloti* and alfalfa (*Medicago sativa*). J Bacteriol **178:**7159-66.

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. Dreano, N. A. Federspiel, R. F. Fisher, S. Gloux, T. Godrie, A. Goffeau, B. Golding, J. Gouzy, M. Gurjal, I. Hernandez-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. Thebault, 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:**668-72.

Gardina, P., C. Conway, M. Kossman, and M. Manson. 1992. Aspartate and maltose-binding protein interact with adjacent sites in the Tar chemotactic signal transducer of *Escherichia coli*. J Bacteriol **174:**1528-36.

Gatz, C., J. Altschmied, and W. Hillen. 1986. Cloning and expression of the *Acinetobacter calcoaceticus* mutarotase gene in *Escherichia coli*. J Bacteriol **168:**31-9.

Gatz, C., and W. Hillen. 1986. *Acinetobacter calcoaceticus* encoded mutarotase: nucleotide sequence analysis of the gene and characterization of its secretion in *Escherichia coli*. Nucleic Acids Res **14:**4309-23.

Gaudet, R., and D. C. Wiley. 2001. Structure of the ABC ATPase domain of human TAP1, the transporter associated with antigen processing. EMBO J **20:**4964-72.

Geer, L. Y., M. Domrachev, D. J. Lipman, and S. H. Bryant. 2002. CDART: protein homology by domain architecture. Genome Res **12:**1619-23.

Georgopoulos, C. 1992. The emergence of the chaperone machines. Trends Biochem Sci **17:**295-9.

Gething, M. J., and J. Sambrook. 1992. Protein folding in the cell. Nature **355:**33-45.

Gibbons, I. R., B. H. Gibbons, G. Mocz, and D. J. Asai. 1991. Multiple nucleotide-binding sites in the sequence of dynein beta heavy chain. Nature **352:**640-3.

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:**3971-4.

Gonzalez, V., R. I. Santamaria, P. Bustos, I. Hernandez-Gonzalez, A. Medrano-Soto, G. Moreno-Hagelsieb, S. C. Janga, M. A. Ramirez, V. Jimenez-Jacinto, J. Collado-Vides, and G. Davila. 2006. The partitioned *Rhizobium etli* genome: genetic and metabolic redundancy in seven interacting replicons. Proc Natl Acad Sci U S A **103:**3834-9.

Gottesman, M. M. 2002. Mechanisms of cancer drug resistance. Annu Rev Med **53:**615-27.

Groeger, W., and W. Koster. 1998. Transmembrane topology of the two FhuB domains representing the hydrophobic components of bacterial ABC transporters involved in the uptake of siderophores, haem and vitamin B12. Microbiology **144:**2759-69.

Hanrahan, J. W., M. Gentzsch, and J. R. Riordan. 2003. The cystic fibrosis transmembrane conductance regulator (ABCC7), p. 589-618. *In* I. B. Holland, Cole, S. P. C., Kuchler, K., and Higgins, C. F. (ed.), ABC Proteins from Bacteria to Man. Academic Press, London.

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

Heidstra, R., G. Nilsen, F. Martinez-Abarca, A. van Kammen, and T. Bisseling. 1997. Nod factor-induced expression of leghemoglobin to study the mechanism of NH_4NO_3 inhibition on root hair deformation. Mol Plant Microbe Interact **10:**215-20.

Hendrick, J. P., and F. U. Hartl. 1993. Molecular chaperone functions of heat-shock proteins. Annu Rev Biochem **62:**349-84.

Hepler, P. K., L. Vidali, and A. Y. Cheung. 2001. Polarized cell growth in higher plants. Annu Rev Cell Dev Biol **17:**159-87.

Heringa, J. 1999. Two strategies for sequence comparison: profile-preprocessed and secondary structure-induced multiple alignment. Comput Chem **23:**341-64.

Heringa, J. 2000. Computational methods for protein secondary structure prediction using multiple sequence alignments. Curr Protein Pept Sci **1:**273-301.

Heringa, J. 2002. Local weighting schemes for protein multiple sequence alignment. Comput Chem **26:**459-77.

Herrmann, C., J. Wray, F. Travers, and T. Barman. 1992. Effect of 2,3-butanedione monoxime on myosin and myofibrillar ATPases. An example of an uncompetitive inhibitor. Biochemistry **31:**12227-32.

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

Higgins, C. F., I. D. Hiles, G. P. Salmond, D. R. Gill, J. A. Downie, I. J. Evans, I. B. Holland, L. Gray, S. D. Buckel, A. W. Bell, and et al. 1986. A family of related ATP-binding subunits coupled to many distinct biological processes in bacteria. Nature **323:**448-50.

Higgins, C. F., I. D. Hiles, K. Whalley, and D. J. Jamieson. 1985. Nucleotide binding by membrane components of bacterial periplasmic binding protein-dependent transport systems. EMBO J **4:**1033-9.

Higgins, C. F., and K. J. Linton. 2003. ABC transporters: an introduction and overview, p. xvii-xxii. *In* J. L. Linton, Cole, S. P. C., Kuchler, K., and Higgins, C. F. (ed.), ABC transporters: from bacteria to man. Academic Press, London.

Higgins, C. F., and K. J. Linton. 2004. The ATP switch model for ABC transporters. Nat Struct Mol Biol **11:**918-26.

Higuchi, R., B. Krummel, and R. K. Saiki. 1988. A general method of *in vitro* preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. Nucleic Acids Res **16:**7351-67.

Hirsch, A. M. 1992. Developmental biology of legume nodulation. New Phytologist **122:**211-237.

Hirsch, A. M. 1999. Role of lectins (and rhizobial exopolysaccharides) in legume nodulation. Curr Opin Plant Biol **2:**320-6.

Hirsch, P. R. 1979. Plasmid-determined bacteriocin production by *Rhizobium leguminosarum*. J. Gen. Microbiol. **113:**219-228.

Hirsch, P. R., Van Montagu, M., Johnston, A. W. B., Brewin, N. J., and Schell, J. 1980. Physical identification of bacteriocinogenic, nodulation, and other plasmids in strains of *Rhizobium leguminosarum.* J.Gen. Microbiol. **120:**403-412.

Ho, S. N., H. D. Hunt, R. M. Horton, J. K. Pullen, and L. R. Pease. 1989. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene **77:**51-9.

Hobson, A. C., R. Weatherwax, and G. F. Ames. 1984. ATP-binding sites in the membrane components of histidine permease, a periplasmic transport system. Proc Natl Acad Sci U S A **81:**7333-7.

Holden, H. M., I. Rayment, and J. B. Thoden. 2003. Structure and function of enzymes of the Leloir pathway for galactose metabolism. J Biol Chem **278:**43885-8.

Holland, I. B. 2003. Introduction to bacterial ABC proteins, p. 149-155. *In* J. L. Linton, Cole, S. P. C., Kuchler, K., and Higgins, C. F. (ed.), ABC proteins: from bacteria to man. Academic Press, London.

Holland, I. B., and M. A. Blight. 1999. ABC ATPases, adaptable energy generators fuelling transmembrane movement of a variety of molecules in organisms from bacteria to humans. J. Mol. Biol. **293:**381-399.

Honeycutt, R. J., M. McClelland, and B. W. Sobral. 1993. Physical map of the genome of *Rhizobium meliloti* 1021. J Bacteriol **175:**6945-52.

Hooykaas, P. J. J., Van Brussel, A. A. N., den Dulk-Ras, H., van Sloteren, G. M., and Schilperooret, R. A. 1981. Sym plasmid of *Rhizobium trifolii* expressed in different rhizobial species and *Agrobacterium tumefaciens.* Nature **291:**351-353.

Hopfner, K. P., A. Karcher, D. S. Shin, L. Craig, L. M. Arthur, J. P. Carney, and J. A. Tainer. 2000. Structural biology of Rad50 ATPase: ATP-driven

conformational control in DNA double-strand break repair and the ABC-ATPase superfamily. Cell **101:**789-800.

Hopfner, K. P., and J. A. Tainer. 2003. Rad50/SMC proteins and ABC transporters: unifying concepts from high-resolution structures. Curr Opin Struct Biol **13:**249-55.

Hor, L. I., and H. A. Shuman. 1993. Genetic analysis of periplasmic binding protein dependent transport in *Escherichia coli*. Each lobe of maltose-binding protein interacts with a different subunit of the MalFGK2 membrane transport complex. J Mol Biol **233:**659-70.

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:**680-9.

Hosie, A. H., and P. S. Poole. 2001. Bacterial ABC transporters of amino acids. Res Microbiol **152:**259-70.

Howard, S. M., and M. R. Heinrich. 1965. The Anomeric Specificity of Yeast Galactokinase. Arch Biochem Biophys **110:**395-400.

Hu, Y., S. Rech, R. P. Gunsalus, and D. C. Rees. 1997. Crystal structure of the molybdate binding protein ModA. Nat Struct Biol **4:**703-7.

Hucho, F., and K. Wallenfels. 1971. The enzymatically catalyzed mutarotation. The mechanism of action of mutarotase (aldose-1-epimerase) from *Escherichia coli*. Eur J Biochem **23:**489-96.

Hung, L. W., I. X. Wang, K. Nikaido, P. Q. Liu, G. F. Ames, and S. H. Kim. 1998. Crystal structure of the ATP-binding subunit of an ABC transporter. Nature **396:**703-7.

Hunke, S., H. Landmesser, and E. Schneider. 2000a. Novel missense mutations that affect the transport function of MalK, the ATP-binding-cassette subunit of the *Salmonella enterica* serovar *typhimurium* maltose transport system. J Bacteriol **182:**1432-6.

Hunke, S., M. Mourez, M. Jehanno, E. Dassa, and E. Schneider. 2000b. ATP modulates subunit-subunit interactions in an ATP-binding cassette transporter (MalFGK2) determined by site-directed chemical cross-linking. J Biol Chem **275:**15526-34.

Hurley, J. H. 1996. The sugar kinase/heat shock protein 70/actin superfamily: implications of conserved structure for mechanism. Annu Rev Biophys Biomol Struct **25:**137-62.

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* IIIGlc with glycerol kinase. Science **259:**673-7.

Hyde, S. C., P. Emsley, M. J. Hartshorn, M. M. Mimmack, U. Gileadi, S. R. Pearce, M. P. Gallagher, D. R. Gill, R. E. Hubbard, and C. F. Higgins. 1990. Structural model of ATP-binding proteins associated with cystic fibrosis, multidrug resistance and bacterial transport. Nature **346:**362-5.

Hynes, M. F. 1990a. The role of plasmids in competition between strains of *Rhizobium leguminosarum*. *In* P. M. Gressshoff, Roth, L. E., Stacey, G., and Newton, W. E. (ed.), Nitrogen fixation: achievements and objectives. Chapman and Hall, New York.

Hynes, M. F., and O'Connell,. 1990b. Host plant effect on competition among strains of *Rhizobium leguminosarum*. Can. J. Microbiol. **36:**864-890.

Hynes, M. F., and N. F. McGregor. 1990. Two plasmids other than the nodulation plasmid are necessary for formation of nitrogen-fixing nodules by *Rhizobium leguminosarum*. Mol Microbiol **4:**567-74.

Hynes, M. F., Simon, R., Müller, P., Niehaus, K., Labes, M., and Pühler, A. 1986. The two megaplasmids of *Rhizobium meliloti* are involved in the effective nodulation of alfalfa. Mol. Gen. Genet. **202:**356-362.

Iwanami, Y. 1956. Protoplasmic movement in pollen grains and pollen tubes. Phytomorphology **6:**288–295.

Janas, E., M. Hofacker, M. Chen, S. Gompf, C. van der Does, and R. Tampe. 2003. The ATP hydrolysis cycle of the nucleotide-binding domain of the mitochondrial ATP-binding cassette transporter Mdl1p. J Biol Chem **278:**26862-9.

Jin, R. Z., R. G. Forage, and E. C. Lin. 1982. Glycerol kinase as a substitute for dihydroxyacetone kinase in a mutant of *Klebsiella pneumoniae*. J Bacteriol **152:**1303-7.

Johnston, A. W. B., Beynon, J. L., Buchanan-Wollaston, A. V., Setchell, S. M., Hirsch, P. R., and Beringer, J. E. 1978. High frequency transfer of nodulating ability between strains and species of *Rhizobium*. Nature **276:**634-636.

Jones, J. D., and N. Gutterson. 1987. An efficient mobilizable cosmid vector, pRK7813, and its use in a rapid method for marker exchange in *Pseudomonas fluorescens* strain HV37a. Gene **61:**299-306.

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

Jones, P. M., and A. M. George. 2002. Mechanism of ABC transporters: a molecular dynamics simulation of a well characterized nucleotide-binding subunit. Proc Natl Acad Sci U S A **99:**12639-44.

Jones, P. M., and A. M. George. 2004. The ABC transporter structure and mechanism: perspectives on recent research. Cell Mol Life Sci **61:**682-99.

Junop, M. S., G. Obmolova, K. Rausch, P. Hsieh, and W. Yang. 2001. Composite active site of an ABC ATPase: MutS uses ATP to verify mismatch recognition and authorize DNA repair. Mol Cell **7:**1-12.

Jurnak, F. 1985. Structure of the GDP domain of EF-Tu and location of the amino acids homologous to Ras oncogene proteins. Science **230:**32-6.

Kabsch, W., and K. C. Holmes. 1995. The actin fold. Faseb J **9:**167-74.

Kabsch, W., H. G. Mannherz, D. Suck, E. F. Pai, and K. C. Holmes. 1990. Atomic structure of the actin:DNase I complex. Nature **347:**37-44.

Kabsch, W., and J. Vandekerckhove. 1992. Structure and function of actin. Annu Rev Biophys Biomol Struct **21:**49-76.

Kalckar, H. M. 1960. Hereditary defects in galactose metabolism in man and microorganisms. Fed Proc **19:**984-90.

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, Y. Mochizuki, S. Nakayama, N. Nakazaki, S. Shimpo, M. Sugimoto, C. Takeuchi, M. Yamada, and S. Tabata. 2000. Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. DNA Res **7:**331-8.

Kaneko, T., Y. Nakamura, S. Sato, K. Minamisawa, T. Uchiumi, S. Sasamoto, A. Watanabe, K. Idesawa, M. Iriguchi, K. Kawashima, M. Kohara, M. Matsumoto, S. Shimpo, H. Tsuruoka, T. Wada, M. Yamada, and S. Tabata. 2002. Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110. DNA Res **9:**189-97.

Karpowich, N., O. Martsinkevich, L. Millen, Y. R. Yuan, P. L. Dai, K. MacVey, P. J. Thomas, and J. F. Hunt. 2001. Crystal structures of the MJ1267 ATP binding cassette reveal an induced-fit effect at the ATPase active site of an ABC transporter. Structure **9:**571-86.

Karpowich, N. K., H. H. Huang, P. C. Smith, and J. F. Hunt. 2003. Crystal structures of the BtuF periplasmic-binding protein for vitamin B12 suggest a functionally important reduction in protein mobility upon ligand binding. J Biol Chem **278:**8429-34.

Kemner, J. M., X. Liang, and E. W. Nester. 1997. The *Agrobacterium tumefaciens* virulence gene *chvE* is part of a putative ABC-type sugar transport operon. J Bacteriol **179:**2452-8.

Kerr, I. D. 2002. Structure and association of ATP-binding cassette transporter nucleotide-binding domains. Biochim Biophys Acta **1561:**47-64.

Ketelaar, T., and A. M. C. Emons. 2001. The cytoskeleton in plant cell growth: lessons from root hairs. New Phytol. **152:**409–418.

Kim, M. S., H. Oh, C. Park, and B. H. Oh. 2001. Crystallization and preliminary X-ray crystallographic analysis of *Escherichia coli* RbsD, a component of the ribose-transport system with unknown biochemical function. Acta Crystallogr D Biol Crystallogr **57:**728-30.

Kim, M. S., J. Shin, W. Lee, H. S. Lee, and B. H. Oh. 2003. Crystal structures of RbsD leading to the identification of cytoplasmic sugar-binding proteins with a novel folding architecture. J Biol Chem **278:**28173-80.

Knee, E. M., F. C. 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:**775-84.

Koebnik, R., K. P. Locher, and P. Van Gelder. 2000. Structure and function of bacterial outer membrane proteins: barrels in a nutshell. Mol Microbiol **37:**239-53.

Konig, J., A. T. Nies, Y. Cui, I. Leier, and D. Keppler. 1999. Conjugate export pumps of the multidrug resistance protein (MRP) family: localization, substrate specificity, and MRP2-mediated drug resistance. Biochim Biophys Acta **1461:**377-94.

Kosslak, R. M., and B. B. Bohlool. 1984. Suppression of nodule development of one side of a split-root system of soybeans caused by prior inoculation of the other side. Plant Physiol **75:**125-130.

Kuzma, M. M., S. Hunt, and D. B. Layzell. 1993. Role of Oxygen in the Limitation and Inhibition of Nitrogenase Activity and Respiration Rate in Individual Soybean Nodules. Plant Physiol **101:**161-169.

la Cour, T. F., J. Nyborg, S. Thirup, and B. F. Clark. 1985. Structural details of the binding of guanosine diphosphate to elongation factor Tu from *E. coli* as studied by X-ray crystallography. EMBO J **4:**2385-8.

Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:**680-5.

Lamers, M. H., A. Perrakis, J. H. Enzlin, H. H. Winterwerp, N. de Wind, and T. K. Sixma. 2000. The crystal structure of DNA mismatch repair protein MutS binding to a G x T mismatch. Nature **407:**711-7.

Lamers, M. H., H. H. Winterwerp, and T. K. Sixma. 2003. The alternating ATPase domains of MutS control DNA mismatch repair. EMBO J **22:**746-56.

Lee, K. H., and T. A. Larue. 1992. Exogenous ethylene inhibits nodulation of *Pisum sativum* L. cv Sparkle. Plant Physiol **100:**1759-1763.

Lesk, A. M., and C. Chothia. 1984. Mechanisms of domain closure in proteins. J Mol Biol **174:**175-91.

Levy, G. B., and E. S. Cook. 1954. A rotographic study of mutarotase. Biochem J **57:**50-5.

Libbenga, K. R., and P. A. A. Harkes. 1973. Initial proliferation of cortical cells in the formation of root nodules in *Pisum sativum*. Planta **114:**17-28.

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

Lin, E. C. 1976. Glycerol dissimilation and its regulation in bacteria. Annu Rev Microbiol **30:**535-78.

Linton, J. L., Rosenberg, M. F., Kerr, I.D., and Higgins, C. F. 2003. Structure of ABC transporters, p. 65-80. *In* J. L. Linton, Cole, S. P. C., Kuchler, K., and Higgins, C. F. (ed.), ABC proteins: from bacteria to man. Academic Press, London.

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

Livingstone, G., Franks, F., and Aspinall, L. J. 1977. The effects of aqueous solvent structure on the mutarotation kinetics of glucose. J. Solution Chem. **6:**203–216.

Lloyd, C., K. Pearce, D. J. Rawlins, R. W. Ridge, and P. J. Shaw. 1987. Endoplasmic microtubules connect the advancing nucleus to the tip of legume root hairs, but F-actin is involved in basipetal migration. Cell Motil. Cytoskel. **8:**27–36.

Locher, K. P. 2004. Structure and mechanism of ABC transporters. Curr Opin Struct Biol **14:**426-31.

Locher, K. P., and E. Borths. 2004. ABC transporter architecture and mechanism: implications from the crystal structures of BtuCD and BtuF. FEBS Lett **564:**264-8.

Locher, K. P., A. T. Lee, and D. C. Rees. 2002. The *E. coli* BtuCD structure: a framework for ABC transporter architecture and mechanism. Science **296:**1091-8.

Long, S. R. 1996. *Rhizobium* symbiosis: nod factors in perspective. Plant Cell **8:**1885-98.

Long, S. R. 2001. Genes and signals in the *Rhizobium*-legume symbiosis. Plant Physiol **125:**69-72.

Loo, T. W., M. C. Bartlett, and D. M. Clarke. 2002. The "LSGGQ" motif in each nucleotide-binding domain of human P-glycoprotein is adjacent to the opposing Walker A sequence. J Biol Chem **277:**41303-6.

Loo, T. W., M. C. Bartlett, and D. M. Clarke. 2003. Drug binding in human P-glycoprotein causes conformational changes in both nucleotide-binding domains. J Biol Chem **278:**1575-8.

Lowe, J., S. C. Cordell, and F. van den Ent. 2001. Crystal structure of the SMC head domain: an ABC ATPase with 900 residues antiparallel coiled-coil inserted. J Mol Biol **306:**25-35.

Luckey, M., and H. Nikaido. 1980a. Diffusion of solutes through channels produced by phage lambda receptor protein of *Escherichia coli*: inhibition by higher oligosaccharides of maltose series. Biochem Biophys Res Commun **93:**166-71.

Luckey, M., and H. Nikaido. 1980b. Specificity of diffusion channels produced by lambda phage receptor protein of *Escherichia coli*. Proc Natl Acad Sci U S A **77:**167-71.

Luecke, H., and F. A. Quiocho. 1990. High specificity of a phosphate transport protein determined by hydrogen bonds. Nature **347:**402-6.

Mao, C., Z. Ozer, M. Zhou, and F. M. Uckun. 1999. X-Ray structure of glycerol kinase complexed with an ATP analog implies a novel mechanism for the ATP-dependent glycerol phosphorylation by glycerol kinase. Biochem Biophys Res Commun **259:**640-4.

Martinez-Romero, E., and M. Rosenblueth. 1990. Increased bean (*Phaseolus vulgaris* L.) nodulation competitiveness of genetically modified *Rhizobium* strains. Appl Environ Microbiol **56:**2384-2388.

Martinoia, E., M. Klein, M. Geisler, L. Bovet, C. Forestier, U. Kolukisaoglu, B. Muller-Rober, and B. Schulz. 2002. Multifunctionality of plant ABC transporters--more than just detoxifiers. Planta **214:**345-55.

Mateos, P. F., D. L. Baker, M. Petersen, E. Velazquez, J. I. Jimenez-Zurdo, E. Martinez-Molina, A. Squartini, G. Orgambide, D. H. Hubbell, and F. B. Dazzo. 2001. Erosion of root epidermal cell walls by *Rhizobium* polysaccharide-

degrading enzymes as related to primary host infection in the *Rhizobium*-legume symbiosis. Can J Microbiol **47:**475-87.

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

Maurel, C., Reizer, J., Schroeder, J. I., Chrispeels, M. J., Saier, M. H., Jr. 1994. Functional characterization of the *Escherichia coli* glycerol facilitator, GlpF, in *Xenopus oocytes*. J Biol Chem **269:**11869-72.

McCabe, E. R. B. 1995. p. 1631–1652. *In* C. R. Scrivner (ed.), The metabolic basis of inherited disease, 7th ed. McGraw–Hill, New York.

McClelland, M., K. E. Sanderson, J. Spieth, S. W. Clifton, P. Latreille, L. Courtney, S. Porwollik, J. Ali, M. Dante, F. Du, S. Hou, D. Layman, S. Leonard, C. Nguyen, K. Scott, A. Holmes, N. Grewal, E. Mulvaney, E. Ryan, H. Sun, L. Florea, W. Miller, T. Stoneking, M. Nhan, R. Waterston, and R. K. Wilson. 2001. Complete genome sequence of *Salmonella enterica* serovar *Typhimurium* LT2. Nature **413:**852-6.

McLaughlin, P. J., J. T. Gooch, H. G. Mannherz, and A. G. Weeds. 1993. Structure of gelsolin segment 1-actin complex and the mechanism of filament severing. Nature **364:**685-92.

McNeil, M., A. G. Darvill, S. C. Fry, and P. Albersheim. 1984. Structure and function of the primary cell walls of plants. Annu Rev Biochem **53:**625-63.

Meade, H. M., S. R. Long, G. B. Ruvkun, S. E. Brown, and F. M. Ausubel. 1982. Physical and genetic characterization of symbiotic and auxotrophic mutants of *Rhizobium meliloti* induced by transposon *Tn*5 mutagenesis. J Bacteriol **149:**114-22.

Miller, D. D., N. C. A. de Ruijter, T. Bisseling, and A. M. C. Emons. 1999. The role of actin in root hair morphogenesis: studies with lipochito-oligosaccharide as a growth stimulator and cytochalasin as an actin perturbing drug. Plant J. **17:**141-154.

Miller, D. M., 3rd, J. S. Olson, J. W. Pflugrath, and F. A. Quiocho. 1983. Rates of ligand binding to periplasmic proteins involved in bacterial transport and chemotaxis. J Biol Chem **258:**13665-72.

Miller, J. H. 1972. p. 1-468. *In* J. H. Miller (ed.), Experiments in molecular genetics. Cold Springs Harbor Laboratory, Cold Springs Harbor, N.Y.

Mimmack, M. L., M. P. Gallagher, S. R. Pearce, S. C. Hyde, I. R. Booth, and C. F. Higgins. 1989. Energy coupling to periplasmic binding protein-dependent

transport systems: stoichiometry of ATP hydrolysis during transport *in vivo*. Proc Natl Acad Sci U S A **86:**8257-61.

Miroux, B., and J. E. Walker. 1996. Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. J Mol Biol **260:**289-98.

Moenne-Locoz, Y., Baldani, I.J., Weaver, R.W. 1995. Sequential heat curing of Tn*5mob-sac* labeled plasmids from *Rhizobium* to obtain derivatives with various combinations of plasmids and no plasmids. Lett Appl Microbiol **20:**175-179.

Moody, J. E., L. Millen, D. Binns, J. F. Hunt, and P. J. Thomas. 2002. Cooperative, ATP-dependent association of the nucleotide binding cassettes during the catalytic cycle of ATP-binding cassette transporters. J Biol Chem **277:**21111-4.

Mortensen, L., G. Dandanell, and K. Hammer. 1989. Purification and characterization of the *deoR* repressor of *Escherichia coli*. EMBO J **8:**325-31.

Moulin, L., A. Munive, B. Dreyfus, and C. Boivin-Masson. 2001. Nodulation of legumes by members of the beta-subclass of Proteobacteria. Nature **411:**948-50.

Mourez, M., M. Hofnung, and E. Dassa. 1997. Subunit interactions in ABC transporters: a conserved sequence in hydrophobic membrane proteins of periplasmic permeases defines an important site of interaction with the ATPase subunits. EMBO J **16:**3066-77.

Mourez, M., M. Jehanno, E. Schneider, and E. Dassa. 1998. In vitro interaction between components of the inner membrane complex of the maltose ABC transporter of *Escherichia coli:* modulation by ATP. Mol Microbiol **30:**353-63.

Mowbray, S. L., and D. E. Koshland, Jr. 1987. Additive and independent responses in a single receptor: aspartate and maltose stimuli on the tar protein. Cell **50:**171-80.

Napoli, C. A., and D. H. Hubbell. 1975. Ultrastructure of *Rhizobium*-induced infection threads in clover root hairs. Appl Microbiol **30:**1003-9.

Newcomer, M. E., B. A. Lewis, and F. A. Quiocho. 1981. The radius of gyration of L-arabinose-binding protein decreases upon binding of ligand. J Biol Chem **256:**13218-22.

Nikaido, H., and J. A. Hall. 1998. Overview of bacterial ABC transporters. Methods Enzymol **292:**3-20.

Nikaido, K., and G. F. Ames. 1992. Purification and characterization of the periplasmic lysine-, arginine-, ornithine-binding protein (LAO) from *Salmonella typhimurium*. J Biol Chem **267:**20706-12.

Novelli, G., and J. K. Reichardt. 2000. Molecular basis of disorders of human galactose metabolism: past, present, and future. Mol Genet Metab **71:**62-5.

Novotny, M. J., W. L. Frederickson, E. B. Waygood, and M. H. Saier, Jr. 1985. Allosteric regulation of glycerol kinase by enzyme IIIglc of the phosphotransferase system in *Escherichia coli* and *Salmonella typhimurium*. J Bacteriol **162:**810-6.

Nutman, P. S. 1952. Studies on the physiology of nodule formation. III. Experiments on the excision of root-tips and nodules. Ann. Bot. **16:**79-101.

Nutman, P. S. 1962. The relation between root hair infection by *Rhizobium* and nodulation in *Trifolium* and *Vicia*. Proc. R. Soc. London Ser. B **156:**122-137.

Obmolova, G., C. Ban, P. Hsieh, and W. Yang. 2000. Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA. Nature **407:**703-10.

Ogawa, K. 1991. Four ATP-binding sites in the midregion of the beta heavy chain of dynein. Nature **352:**643-5.

Oh, B. H., G. F. Ames, and S. H. Kim. 1994. Structural basis for multiple ligand specificity of the periplasmic lysine-, arginine-, ornithine-binding protein. J Biol Chem **269:**26323-30.

Oh, B. H., J. Pandit, C. H. Kang, K. Nikaido, S. Gokcen, G. F. Ames, and S. H. Kim. 1993. Three-dimensional structures of the periplasmic lysine/arginine/ornithine-binding protein with and without a ligand. J Biol Chem **268:**11348-55.

Oldroyd, G. E., E. M. Engstrom, and S. R. Long. 2001. Ethylene inhibits the Nod factor signal transduction pathway of *Medicago truncatula*. Plant Cell **13:**1835-49.

Olsson, P. A., P. Kjellbom, and L. Rosendahl. 2002. *Rhizobium* colonization induced changes in membrane-bound and soluble hydroxyproline-rich glycoprotein composition in pea. Physiol Plant **114:**652-660.

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:**1233-43.

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 leguminosaurm* bv. *trifolii*: evidence for plant-inducible rhamnose locus involved in competition for nodulation. Mol. Plant-Microbe Intract. **186:**8433-8442.

Oresnik, I. J., S. Twelker, and M. F. Hynes. 1999. Cloning and characterization of a *Rhizobium leguminosarum* gene encoding a bacteriocin with similarities to RTX toxins. Appl Environ Microbiol **65:**2833-40.

Ormo, M., C. E. Bystrom, and S. J. Remington. 1998. Crystal structure of a complex of *Escherichia coli* glycerol kinase and an allosteric effector fructose 1,6-bisphosphate. Biochemistry **37:**16565-72.

Ostermeier, C., and H. Michel. 1997. Crystallization of membrane proteins. Curr Opin Struct Biol **7:**697-701.

Ouellette, M., D. Legare, and B. Papadopoulou. 2001. Multidrug resistance and ABC transporters in parasitic protozoa. J Mol Microbiol Biotechnol **3:**201-6.

Page, R. D. M. 1996. TREEVIEW: An application to display phylogenetic trees on personal computers. Computer Applications in the Biosciences **12:**357-358.

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:**535-46.

Paquet, J. Y., M. A. Diaz, S. Genevrois, M. Grayon, J. M. Verger, X. de Bolle, J. H. Lakey, J. J. Letesson, and A. Cloeckaert. 2001. Molecular, antigenic, and functional analyses of Omp2b porin size variants of *Brucella* spp. J Bacteriol **183:**4839-47.

Park, S. H., I. Pastuszak, R. Drake, and A. D. Elbein. 1998. Purification to apparent homogeneity and properties of pig kidney L-fucose kinase. J Biol Chem **273:**5685-91.

Parkhill, J. 2002. p. 33. *In* B. Wren, Dorrell, N. (ed.), Annotation of microbial genomics. Academic Press, London.

Pauly, H. E., and G. Pfleiderer. 1975. D-glucose dehydrogenase from *Bacillus megaterium* M 1286: purification, properties and structure. Hoppe Seylers Z Physiol Chem **356:**1613-23.

Pavlik, P., M. Simon, T. Schuster, and H. Ruis. 1993. The glycerol kinase (GUT1) gene of *Saccharomyces cerevisiae*: cloning and characterization. Curr Genet **24:**21-5.

Pawlyk, A. C., and D. W. Pettigrew. 2002. Transplanting allosteric control of enzyme activity by protein-protein interactions: coupling a regulatory site to the conserved catalytic core. Proc Natl Acad Sci U S A **99:**11115-20.

Payen, L. F., M. Gao, C. J. Westlake, S. P. Cole, and R. G. Deeley. 2003. Role of carboxylate residues adjacent to the conserved core Walker B motifs in the

catalytic cycle of multidrug resistance protein 1 (ABCC1). J Biol Chem **278:**38537-47.

Pelham, H. 1988. Heat-shock proteins. Coming in from the cold. Nature **332:**776-7.

Pelham, H. R. 1986. Speculations on the functions of the major heat shock and glucose-regulated proteins. Cell **46:**959-61.

Peng, W. T., Y. W. Lee, and E. W. Nester. 1998. The phenolic recognition profiles of the *Agrobacterium tumefaciens* VirA protein are broadened by a high level of the sugar binding protein ChvE. J Bacteriol **180:**5632-8.

Penmetsa, R. V., and D. R. Cook. 1997. A legume ethylene-insensitive mutant hyperinfected by its rhizobial symbiont. Science **275:**527-30.

Penmetsa, R. V., J. A. Frugoli, L. S. Smith, S. R. Long, and D. R. Cook. 2003. Dual genetic pathways controlling nodule number in *Medicago truncatula*. Plant Physiol **131:**998-1008.

Peters, N. K., and D. K. Crist-Estes. 1989. Nodule formation is stimulated by the ethylene inhibitor aminoethoxyvinylglycine. Plant Physiol **91:**690-693.

Petry, K. G., and J. K. Reichardt. 1998. The fundamental importance of human galactose metabolism: lessons from genetics and biochemistry. Trends Genet **14:**98-102.

Pettigrew, D. W., G. B. Smith, K. P. Thomas, and D. C. Dodds. 1998. Conserved active site aspartates and domain-domain interactions in regulatory properties of the sugar kinase superfamily. Arch Biochem Biophys **349:**236-45.

Pettigrew, D. W., G. J. Yu, and Y. Liu. 1990. Nucleotide regulation of *Escherichia coli* glycerol kinase: initial-velocity and substrate binding studies. Biochemistry **29:**8620-7.

Pflugrath, J. W., and F. A. Quiocho. 1985. Sulphate sequestered in the sulphate-binding protein of *Salmonella typhimurium* is bound solely by hydrogen bonds. Nature **314:**257-60.

Phillips, D. A., and Streit, W. R. 1996. p. 236–271. *In* G. Stacey, and Keen, N.T. (ed.), Plant microbe interactions. Chapman and Hall, New York.

Phillips, D. A., Dakora, F.D., Sande, E., Joseph, C.M., and Zon, J. 1994. Synthesis, release and transmission of alfalfa signals to rhizobial symbionts. Plant Soil **161:**69-80.

Phillips, D. A., E. S. Sande, J. A. C. Vriezen, F. J. de Bruijn, D. Le Rudulier, and C. M. Joseph. 1998. A new genetic locus in *Sinorhizobium meliloti* is involved in stachydrine utilization. Appl Environ Microbiol **64:**3954-60.

Poelarends, G. J., Mazurkiewicz, P., Konings, W. N. 2002. Multidrug transporters and antibiotic resistance in *Lactococcus lactis*. Biochim Biophys Acta **1555:**1-7.

Poole, P., and D. Allaway. 2000. Carbon and nitrogen metabolism in *Rhizobium*. Adv Microb Physiol **43:**117-63.

Postma, P. W., Epstein, W., Schuitema, A. R., Nelson, S. O. 1984. Interaction between IIIGlc of the phosphoenolpyruvate:sugar phosphotransferase system and glycerol kinase of *Salmonella typhimurium*. J Bacteriol **158:**351-3.

Power, J. 1967. The L-rhamnose genetic system in *Escherichia coli* K-12. Genetics **55:**557-68.

Priefer, U. B. 1989. Genes involved in lipopolysaccharide production and symbiosis are clustered on the chromosome of *Rhizobium leguminosarum* biovar *viciae* VF39. J Bacteriol **171:**6161-8.

Prossnitz, E., K. Nikaido, S. J. Ulbrich, and G. F. Ames. 1988. Formaldehyde and photoactivatable cross-linking of the periplasmic binding protein to a membrane component of the histidine transport system of *Salmonella typhimurium*. J Biol Chem **263:**17917-20.

Pugsley, A. P. 1993. The complete general secretory pathway in gram-negative bacteria. Microbiol Rev **57:**50-108.

Purchase, H. F. 1958. Restriction of infection threads in nodulation of clover and lucerne. Aust. J. Biol. Sci. **11:**155-161.

Qu, Q., and F. J. Sharom. 2001. FRET analysis indicates that the two ATPase active sites of the P-glycoprotein multidrug transporter are closely associated. Biochemistry **40:**1413-22.

Quandt, J., and M. F. Hynes. 1993. Versatile suicide vectors which allow direct selection for gene replacement in gram-negative bacteria. Gene **127:**15-21.

Quiocho, F. A., and P. S. Ledvina. 1996. Atomic structure and specificity of bacterial periplasmic receptors for active transport and chemotaxis: variation of common themes. Mol Microbiol **20:**17-25.

Quiocho, F. A., J. C. Spurlino, and L. E. Rodseth. 1997. Extensive features of tight oligosaccharide binding revealed in high-resolution structures of the maltodextrin transport/chemosensory receptor. Structure **5:**997-1015.

Quispel, A. 1974. *In* A. Quispel (ed.), The biology of nitrogen fixation. North-Holland Press Co., Amsterdam.

Rae, A. L., P. Bonfante-Fasolo, and N. J. Brewin. 1992. Structure and growth of infection threads in the legume symbiosis with *Rhizobium leguminosarum*. Plant J. **2:**385-395.

Rayment, I. R., W. R. Schmidt-Base, K. Smith, R. Tomchick, D. R. Benning, M. M. Winkelmann, D. A. Wesenberg, G. Holden, H. M. 1993. Three-dimensional structure of myosin subfragment-1: a molecular motor. Science **261:**50-8.

Rea, P. A., Li, Z.S., Lu, Y.P., Drozdowicz, Y.M., Martinoia, E. 1998. From vacuolar Gs-X pumps to multispecific ABC transporters. Annu Rev Plant Physiol Plant Mol Biol **49:**727-760.

Reizer, A., J. Deutscher, M. H. Saier, Jr., and J. Reizer. 1991. Analysis of the gluconate (*gnt*) operon of *Bacillus subtilis*. Mol Microbiol **5:**1081-9.

Richardson, J. S., X. Carpena, J. Switala, I. Fita, P. C. Loewen, and I. J. Oresnik. 2007. RhaU of *Rhizobium leguminosarum* is a rhamnose mutarotase. Submitted to the Journal of Bacteriology.

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:**8433-42.

Richardson, J. S., and I. J. Oresnik. 2007. Rhamnose transport is sugar-kinase (RhaK) dependent in *Rhizobium leguminosarum* bv. *trifolii*. Submitted to the Journal of Bacteriology.

Ridge, R. W. 1992. A model of legume root hair growth and *Rhizobium* infection. Symbiosis **14:**349–373.

Ridge, R. W., and B. G. Rolfe. 1985. *Rhizobium* sp. degradation of legume root hair cell wall at the site of infection thread origin. Appl Environ Microbiol **50:**717-720.

Riordan, J. R. 1993. The cystic fibrosis transmembrane conductance regulator. Annu Rev Physiol **55:**609-30.

Riordan, J. R., J. M. Rommens, B. Kerem, N. Alon, R. Rozmahel, Z. Grzelczak, J. Zielenski, S. Lok, N. Plavsic, J. L. Chou, and et al. 1989. Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. Science **245:**1066-73.

Robleto, E. A., K. Kmiecik, E. S. Oplinger, J. Nienhuis, and E. W. Triplett. 1998. Trifolitoxin production increases nodulation competitiveness of *Rhizobium etli* CE3 under agricultural conditions. Appl Environ Microbiol **64:**2630-3.

Rosenberg, C., P. Boistard, J. Denarie, and F. Casse-Delbart. 1981. Genes controlling early and late functions in symbiosis are located on a megaplasmid in *Rhizobium* meliloti. Mol Gen Genet **184:**326-33.

Rust, S., M. Rosier, H. Funke, J. Real, Z. Amoura, J. C. Piette, J. F. Deleuze, H. B. Brewer, N. Duverger, P. Denefle, and G. Assmann. 1999. Tangier disease is caused by mutations in the gene encoding ATP-binding cassette transporter 1. Nat Genet **22:**352-5.

Rutberg, B. 1997. Antitermination of transcription of catabolic operons. Mol Microbiol **23:**413-21.

Ruvkun, G. B., and F. M. Ausubel. 1981. A general method for site-directed mutagenesis in prokaryotes. Nature **289:**85-8.

Ryu, K. S., C. Kim, I. Kim, S. Yoo, B. S. Choi, and C. Park. 2004. NMR application probes a novel and ubiquitous family of enzymes that alter monosaccharide configuration. J Biol Chem **279:**25544-8.

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:**153-62.

Sack, J. S., M. A. Saper, and F. A. Quiocho. 1989. Periplasmic binding protein structure and function. Refined X-ray structures of the leucine/isoleucine/valine-binding protein and its complex with leucine. J Mol Biol **206:**171-91.

Saier, M. H., Jr. 1999. A functional-phylogenetic system for the classification of transport proteins. J Cell Biochem **Suppl 32-33:**84-94.

Saier, M. H., Jr. 2000. A functional-phylogenetic classification system for transmembrane solute transporters. Microbiol Mol Biol Rev **64:**354-411.

Sambrook, J., E.F. Fritsch, and T. A. Maniatis 1989. *In* J. Sambrook, E.F. Fritsch, and T. A. Maniatis (ed.), Molecular Cloning A Laboratory Manual 2nd edition Cold Spring Harbour Laboratory Press, Cold Spring Harbor, NY.

Sambrook, J., and D. W. Russell. 2001. *In* J. Sambrook, and D. W. Russell (ed.), Molecular cloning a laboratory manual 3rd edition. Cold Spring Harbour Laboratory Press, Cold Spring Harbor, NY.

Saraste, M., P. R. Sibbald, and A. Wittinghofer. 1990. The P-loop--a common motif in ATP- and GTP-binding proteins. Trends Biochem Sci **15:**430-4.

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:**22-41.

Sawada, H., and Y. Takagi. 1964. The Metabolism of L-Rhamnose in *Escherichia coli*. 3. L-Rhamnulose-Phosphate Aldolase. Biochim Biophys Acta **92:**26-32.

Schmitt, L., H. Benabdelhak, M. A. Blight, I. B. Holland, and M. T. Stubbs. 2003. Crystal structure of the nucleotide-binding domain of the ABC-transporter haemolysin B: identification of a variable region within ABC helical domains. J Mol Biol **330:**333-42.

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

Schoepfer, R. 1993. The pRSET family of T7 promoter expression vectors for *Escherichia coli*. Gene **124:**83-5.

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

Schulz, G. E., M. Elzinga, F. Marx, and R. H. Schrimer. 1974. Three dimensional structure of adenyl kinase. Nature **250:**120-3.

Schutt, C. E., J. C. Myslik, M. D. Rozycki, N. C. Goonesekere, and U. Lindberg. 1993. The structure of crystalline profilin-beta-actin. Nature **365:**810-6.

Senior, A. E. 1998. Catalytic mechanism of P-glycoprotein. Acta Physiol Scand Suppl **643:**213-8.

Senior, A. E., and D. C. Gadsby. 1997. ATP hydrolysis cycles and mechanism in P-glycoprotein and CFTR. Semin Cancer Biol **8:**143-50.

Serrano, L., J. Sancho, M. Hirshberg, and A. R. Fersht. 1992. Alpha-helix stability in proteins. I. Empirical correlations concerning substitution of side-chains at the N and C-caps and the replacement of alanine by glycine or serine at solvent-exposed surfaces. J Mol Biol **227:**544-59.

Sharff, A. J., L. E. Rodseth, and F. A. Quiocho. 1993. Refined 1.8-A structure reveals the mode of binding of beta-cyclodextrin to the maltodextrin binding protein. Biochemistry **32:**10553-9.

Sharma, S., and A. L. Davidson. 2000. Vanadate-induced trapping of nucleotides by purified maltose transport complex requires ATP hydrolysis. J Bacteriol **182:**6570-6.

Shaw, P. D. 1987. Methods for multilocus enzyme electrophoresis for bacterial population genetics and systematics., p. 3-39. *In* T. Kosuge, and Nester, E. W. (ed.), Plant-microbe interactions: Molecular and genetics perspectives, vol. 2. Macmillan Publishing Co., New York.

Sheterline, P., J. Clayton, and J. Sparrow. 1995. Actin. Protein Profile **2:**1-103.

Shilton, B. H., M. M. Flocco, M. Nilsson, and S. L. Mowbray. 1996. Conformational changes of three periplasmic receptors for bacterial chemotaxis and transport: the maltose-, glucose/galactose- and ribose-binding proteins. J Mol Biol **264:**350-63.

Sieberer, B., and A. M. C. Emons. 2000. Cytoarchitecture and pattern of cytoplasmic streaming in root hairs of *Medicago truncatula* during development and deformation by nodulation factors. Protoplasma **214:**118-127.

Sieberer, B. J., A. C. Timmers, F. G. Lhuissier, and A. M. Emons. 2002. Endoplasmic microtubules configure the subapical cytoplasm and are required for fast growth of *Medicago truncatula* root hairs. Plant Physiol **130:**977-88.

Sigrell, J. A., A. D. Cameron, and S. L. Mowbray. 1999. Induced fit on sugar binding activates ribokinase. J Mol Biol **290:**1009-18.

Simon, R., Priefer, U., and Pühler, A. . 1983. A broad host range mobilization system for *in vivo* engineering: transposon mutagenesis in gram-negative bacteria. Bio/Techniques **1:**784-791.

Simossis, V. A., and J. Heringa. 2003. The PRALINE online server: optimizing progressive multiple alignment on the web. Comput Biol Chem **27:**511-9.

Simossis, V. A., and J. Heringa. 2004. Integrating protein secondary structure prediction and multiple sequence alignment. Curr Protein Pept Sci **5:**249-66.

Simossis, V. A., and J. Heringa. 2005. PRALINE: a multiple sequence alignment toolbox that integrates homology-extended and secondary structure information. Nucleic Acids Res **33:**W289-94.

Simossis, V. A., J. Kleinjung, and J. Heringa. 2005. Homology-extended sequence alignment. Nucleic Acids Res **33:**816-24.

Smit, G., J. W. Kijne, and B. J. Lugtenberg. 1987. Involvement of both cellulose fibrils and a Ca^{2+}-dependent adhesin in the attachment of *Rhizobium leguminosarum* to pea root hair tips. J Bacteriol **169:**4294-301.

Smit, G., J. W. Kijne, and B. J. Lugtenberg. 1989. Roles of flagella, lipopolysaccharide, and a Ca^{2+}-dependent cell surface protein in attachment of *Rhizobium leguminosarum* biovar viciae to pea root hair tips. J Bacteriol **171:**569-72.

Smith, A. J. 1982. p. 47-85. *In* N. G. Carr, and Whitton, B. A. (ed.), The biology of cyanobacteria. University of California Press, Berkeley, Ca.

Smith, C. A., and I. Rayment. 1996. Active site comparisons highlight structural similarities between myosin and other P-loop proteins. Biophys J **70:**1590-602.

Smith, P. C., N. Karpowich, L. Millen, J. E. Moody, J. Rosen, P. J. Thomas, and J. F. Hunt. 2002. ATP binding to the motor domain from an ABC transporter drives formation of a nucleotide sandwich dimer. Mol Cell **10:**139-49.

Spaink, H. P., A. Kondorosi, and P. J. J. Hooykaas. 1998. p. 1-566. *In* H. P. Spaink, A. Kondorosi, and P. J. J. Hooykaas (ed.), The *Rhizobiaceae*: molecular biology of model plant-associated bacteria. Springer, Kluwer Academic Publishers, Dordrecht, The Netherlands.

Staiger, C. J., M. Yuan, R. Valenta, P. J. Shaw, R. M. Warn, and C. W. Lloyd. 1994. Microinjected profilin affects cytoplasmic streaming in plant cells by rapidly depolymerizing actin microfilaments. Curr Biol **4:**215-9.

Steitz, T. A., M. Shoham, and W. S. Bennett, Jr. 1981. Structural dynamics of yeast hexokinase during catalysis. Philos Trans R Soc Lond B Biol Sci **293:**43-52.

Streeter, J. 1988. Inhibition of legume nodule formation and nitrogen fixation by nitrate. CRC Crit. Rev. Plant Sci. **7:**7-23.

Streit, W. R., C. M. Joseph, and D. A. Phillips. 1996. Biotin and other water-soluble vitamins are key growth factors for alfalfa root colonization by *Rhizobium meliloti* 1021. Mol Plant Microbe Interact **9:**330-8.

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 U S A **95:**5145-9.

Sundararajan, T. A., A. M. Rapin, and H. M. Kalckar. 1962. Biochemical observations on *E. coli* mutants defective in uridine diphosphoglucose. Proc Natl Acad Sci U S A **48:**2187-93.

Sutcliffe, I. C., and R. R. Russell. 1995. Lipoproteins of gram-positive bacteria. J Bacteriol **177:**1123-8.

Swain, C. G., and Brown, J. F. 1952. Concerted displacement reactions VII: the mechanism of acid-base catalysis in non-aqueous solvents. J. Amer. Chem. Soc. **74:**2534.

Sy, A., E. Giraud, P. Jourand, N. Garcia, A. Willems, P. de Lajudie, Y. Prin, M. Neyra, M. Gillis, C. Boivin-Masson, and B. Dreyfus. 2001. Methylotrophic *Methylobacterium* bacteria nodulate and fix nitrogen in symbiosis with legumes. J Bacteriol **183:**214-20.

Szakacs, G., C. Ozvegy, E. Bakos, B. Sarkadi, and A. Varadi. 2000. Transition-state formation in ATPase-negative mutants of human MDR1 protein. Biochem Biophys Res Commun **276:**1314-9.

Takagi, Y., and H. Sawada. 1964a. The Metabolism of L-Rhamnose in *Escherichia coli*. Biochim Biophys Acta **92:**10-7.

Takagi, Y., and H. Sawada. 1964b. The Metabolism of L-Rhamnose in *Escherichia coli*. Biochim Biophys Acta **92:**18-25.

Tatusov, R. L., E. V. Koonin, and D. J. Lipman. 1997. A genomic perspective on protein families. Science **278:**631-7.

Thoden, J. B., and H. M. Holden. 2002. High resolution X-ray structure of galactose mutarotase from *Lactococcus lactis*. J Biol Chem **277:**20854-61.

Thoden, J. B., J. Kim, F. M. Raushel, and H. M. Holden. 2002. Structural and kinetic studies of sugar binding to galactose mutarotase from *Lactococcus lactis*. J Biol Chem **277:**45458-65.

Thoden, J. B., J. Kim, F. M. Raushel, and H. M. Holden. 2003. The catalytic mechanism of galactose mutarotase. Protein Sci **12:**1051-9.

Thomas, P. J., and J. F. Hunt. 2001. A snapshot of Nature's favorite pump. Nat Struct Biol **8:**920-3.

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:**4876-82.

Thorner, J. W., and H. Paulus. 1973. Catalytic and allosteric properties of glycerol kinase from *Escherichia coli*. J Biol Chem **248:**3922-32.

Thurman, N. P., D. M. Lewis, and D. G. Jones. 1985. The relationship of plasmid number to growth, acid tolerance and symbiotic efficiency in isolates of *R. trifolii.* J. Appl. Bacteriol. **58:**1-6.

Timmers, A. C., M. C. Auriac, F. de Billy, and G. Truchet. 1998. Nod factor internalization and microtubular cytoskeleton changes occur concomitantly during nodule differentiation in alfalfa. Development **125:**339-49.

Timmers, A. C., M. C. Auriac, and G. Truchet. 1999. Refined analysis of early symbiotic steps of the *Rhizobium-Medicago* interaction in relationship with microtubular cytoskeleton rearrangements. Development **126:**3617-28.

Timson, D. J., and R. J. Reece. 2003. Identification and characterization of human aldose-1-epimerase. FEBS Lett **543:**21-4.

Tominaga, M., E. Yokota, S. Sonobe, and T. Shimmen. 2000. Mechanism of inhibition of cytoplasmic streaming by a myosin inhibitor, 2,3-butanedione monoxime. Protoplasma **213:**46-54.

Tominaga, M., K. Morita, S. Sonobe, E. Yokota, and T. Shimmen. 1997. Microtubules regulate the organization of actin filaments at the cortical region in root hair cells of *Hydrocharis*. Protoplasma **199:**83–92.

Toro, N., and J. Olivares. 1986. Analysis of *Rhizobium meliloti* sym mutants obtained by heat treatment. Appl Environ Microbiol **51:**1148-1150.

Triplett, E. W., and M. J. Sadowsky. 1992. Genetics of competition for nodulation of legumes. Annu Rev Microbiol **46:**399-428.

Tsuruo, T., H. Iida, S. Tsukagoshi, and Y. Sakurai. 1981. Overcoming of vincristine resistance in P388 leukemia in vivo and in vitro through enhanced cytotoxicity of vincristine and vinblastine by verapamil. Cancer Res **41:**1967-72.

Turgeon, B. G., and W. D. Bauer. 1985. Ultrastructure of infection-thread development during the infection of soybean by *Rhizobium japonicum.* Planta **163:**328–349.

Urbatsch, I. L., K. Gimi, S. Wilke-Mounts, and A. E. Senior. 2000. Investigation of the role of glutamine-471 and glutamine-1114 in the two catalytic sites of P-glycoprotein. Biochemistry **39:**11921-7.

Urbatsch, I. L., G. A. Tyndall, G. Tombline, and A. E. Senior. 2003. P-glycoprotein catalytic mechanism: studies of the ADP-vanadate inhibited state. J Biol Chem **278:**23171-9.

Van Bastelaere, E., M. Lambrecht, H. Vermeiren, A. Van Dommelen, V. Keijers, P. Proost, and J. Vanderleyden. 1999. Characterization of a sugar-binding protein from *Azospirillum brasilense* mediating chemotaxis to and uptake of sugars. Mol Microbiol **32:**703-14.

van Brussel, A. A. N., R. Bakhuizen, P. C. van Spronsen, H. P. Spaink, T. Tak, B. J. J. Lugtenberg, and J. W. Kijne. 1992. Induction of pre-infection thread structures in the leguminous host plant by mitogenic lipo-oligosaccharidess of *Rhizobium*. Science **257:**70-72.

Van den bosch, K. A., D. J. Bradley, J. P. Knox, S. Perotto, G. W. Butcher, and N. J. Brewin. 1989. Common components of the infection thread matrix and the intercellular space identified by immunocytochemical analysis of pea nodules and uninfected roots. Embo J **8:**335-341.

van den Ent, F., and J. Lowe. 2000. Crystal structure of the cell division protein FtsA from *Thermotoga maritima*. EMBO J **19:**5300-7.

van der Heide, T., and B. Poolman. 2000. Osmoregulated ABC-transport system of *Lactococcus lactis* senses water stress via changes in the physical state of the membrane. Proc Natl Acad Sci U S A **97:**7102-6.

van der Heide, T., and B. Poolman. 2002. ABC transporters: one, two or four extracytoplasmic substrate-binding sites? EMBO Rep **3:**938-43.

van Rhijn, P., N. A. Fujishige, P. O. Lim, and A. M. Hirsch. 2001. Sugar-binding activity of pea lectin enhances heterologous infection of transgenic alfalfa plants by *Rhizobium leguminosarum* biovar *viciae*. Plant Physiol **126:**133-44.

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

van Spronsen, P. C., R. Bakhuizen, A. A. van Brussel, and J. W. Kijne. 1994. Cell wall degradation during infection thread formation by the root nodule bacterium *Rhizobium leguminosarum* is a two-step process. Eur J Cell Biol **64:**88-94.

van Veen, H. W., and W. N. Konings. 1997. Multidrug transporters from bacteria to man: similarities in structure and function. Semin Cancer Biol **8:**183-91.

van Veen, H. W., A. Margolles, M. Muller, C. F. Higgins, and W. N. Konings. 2000. The homodimeric ATP-binding cassette transporter LmrA mediates multidrug transport by an alternating two-site (two-cylinder engine) mechanism. EMBO J **19:**2503-14.

van Workum, W. A. T., S. van Slageren, A. A. N. van Brussel, and J. W. Kijne. 1998. Role of exopolysaccharides of *Rhizobium leguminosarum* bv. *viciae* as host plant-specific molecules required for infection thread formation during nodulation of *Vicia sativa.* Mol. Plant-Microbe Interact. **11:**1233-1241.

Vasse, J., F. de Billy, S. Camut, and G. Truchet. 1990. Correlation between ultrastructural differentiation of bacteroids and nitrogen fixation in alfalfa nodules. J Bacteriol **172:**4295-306.

Vasse, J., F. de Billy, and G. Truchet. 1993. Abortion of infection during the *Rhizobium meliloti*-alfalfa symbiotic interaction is accompanied by a hypersensitive reaction. Plant J. **4:**555-566.

Venter, A. P., S. Twelker, I. J. Oresnik, and M. F. Hynes. 2001. Analysis of the genetic region encoding a novel rhizobiocin from *Rhizobium leguminosarum* bv. *viciae* strain 306. Can J Microbiol **47:**495-502.

Verdon, G., S. V. Albers, B. W. Dijkstra, A. J. Driessen, and A. M. Thunnissen. 2003. Crystal structures of the ATPase subunit of the glucose ABC transporter from *Sulfolobus solfataricus*: nucleotide-free and nucleotide-bound conformations. J Mol Biol **330:**343-58.

Verma, D. P. S. 1988. *In* J. W. N. Akkerman (ed.), Molecular signals in plant microbe communities. CRC Press, Boca Raton, FL.

Vidali, L., S. T. McKenna, and P. K. Hepler. 2001. Actin polymerization is essential for pollen tube growth. Mol Biol Cell **12:**2534-45.

Vincent, J. M. 1970. A manual for the practical study of root-nodule bacteria. Blackwell Scientific Publications, Oxford, England.

Vitousek, P. M., J. D. Aber, R. W. Howarth, G. E. Likens, P. A. Matson, D. W. Schindler, W. H. Schlesinger, and D. G. Tilman. 1997. Human alteration of the global nitrogen cycle: sources and consequences. Ecol. Appl. **7:**737–750.

Voegele, R. T., G. D. Sweet, and W. Boos. 1993. Glycerol kinase of *Escherichia coli* is activated by interaction with the glycerol facilitator. J Bacteriol **175:**1087-94.

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:**945-51.

Wallenfels, K., F. Hucho, and K. Herrmann. 1965. Enzymatically catalyzed mutarotation of aldoses. Studies on aldose-1-epimerase from *E. coli*. Biochem Z **343:**307-25.

Wallenfels, K., and G. Kurz. 1962. On the specificity of galactose dehydrogenase from *Pseudomonas saccharophila* and its use as an analytical aid. Biochem Z **335:**559-73.

Walshaw, D. L., and P. S. Poole. 1996. The general L-amino acid permease of *Rhizobium leguminosarum* is an ABC uptake system that influences efflux of solutes. Mol. Microbiol. **21:**1239-1252.

Wang, E. J., C. N. Casciano, R. P. Clement, and W. W. Johnson. 2000. Cooperativity in the inhibition of P-glycoprotein-mediated daunorubicin transport: evidence for half-of-the-sites reactivity. Arch Biochem Biophys **383:**91-8.

Watson, R. J., Y. K. Chan, R. Wheatcroft, A. F. Yang, and S. H. Han. 1988. *Rhizobium meliloti* genes required for C4-dicarboxylate transport and symbiotic nitrogen fixation are located on a megaplasmid. J Bacteriol **170:**927-34.

Widmer, F., B. T. Shaffer, L. A. Porteous, and R. J. Seidler. 1999. Analysis of *nifH* gene pool complexity in soil and litter at a Douglas fir forest site in the Oregon cascade mountain range. Appl Environ Microbiol **65:**374-80.

Wieland, O., and M. Suyter. 1957. Glycerokinase; isolation and properties of the enzyme. Biochem Z **329:**320-31.

Wilkinson, A. J., and K. H. G. Verschueren. 2003. Crystal structures of periplasmic solute-binding protein in ABC transport complexes illuminate their function, p.

187-207. *In* J. L. Linton, Cole, S. P. C., Kuchler, K., and Higgins, C. F. (ed.), ABC proteins from bacteria to man. Academic Press, London.

Wilkinson, J. F. 1949. The pathway of the adaptive fermentation of galactose by yeast. Biochem J **44:**460-7.

Wilson, D. B., and D. S. Hogness. 1969. The enzymes of the galactose operon in *Escherichia coli*. II. The subunits of uridine diphosphogalactose 4-epimerase. J Biol Chem **244:**2132-6.

Wolf, A., E. W. Shaw, K. Nikaido, and G. F. Ames. 1994. The histidine-binding protein undergoes conformational changes in the absence of ligand as analyzed with conformation-specific monoclonal antibodies. J Biol Chem **269:**23051-8.

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

Yanisch-Perron, C., Viera, J., Messing, J. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of M13mp18 and pUC19 Gene **33:**103-119.

Yarosh, O. K., T. C. Charles, and T. M. Finan. 1989. Analysis of C4-dicarboxylate transport genes in *Rhizobium meliloti*. Mol Microbiol **3:**813-23.

Yost, C. K., P. Rochepeau, and M. F. Hynes. 1998. *Rhizobium leguminosarum* contains a group of genes that appear to code for methyl-accepting chemotaxis proteins. Microbiology **144:**1945-56.

Young, J., and I. B. Holland. 1999. ABC transporters: bacterial exporters-revisited five years on. Biochim Biophys Acta **1461:**177-200.

Young, J. P., L. C. Crossman, A. W. Johnston, N. R. Thomson, Z. F. Ghazoui, K. H. Hull, M. Wexler, A. R. 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. Moule, K. Mungall, H. Norbertczak, E. Rabbinowitsch, M. Sanders, M. Simmonds, S. Whitehead, and J. Parkhill. 2006. The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. Genome Biol **7:**R34.

Yuan, Y. R., S. Blecker, O. Martsinkevich, L. Millen, P. J. Thomas, and J. F. Hunt. 2001. The crystal structure of the MJ0796 ATP-binding cassette. Implications for the structural consequences of ATP hydrolysis in the active site of an ABC transporter. J Biol Chem **276:**32313-21.

Zadrazil, S. 1973. Chromatography of nucleic acid components. *In* W. W. Zorbach, and Tipson, R. S. (ed.), Synthetic procedures in nucleic acid chemistry: physical and physiochemical aids in characterization and in determination of structure, vol. 2. Wiley-Interscience, New York.

Zehr, J. P., B. D. Jenkins, S. M. Short, and G. F. Steward. 2003. Nitrogenase gene diversity and microbial community structure: a cross-system comparison. Environ Microbiol **5:**539-54.

Zhang, Y., C. Conway, M. Rosato, Y. Suh, and M. D. Manson. 1992. Maltose chemotaxis involves residues in the N-terminal and C-terminal domains on the same face of maltose-binding protein. J Biol Chem **267:**22813-20.

Zhu, Y., and E. C. Lin. 1988. A mutant *crp* allele that differentially activates the operons of the *fuc* regulon in *Escherichia coli*. J Bacteriol **170:**2352-8.

Appendix A

A.1 Copyrighted Material Permission Confirmation

A.1.1 Elsevier Publishing Ltd. from Biochimica et biophysica acta

Dear Mr Richardson,

We hereby grant you permission to reprint the aforementioned material at no charge **in your thesis, in print and on the University of Manitoba web site** subject to the following conditions:

1. If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies.

2. Suitable acknowledgment to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows: "Reprinted from Publication title, Vol number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier".

3. Your thesis may be submitted to your institution in either print or electronic form.

4. Reproduction of this material is confined to the purpose for which permission is hereby given.

5. This permission is granted for non-exclusive world **English** rights only. For other languages please reapply separately for each one required. Permission excludes use in an electronic form. Should you have a specific electronic project in mind please reapply for permission.

This includes permission for the Library and Archives of Canada and the UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission.

Yours sincerely

Manuela Beis
Rights Assistant

Material requested:
From "Structure and association of ATP-binding cassette transporter nucleotide-binding domains"; Kerr, I. D.; 2002; Vol. 1561; 1; pp. 47-64.
Figure 7. Schematic representation of the NBD dimers (panel (C) Rad50 only).
Figure 6. Structure of a DNA repair ABC transporter-like domain (panel (A) only).

Figure 1.2. Model of ATPase Rad50 dimer association. Reprinted from Biochimica et biophysica acta, Volume 1561, Kerr, I. D., Structure and association of ATP-binding cassette transporter nucleotide-binding domains, pp. 47-64, copyright 2002, with permission from Elsevier.

A.1.2 Birkhäuser Publishing Ltd. from Cellular and Molecular Life Sciences (CMLS)

Dear Mr. Richardson,

Thank you very much for your permission request. We herewith grant you the desired permission under condition that full reference to the original source of publication is made. Please refer to Birkhäuser Verlag, Switzerland.

Best wishes for your work and kind regards,

Liv Etienne Rights & Licenses Manager
Birkhäuser Verlag AG
Viaduktstrasse 42
CH-4051 Basel
T +41 61 205 07 13
F +41 61 205 07 92
liv.etienne@birkhauser.ch
www.birkhauser.ch
-----Ursprüngliche Nachricht-----
Von: SpringerAlerts@springeronline.com
[mailto:SpringerAlerts@springeronline.com]
Gesendet: Mittwoch, 7. February 2007 19:54
An: Etienne, Liv, Birkhaeuser
Betreff: Copyright Request

Material requested:
From "The ABC transporter structure and mechanism: perspectives on recent research"; Jones, P. M., and George, A. M.; 2004; Vol. 61; 6; pp. 682-699.
Figure 2. ABC-ATPase catalytic site.
Figure 3. (A) Topology diagram of the consensus ABC transporter NBD fold.(B) Three-dimensional structure of ABC transporter NBD.

Figure 1.3. Consensus of the NBD fold and 3-D structure of ABC transporter ABC proteins NBD monomer. Reprinted from Cellular and Molecular Life Sciences (CMLS), volume **61**(6), Jones, P. M. and George, A. M., The ABC transporter structure and mechanism: perspectives on recent research, pp. 682-699, copyright 2004, with permission from Birkhäuser Publishing Ltd.

Figure 1.4. ABC ATPase catalytic core. Reprinted from Cellular and Molecular Life Sciences (CMLS), volume **61**(6), Jones, P. M. and George, A. M., The ABC transporter structure and mechanism: perspectives on recent research, pp. 682-699, copyright 2004, with permission from Birkhäuser Publishing Ltd.

A.1.3 Elsevier Publishing Ltd. from FEBS letters

We hereby grant you permission to reprint the aforementioned material at no charge in your thesis, in print and on the University of Manitoba web site subject to the following conditions:

1. If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies.

2. Suitable acknowledgment to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows: "Reprinted by permission of Federation of the European Biochemical Societies from (title of article), by (author) FEBS Letters, Vol No, Pages, Copyright year"

3. Your thesis may be submitted to your institution in either print or electronic form.

4. Reproduction of this material is confined to the purpose for which permission is hereby given.

5. This permission is granted for non-exclusive world English rights only. For other languages please reapply separately for each one required. Permission excludes use in an electronic form. Should you have a specific electronic project in mind please reapply for permission.

6. This includes permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission.

Yours sincerely,

Steph Smith (Rights Assistant, Elsevier Ltd., The Boulevard Langford Lane, Kidlington, Oxford, OX5 1GB)

Material requested:
From "ABC transporter architecture and mechanism: implications from the crystal structures of BtuCD and BtuF"; Locher, K. P., and Borths, E.; 2004; Vol. 564; 3; pp. 264-268.
Figure 2. Transmission interface.

Figure 1.5. TMD to ABC protein interface of BtuC-BtuD vitamin B_{12} ABC transporter. Reprinted by permission of Federation of the European Biochemical Societies from ABC transporter architecture and mechanism: implications from the crystal structures of BtuCD and BtuF, by Locher, K. P. and Borths, E., FEBS Letters, Volume **564**(3), pp. 264-268, copyright 2004.

A.1.4 Elsevier Publishing Ltd. from the Journal of Molecular Biology

We hereby grant you permission to reprint the aforementioned material at no charge in your thesis, in print and on the University of Manitoba web site subject to the following conditions:

1. If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies.
2. Suitable acknowledgment to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows: "Reprinted from Publication title, Vol number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier".
3. Your thesis may be submitted to your institution in either print or electronic form.
4. Reproduction of this material is confined to the purpose for which permission is hereby given.
5. This permission is granted for non-exclusive world English rights only. For other languages please reapply separately for each one required. Permission excludes use in an electronic form. Should you have a specific electronic project in mind please reapply for
permission.
6. This includes permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission.

Yours sincerely,

Steph Smith (Rights Assistant, Elsevier Ltd., The Boulevard Langford Lane, Kidlington, Oxford, OX5 1GB)

Material requested:
From "Structural insights into the monosaccharide specificity of *Escherichia coli* rhamnose mutarotase"; Ryu, K. S., Kim, J. I., Cho, S.J., Park, D., Park, C., Cheong, H. K., Lee, J. O., and Choi, B. S.; 2005; Vol. 349; 1; pp. 153-162.
Figure 5. The active site configurations of YiiL (stereo-view), RbsD, and galactose mutarotases. (panel (A) only).
Figure 7. Configuration of the active-site core of YiiL.

Figure 1.7. Active site configuration of L-rhamnose mutarotase YiiL in *E. coli*. Reprinted from the Journal of Molecular Biology, volume **349**(1), Ryu, K. S., Kim, J. I., Cho, S. J., Park, D., Park, C., Cheong, H. K., Lee, J. O. and Choi, B. S., Structural insights into the monosaccharide specificity of *Escherichia coli* rhamnose mutarotase, pp. 153-162, copyright 2005, with permission from Elsevier.

A.1.5 Proceedings of the National Academy of Sciences of the United States of America and Dr. Peer Bork

From: Josie Venturillo jventuri@highwire.stanford.edu
Dear Jason S. Richardson,

Thank you for contacting PNAS. To request for copyright permission, please see the permissions page at: http://www.pnas.org/misc/rightperm.shtml.

Please let us know if you have any further questions.

Regards,
Josie Venturillo

Rights and Permissions Copyright (http://www.pnas.org/misc/rightperm.shtml)
Volumes 90-102, copyright 1993-2005 by The National Academy of Sciences of the United States of America, all rights reserved. Volumes 1-89, copyright as a collective work only; author(s) retains copyright to individual articles.

To: Dr. Peer Bork

I am writing to request permission to include the following material from Proceedings of the National Academy of Sciences of the United States of America, in a literature review section of a graduate thesis. The thesis, entitled Characterization of L-Rhamnose Catabolism in Rhizobium leguminosarum is part of my requirements as a student in the Ph. D. of Science department. My thesis/practicum will be microfiched by Library and Archives Canada (LAC). A microfiche copy of the theses/practicum will be available for loan from LAC, and will be reproduced and sold in various formats through LACs agent, University Microfilms International (UMI). My thesis/practicum may also be posted electronically in the University of Manitoba digital repository (https://mspace.lib.umanitoba.ca/dspace/index.jsp), and be accessible to a worldwide audience from LACs Theses Canada (www.collectionscanada.ca/thesescanada). Access to the thesis in digital format is free of charge and therefore, no profit will be realized from the work. Please respond in writing to confirm whether or not you will grant permission for the above-mentioned work to be included in my thesis, as well as any stipulations or royalties that you request.

Thank you very much for your consideration of this request.

Jason Richardson

Material requested:
From "An ATPase domain common to prokaryotic cell cycle proteins, sugar kinases, actin, and hsp70 heat shock proteins"; Bork, P., Sander, C., and Valencia, A.; 1992; Vol. 89; 16; pp. 7290-7294.
Figure 1. Diagrammatic representation of actin.

From: Peer Bork bork@embl.de
Date: Thu, 15 Feb 2007 22:36:13 +0100 (CET)

fine
bw
Peer .p

Figure 1.8. Representation of actin highlighting the common motifs of the kinase-hsp70-actin protein superfamily. Reprinted from The Proceedings of the National Academy of Sciences of the United States of America, volume 89, Bork, P., C. Sander, and A. Valencia, An ATPase domain common to prokaryotic cell cycle proteins, sugar kinases, actin, and Hsp70 heat shock proteins, pp. 7290-7294, copyright 1992, with permission by Peer Bork.

A.1.6 Annual Review of Biophysics and Biomolecular Structure

Dear Dr. Richardson:

Thank you for your request for permission to reprint the figure cited in your e-mail below. We are happy to grant you permission to use this material in your thesis. Please use the following acknowledgment: "Reprinted, with permission, from the Annual Review of Biophysics and Biomolecular Structure, Volume 25 (c) 1996 by Annual Reviews www.annualreviews.org"

This permission to reprint is for a one-time usage only and any subsequent use of this material requires submission of a new permission request. Fees for this noncommercial usage have been waived for you. If I can be of further assistance, please do not hesitate to contact me.

Best wishes for success with your thesis.

Sincerely,
Laura Folkner (Permissions Department, ANNUAL REVIEWS, A Nonprofit Scientific Publisher, 4139 El Camino Way, Palo Alto, CA., 94306, Ph: (650) 843-6636, Fax: (650) 855-9815, lfolkner@annualreviews.org)

Material requested:
From "The sugar kinase/heat shock protein 70/actin superfamily: implications of conserved structure for mechanism"; Hurley, J. H.; 1996; Vol. 25; pp. 137-162
Figure 2. Schematic of the topology of the conserved core subdomains.

Figure 1.9. Diagrammatic representation of the topology of the conserved core in members of the kinase-hsp70-actin protein superfamily. Reprinted, with permission, from the Annual Review of Biophysics and Biomolecular Structure, Volume 25 © 1996 by Annual Reviews www.annualreviews.org.

A.1.7 AAAS Science and Dr. James H. Hurley

Dear Dr. Richardson,

Our permission letter with terms and conditions for use is attached below. Please let me know if you have trouble accessing it.

Best wishes,

Emilie L. David (Administrator, Rights, Permissions & Licensing, The American Association for the Advancement of Science, 1200 New York Ave., NW, Washington, DC 20005, Ph. (202) 326-7074, fax (202) 682-0816, edavid@aaas.org)

To: Dr. James H. Hurley:

I am writing to request permission to include the following material from Science, in a graduate thesis. The thesis, entitled Cataboilsm of L-rhamnose in Rhizobium leguminosarum is part of my requirements as a student in the Ph. D. department. My thesis/practicum will be microfiched by Library and Archives Canada (LAC). A microfiche copy of the theses/practicum will be available for loan from LAC, and will be reproduced and sold in various formats through LAC's agent, University Microfilms International (UMI). My thesis/practicum may also be posted electronically in the University of Manitoba digital repository (https://mspace.lib.umanitoba.ca/dspace/index.jsp), and be accessible to a worldwide audience from LAC's Theses Canada (www.collectionscanada.ca/thesescanada). I am requesting permission to reprint the figure of schematic diagram of glycerol and ADP binding in glycerol kinase catalytic core. Sciene has granted me permissio however, I also require the authors permission to reproduce any material pre 1994. Access to the thesis in digital format is free of charge and therefore, no profit will be realized from the work. Please respond in writing to confirm whether or not you will grant permission for the above-mentioned work to be included in my thesis, as well as any stipulations or royalties that you request.

Thank you very much for your consideration of this request.

Jason Richardson

From: James Hurley hurley@helix.nih.gov
Date: Thu, 15 Feb 2007 15:03:33 -0500

I am fine with it, but I believe AAAS/Science owns the copyright, not me.

Best,
Jim

Permissions Letter

Ref # 07-21221

DATE: 8 February 2007

TO: Jason Richardson
University of Manitoba
suite 108 115 Swindon Way
Winnipeg, Manitoba
Canada R3P 0W3

FROM: Emilie David, Rights and Permissions
RE: Your request for permission dated 02/07/07, submission id# 33388

Regarding your request, we are pleased to grant you non-exclusive, non-transferable permission to use the AAAS material identified below in your dissertation or thesis identified below, but limited to publication and distribution in the formats identified below, and provided that you meet the criteria below. Such permission is for one-time use and therefore does not include permission for future editions, revisions, additional printings, updates, ancillaries, customized forms, any electronic forms, braille editions, translations, or promotional pieces unless expressly permitted below. We must be contacted for permission each time such use is planned. This permission does not apply to figures / artwork that are credited to non-AAAS sources. This permission does not include the right to modify AAAS material.

Print the required copyright credit line on the first page that the material appears: "Reprinted (abstracted/excerpted) with permission from [FULL REFERENCE CITATION]. Copyright [YEAR] AAAS." Insert the appropriate information in place of the capitalized words.

Obtain permission from the author, who may grant or withhold permission, and who may or may not charge a fee if permission is granted. See original article for author's address. This requirement applies only to material published in Science during the years 1974 - 1994.

AAAS agrees that ProQuest/UMI may supply copies of the dissertation/thesis on demand.

AAAS must publish the full paper prior to use of any text.

AAAS does not supply photos or artwork. AAAS material must not be used in a derogatory manner. Use of the AAAS material must not imply any endorsement by the American Association for the Advancement of Science. This permission is not valid for the use of the AAAS and/or SCIENCE logos.

Permission is valid for use of the following AAAS material only:
Fig 3A from Hurley, J. H., SCIENCE 259:673-677 (1993)
In the following work only:
Literature review section of the thesis / dissertation titled CHARACTERIZATION OF L-RHAMNOSE CATABOLISM IN RHIZOBIUM LEGUMINOSARUM published by University of Manitoba in 2007
In the formats checked below only:
[x] Print
[x] Microform
[x] World Wide Web

Thank you for writing. If you have any questions please call me at (202) 326-7074 or write to me via fax at (202) 682-0816. For international calls, +1 is the country code for the United States.

Material requested:
From "Structure of the regulatory complex of Escherichia coli IIIGlc with glycerol kinase"; Hurley, J. H., Faber, H. R., Worthylake, D., Meadow, N. D., Roseman, S., Pettigrew, D. W., and Remington, S. J.; 1993; Vol. 259; 5095; pp. 673-677.
Figure of schematic diagram of glycerol and ADP binding in glycerol kinase catalytic core.

Figure 1.10. Glycerol and ADP binding to glycerol kinase within binding pocket. Reprinted (abstracted/excerpted) with permission from Science, volume 259(5059), pp. 673-677, Hurley, J. H., Faber, H. R., Worthylake, D., Meadow, N. D., Roseman, S., Pettigrew, D. W. and Remington, S. J., Structure of the regulatory complex of *Escherichia coli* IIIGlc with glycerol kinase, Copyright 1993, AAAS. Reprint permission also granted by James Hurley.

A.1.8 Elsevier Publishing Ltd. from Biochemical and Biophysical Research Communications

We hereby grant you permission to reprint the aforementioned material at no charge in your thesis, in print and on the University of Manitoba web site subject to the following conditions:

1. If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies.

2. Suitable acknowledgment to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows: "Reprinted from Publication title, Vol number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier".

3. Your thesis may be submitted to your institution in either print or electronic form.

4. Reproduction of this material is confined to the purpose for which permission is hereby given.

5. This permission is granted for non-exclusive world English rights only. For other languages please reapply separately for each one required. Permission excludes use in an electronic form. Should you have a specific electronic project in mind please reapply for permission.

6. This includes permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission.

Yours sincerely,

Steph Smith (Rights Assistant, Elsevier Ltd., The Boulevard Langford Lane, Kidlington, Oxford, OX5 1GB)

Material requested:
From "X-Ray structure of glycerol kinase complexed with an ATP analog implies a novel mechanism for the ATP-dependent glycerol phosphorylation by glycerol kinase"; Mao, C., Ozer, Z., Zhou, M. and Uckun, F. M.; 1999; Vol. 259; 3; pp. 640-644
Figure 3. Stereo view and Connolly surface of AMP-PNP in the GK/AMP-PNP/Glycerol structure.
Figure 5. The model of GK-induced phosphorylation of glycerol in the presence of ATP following an in-line nucleophillic attack mechanism.

Figure 1.11. A β-hairpin in glycerol kinase containing crucial for γ-phosphate contact of ATP. Reprinted from Biochemical and Biophysical Research Communications, volume 259(3), Mao, C., Ozer, Z., Zhou, M. and Uckun, F. M., X-Ray structure of glycerol kinase complexed with an ATP analog implies a novel mechanism for the ATP-dependent glycerol phosphorylation by glycerol kinase, pp. 640-644, Copyright 2004, with permission from Elsevier.

Figure 1.12. Model for catalytic mechanism of phosphotransferase in glycerol kinase. Reprinted from Biochemical and Biophysical Research Communications, volume 259(3), Mao, C., Ozer, Z., Zhou, M. and Uckun, F. M., X-Ray structure of glycerol kinase complexed with an ATP analog implies a novel mechanism for the ATP-dependent glycerol phosphorylation by glycerol kinase, 640-644, copyright 2004, with permission from Elsevier.

A.1.9 American Society for Microbiology from the Journal of Bacteriology

From: "Reels, Adar" areels@asmusa.org
Date: Wed, 21 Mar 2007 14:19:04 -0400

Here is your approved request.

Thank you

Adar L. Reels (Administrative Secretary/Program Assistant, Journals Department, American Society for Microbiology, 1752 N Street, NW Washington, DC 20036, (202) 942-9244 office (202) 942-9355 fax)

Material Requested:
Table 2.1. Bacterial strains and plasmids
Table 2.2. Complementation of L-rhamnose mutants
Table 2.3. Effect of L-rhamnose and glucose on β-galactosidase fusions to genes in the L-rhamnose transport operon
Table 2.4. Effect of RhaR on *rhaR* expression
Table 2.5. Growth of representative mutant *Rhizobium* strains on glycerol, rhamnose, and glycerol/rhamnose as the carbon sources
Figure 2.1. Rhamnose region of *R. leguminosarum* bv. *trifolii* strain Rlt100
Figure 2.2. RNA dot blot experiments
Figure 2.3. Rhamnose transport experiments
All text, figures and tables were reproduced from the Journal of Bacteriology, volume 186, Richardson, J. S., M. F. Hynes, and I. J. Oresnik, A genetic locus necessary for rhamnose uptake and catabolism in *Rhizobium leguminosarum* bv. *trifolii*, pp. 8433-8442,

02/05/2007

ATTN: Journals Dept., American Society for Microbiology, 1752 N Street, N.W., Washington, DC 20036-2904, USA; fax, (202) 942-9355.

To Whom It May Concern:

I am writing to request permission to include the following material from Journal of bacteriology, in a graduate thesis. The thesis, entitled Characterization of L-Rhamnose Catabolism in *Rhizobium leguminosarum* is part of my requirements as a student in the Ph. D. of Science department. My thesis/practicum will be microfiched by Library and Archives Canada (LAC). A microfiche copy of the theses/practicum will be available for loan from LAC, and will be reproduced and sold in various formats through LAC's agent, University Microfilms International (UMI). My thesis/practicum may also be posted electronically in the University of Manitoba digital repository (https://mspace.lib.umanitoba.ca/dspace/index.jsp), and be accessible to a worldwide audience from LAC's Theses Canada (www.collectionscanada.ca/thesescanada).

Material requested:
A genetic locus necessary for rhamnose uptake and catabolism in *Rhizobium leguminosarum* bv. *trifolii*; Richardson, J. S., Hynes, M. F. and Oresnik, I. J.; 2004; Vol. 186; 24; pp. 8433-8442.

URL:
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15576793

Access to the thesis in digital format is free of charge and therefore, no profit will be realized from the work. Please respond in writing to confirm whether or not you will grant permission for the above-mentioned work to be included in my thesis, as well as any stipulations or royalties that you request.

Thank you very much for your consideration of this request.

Jason Richardson
Ph. D. of Science
University of Manitoba
Winnipeg, MB Canada
(2
Fax

PERMISSION GRANTED
CONTINGENT ON AUTHOR PERMISSION (which you must obtain)
AND APPROPRIATE CREDIT
American Society for Microbiology
Journals Department
Date 2-7-07

A.1.10. RhaU protein crystallization data

Table A.1.10 Data collection, phasing and structural refinement statistics for a RhaU SeMet derivative and a L-rhamnose-soaked crystal

	pk	rm	ip	sk
Data collection statistics				
Unit Cell parameters				
Space group	$P3_212$	$P3_212$	$P3_212$	$P3_212$
a (Å)	69.1	69.2	69.3	68.7
b (Å)	69.1	69.2	69.3	68.7
c (Å)	100.9	101.1	101.1	100.7
α, β, γ (deg.)	90 90 120	90 90 120	90 90 120	90 90 120
Wavelength	0.9797	0.9079	0.9798	0.9330
Resolution (Å)	30 - 1.6 (1.66 1.60)[a]	30 – 1.6 (1.66 1.60)	30 - 1.7 (1.76 1.70)	30 - 1.85 (2.10 2.05)
Unique reflections	35,848 (3,487)	35,978 (3,574)	29,176 (2,848)	23,317 (2169)
Completeness (%)	97.6 (96.1)	97.1 (97.0)	95.1 (93.6)	98.3 (92.5)
R_{sym} (%)[b]	7.6 (51.0)	9.7 (57.6)	6.8 (46.8)	9.1 (30.7)
$\langle I/\sigma I \rangle$	13.9 (3.1)	11.0(3.2)	7.6 (1.7)	8.0(3.7)
Redundancy	5.6 (4.9)	5.7 (5.6)	1.7 (1.7)	4.2 (2.8)
Model refinement statistics				
Resolution	20 - 1.6 (1.64 - 1.60)		29 - 2.00 (2.05 - 2.00)	
No. of reflections	33,997 (2,483)			17,475 (1,281)
Free reflections	1,786 (118)			942 (64)
R_{cryst} (%)[c]	15.2 (21.3)			14.4 (14.8)
R_{free} (%)[d]	18.0 (27.2)			19.1 (23.6)
No. residues	1428			216
No. waters	220			204
Average B-factor (Å^2)				
Protein	12.7			19.9
Water	27.4			33.0
All atoms	17.3			21.9
rms deviations				
Bond lengths (Å)	0.018			0.013
Bond angles (deg.)	1.69			1.43

[a] Values in parentheses correspond to the highest resolution shell

[b] $R_{sym} = E_{hkl}\, E_j |I_{hkl,j} - \langle I_{hkl} \rangle| \,/\, E_{hkl} \langle I_{hkl} \rangle$, where j extends to all the observed hkl symmetry related reflections

[c] $R_{cryst} = \Sigma ||F_{obs}| - |F_{calc}|| / \Sigma |F_{obs}|$.

[d] R_{free} is as for R_{cryst} but calculated for a test set comprising reflections not used in the refinement.

(Communicated from P. Loewen) (Richardson *et al.* 2007)

A.1.11 Figure of mass spectrometry analysis of RhaU

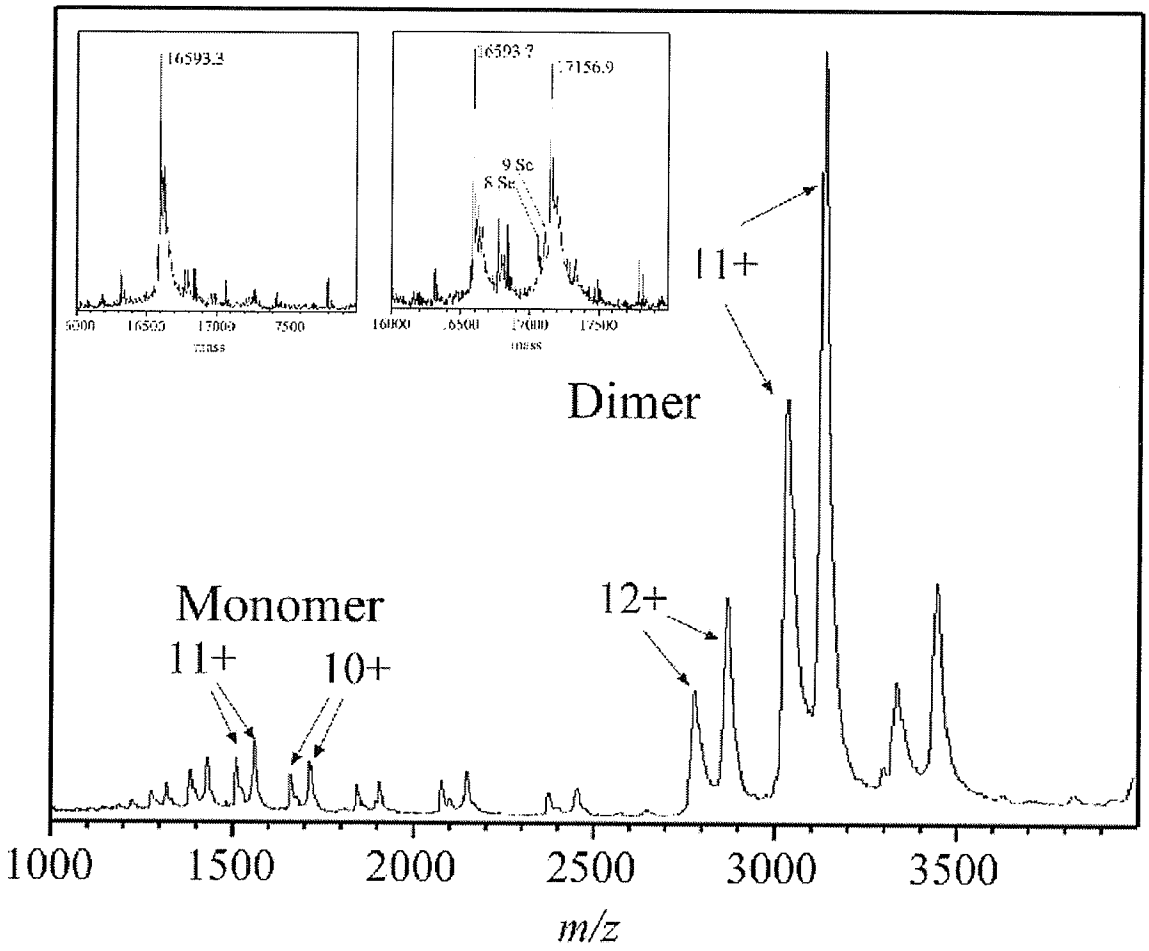


Figure A.1.11. Mass spectrometry analysis of RhaU monomer and dimer. (Communicated from P. Loewen) (Richardson *et al.* 2007)

A.1.12 Figure of RhaU dimer

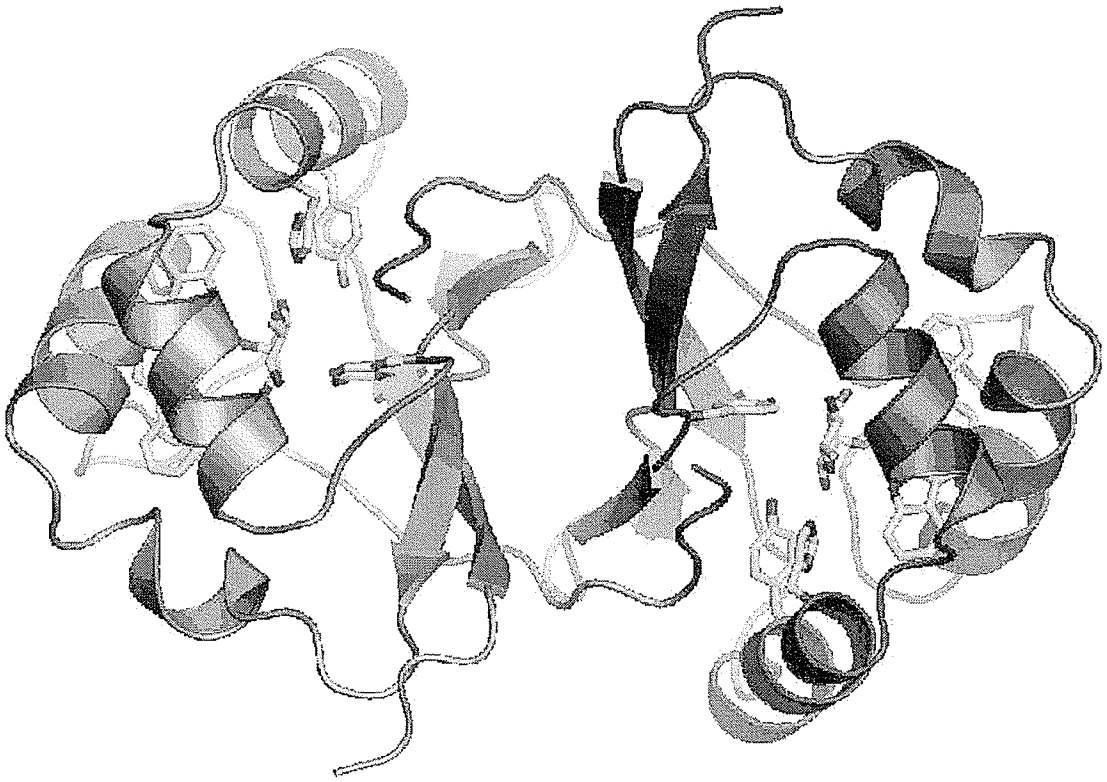


Figure A.1.12. RhaU dimer with catalytic core amino acid residues. Monomers are represented in blue and red β-barrels are drawn as ribbons and α-helices are tubular. (Communicated from P. Loewen) (Richardson *et al.* 2007)

A.1.13 Stereoisomer view of active site without and with L-rhamnose bound

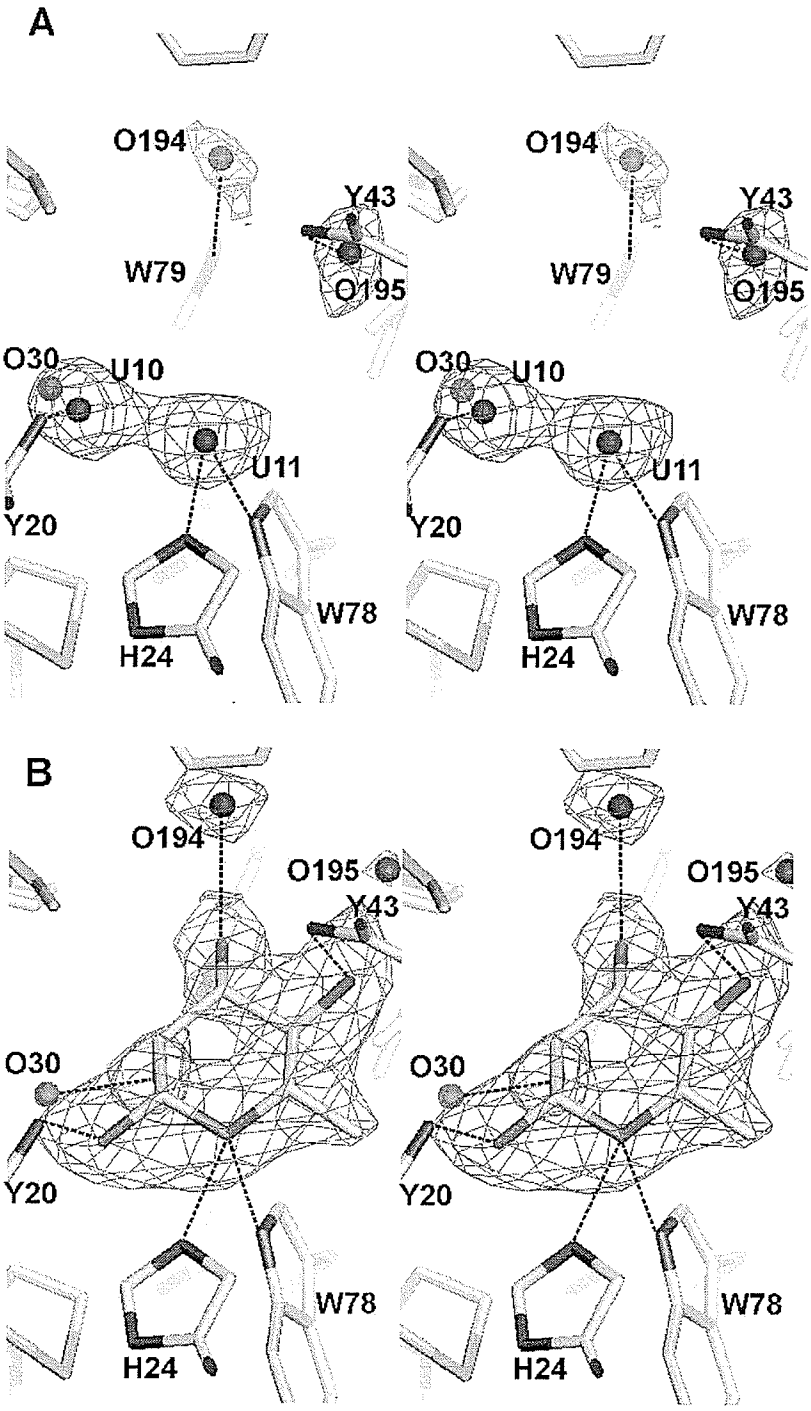


Figure A.1.13. Stereoisomer view of active site with amino acid residues and water molecules proposed to be involved in mutarotation catalysis. Dotted lines represent possible bonding arrangements. **A**. Water molecules in active site. **B**. β-L-rhamnose molecules in active site. (Communicated from P. Loewen) (Richardson *et al.* 2007).

A.1.14 Expression of rhaDI and rhaI fusions

Table A.1.14. β-Galactosidase assay expression of *rhaDI* and *rhaI* fusions

Strain	Relevant Characteristics	*rhaK* Genotype	Glucose[a b]	Rhamnose
Rlt100	wild-type	+	12	9
Rlt101	Δ *rha* region	-	9	7
Rlt114	*rhaK22*::Tn*5*	-	9	9
Rlt124	*rhaI31*::Tn*5*B20; genomic fusion	-	1071 (37)	1210 (167)
Rlt100(pMR79)[d]	*rhaI31*::Tn*5*B20; cosmid fusion	+; from genome	211 (21)	193 (1)
Rlt101(pMR79)[e]	*rhaI31*::Tn*5*B20; cosmid fusion	-	100 (3)	113 (13)
Rlt114(pMR79)[f]	*rhaK22*::Tn*5*/*rhaI31*::Tn*5*B20; cosmid fusion	-	209 (9)	171 (6)
Rlt124(pW3C1)[g]	*rhaI31*::Tn5B20; genomic fusion	+; from cosmid	92 (9)	1016 (64)
Rlt124(pMR110)[h]	*rhaI31*::Tn*5*B20; genomic fusion	+; from plasmid	1663 (117)	2412 (88)
Rlt124(pMR125)[i]	*rhaI31*::Tn*5*B20; genomic fusion	-; bp mutation on plasmid	1303 (84)	1235 (30)
Rlt105	*rhaD1*::Tn*5*B20; genomic fusion	-	140 (4)	259 (40)
Rlt100(pW3A1)[j]	*rhaD1*::Tn*5*B20; cosmid fusion	+; from genome	153 (10)	832 (27)
Rlt101(pW3A1)[k]	*rhaD1*::Tn*5*B20; cosmid fusion	-	1167 (34)	3956 (103)
Rlt114(pW3A1)[l]	*rhaK22*::Tn*5*/*rhaD1*::Tn*5*B20; cosmid fusion	-	2004 (138)	5539 (433)
Rlt105(pW3C1)[m]	*rhaD1*::Tn5B20; genomic fusion	+; from cosmid	15 (2)	125 (6)
Rlt105(pMR110)[n]	*rhaD1*::Tn*5*B20; genomic fusion	+; from plasmid	255 (4)	286 (40)
Rlt105(pMR125)[o]	*rhaD1*::Tn*5*B20; genomic fusion	-; bp mutation on plasmid	166 (4)	273 (38)

[a] The values given represent β-galactosidase activity expressed in Miller units following growth in defined media with either glucose/glycerol or rhamnose/glycerol as carbon sources.

[b] Values are the average of at least three independent replicates; values in parenthesis represent standard error. Rlt100, Rlt101, and Rlt114 do not contain transcriptional fusions; values are representative of one data set and represent background readings in the β-galactosidase activity assay.

[d, e, f, g, h, i, j, k, l, m, n, o] Are referred to in the strain list as Rlt222, Rlt229, Rlt238, Rlt230, Rlt247, Rlt224, Rlt251, Rlt252, Rlt225, Rlt231, Rlt248, and Rlt239 respectively.

A.1.15 Transport of L-rhamnose and RhaK activity in rhaDI transcript

Table A.1.15. Transport of L-rhamnose and RhaK activity in *rhaDI* transcript

Strain	Relevant Genotype	Glucose [a]	Rhamnose [b]	Transport (nmol/gFW/min) [cd]	Activity (μmol/min/mg)[a]
Rlt100	wild-type	+	+	28 ± 4	136.8 ± 29.6
Rlt146	*rhaK52*	+	-	< 0.05	n.d. [f]
Rlt200	*rhaK52; rhaK*$^{+}$	+	+	31 ± 1	32.9 ± 5.8
Rlt105	*rhaD1*	+	-	19 [e]	33.0 ± 10.4
Rlt124	*rhaI31*	+	-	174 ± 10	101.0 ± 40.6

[a b] Strains were tested for their ability to grow on minimal medium (VMM) supplemented with 15 mM rhamnose or glucose.
[c] Strains were grown as broth cultures in minimal medium (VMM) supplemented with 15 mM rhamnose/glycerol. initial transport rates were determined using tritiated rhamnose.
[d] ± values represent standard error.
[e] Represents an average of two replicates.
[f] n.d. not detected