Molecular genetic analysis of multiple *rpoH* and *groEL* genes in *Sinorhizobium meliloti*

by

Alycia Nacole Bittner

B.S., Biology, Juniata College, 2002

Submitted to the Graduate Faculty of

Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2008

UNIVERSITY OF PITTSBURGH

School of Arts and Sciences

This dissertation was presented

by

Alycia Nacole Bittner

It was defended on

April 7, 2008

and approved by

Karen M. Arndt, Associate Professor, Biological Sciences, University of Pittsburgh

Jeffrey L. Brodsky, Professor, Biological Sciences, University of Pittsburgh

Roger W. Hendrix, Professor, Biological Sciences, University of Pittsburgh

Joseph R. McCormick, Associate Professor, Biological Sciences, Duquesne University

Advisor: Valerie Oke, Assistant Professor, Biological Sciences, University of Pittsburgh

Molecular genetic analysis of multiple *rpoH* and *groEL* genes in *Sinorhizobium meliloti*

Alycia Bittner, PhD

University of Pittsburgh, 2008

The genomes of root-nodulating, nitrogen-fixing rhizobia that have been completely sequenced contain multiple copies of genes that encode the heat-shock transcription factor RpoH and the chaperone GroEL-GroES. *Sinorhizobium meliloti* maintains two *rpoH* genes, four *groESL* operons, and a single *groEL* gene. Mutations in some of these genes result in symbiotic defects: an *rpoH1* mutant cannot fix nitrogen in nodules, an *rpoH1 rpoH2* double mutant cannot form nodules, and a *groEL1* mutant cannot fix nitrogen in nodules. My work has sought to further characterize the roles of multiple *rpoH* and *groEL* genes during growth and symbiosis. In *E. coli*, *groESL* is the key target of RpoH. However, I showed that *S. meliloti rpoH* suppressor mutants do not overproduce GroEL, and overexpression of *groESL* does not bypass the *rpoH* mutant defects. In addition, RpoH1 controls expression of only *groEL5*, which is not required for symbiosis, and RpoH2 does not control expression of any of the *groEL* genes. Therefore, the requirements for RpoH1 and RpoH2 during symbiosis cannot be explained solely by loss of GroEL-GroES production, and there must be other crucial targets. To determine what genes are controlled by RpoH1 and RpoH2, I performed microarray experiments to compare global gene expression profiles between wild-type and *rpoH* mutant cells. Although the regulon of RpoH1 is incomplete, the results indicate that the RpoH1 and RpoH2 regulons at least partially overlap with each other and with the *E. coli* RpoH regulon. To uncover functional redundancies among the *groE* genes during growth and symbiosis, I constructed strains containing all possible

combinations of *groEL* mutations. Although a *groEL1 groEL2* double mutant could not be constructed, the *1⁻3⁻4⁻5⁻* and the *2⁻3⁻4⁻5⁻* quadruple mutants are viable, demonstrating that like other bacteria *S. meliloti* requires one *groEL* for growth. Analysis of the quadruple mutants during symbiosis indicates that only *groEL1* is necessary and sufficient for symbiosis. The *groEL1 groESL5* double mutant is temperature sensitive unlike either single mutant, suggesting overlapping roles during stress response. I conclude that *groESL1* encodes the housekeeping GroEL-GroES and that *groESL5* is specialized for stress response.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would first like to thank my advisor, Valerie Oke, for basically everything. I have probably spent more time with you than with anyone else over the past six years. Hopefully I've managed to avoid driving you crazy! I could not ask for a better advisor, and I have learned so much from you. I only hope that someday I can wield a red pen half as effectively. It's a given that I've learned much of what I know about science, teaching, writing and speaking from you. However, I have also learned so much more about balancing work with family and managing life in general. I will miss the chocolate, the ice cream, and you very much.

I also want to thank all of the members of my committee, Karen Arndt, Jeff Brodsky, Roger Hendrix, and Joe McCormick for their support. You have been very generous with your time, lab equipment, and advice, which I greatly appreciate. I must also thank our collaborators in Sharon Long's lab at Stanford University. Melanie Barnett and Carol Toman have worked tirelessly with me on the microarray experiments. I cannot thank you enough for your help. The Department of Biological Sciences has been a wonderful environment for a graduate student. I want to thank everyone for sharing their expertise and, particularly, their lab equipment! Also, the staff in the main and fiscal offices have always been helpful, particularly Cathy Barr. Thank you.

The Oke lab has had several talented undergraduates, rotation students and technicians in the past six years, who have made working in the lab much more entertaining. I particularly want to thank Amanda Foltz, Nick Jones, Elan Alford, and Anne Maly. I miss all of you!

Also, I must thank all of my friends and family for their support. To Marcie Warner, Maggie Braun, and Kathryn Sheldon, thank you for all of the awesome times and random adventures. I won't forget any of them. The graduate students in this department have been the greatest. I don't want to leave anyone out, so I won't be listing you here. Just know that I appreciate everything, that I will miss nearly all of you, and that you all rock! I also include all of my friends from the department who aren't graduate students in that statement. I would also like to thank my best friends outside of the department, Gina Taylor and Eric Bowers, and my family, particularly Dave and Adrian Bittner (my dad and my brother), for patiently allowing me to vent. In retrospect, things were never that bad!

I also want to thank Pitt Basketball and Football, the Pens, and the Steelers for providing welcome distractions. Finally, I thank Chuck Norris and Angela Lansbury for keeping me company late into the night while I was writing my dissertation.

1.0 INTRODUCTION

All organisms require nitrogen in order to synthesize biomolecules, like DNA and protein. Although atmospheric nitrogen is abundant, it must be converted, or "fixed," to a usable form, like ammonia, prior to assimilation by living organisms. Fixed nitrogen can be obtained through abiotic processes, such as lightning, which does not generate sufficient quantities to support all life on earth, and industrial production of ammonia, which is energetically costly. Another source of fixed nitrogen is biological: some bacteria and archaea are capable of catalyzing the conversion of dinitrogen to ammonia (309, 322). Plants of the legume family can utilize this biological source by entering into symbiotic relationships with nitrogen-fixing bacteria in which they exchange nutrients for fixed nitrogen. These symbiotic relationships are highly specific, and both partners have evolved mechanisms to maintain the interaction.

I am interested in the molecular requirements for bacteria to engage effectively in symbioses with host plants. In the following sections, I first introduce the *Rhizobium*-legume symbiosis, with emphasis on the model symbiotic bacterium *Sinorhizobium meliloti* and its interaction with host plants. I then describe the functions of the RpoH and GroEL proteins, which have been implicated in symbiosis. *S. meliloti*, like many other root-nodulating bacteria, maintains multiple copies of genes that encode the alternative sigma factor RpoH and the chaperonin GroEL, and mutations in some of these genes are associated with defects in symbiosis. Both RpoH and GroEL are involved in the cellular response to environmental stress,

which suggests a link between stress resistance and symbiosis. Finally, I discuss how multiple copies of these stress response genes might function in *S. meliloti* and why they might be important for symbiosis.

1.1 THE *RHIZOBIUM*-LEGUME SYMBIOSIS

Sinorhizobium meliloti and other *Rhizobium* species are members of the α-proteobacteria that can survive in the soil as free-living bacteria or form beneficial symbioses with host leguminous plants, including several agriculturally important crop species such as alfalfa and soybean. Rhizobia and their specific plant partners identify each other through the exchange of molecular signals. The plant partner secretes many chemical compounds that can be sensed by specific bacterial partners. These plant-derived compounds stimulate the bacteria to produce and secrete chemical compounds called Nod factors. Nod factor is perceived by the plant partner and elicits the formation of specialized organs called nodules on the roots of the host plant. Leguminous plants can form one of two kinds of nodule: determinate and indeterminate. Indeterminate nodules have a persistent meristem in which plant cells continue to divide throughout nodule development, whereas the meristem in determinate nodules is transient. Bacteria infect nodules during development and enter plant cells where they differentiate into a specialized cell type, called a bacteroid, which is capable of fixing nitrogen. The symbiosis is beneficial in that the bacteria receive nutrients from the plant and the plant receives fixed nitrogen from the bacteria.

In this section I describe the *Rhizobium*-legume symbiosis, with special emphasis on the interaction between *S. meliloti* and its host plants. I then describe the environmental challenges

faced by bacteria during symbiosis. In particular, I discuss the evidence demonstrating that functional stress responses are important to bacterial survival within nodules.

1.1.1 Bacterial-plant recognition

To initiate symbiosis, the plant and bacteria engage in a molecular dialogue such that chemicals produced by one partner are recognized by, and elicit a response in, the other. In the first step, plant roots exude chemical compounds, including flavonoids and betaines, which are sensed by specific rhizobial partners (Fig. 1) (60, 96, 233-235). Different plants produce different suites of compounds, which are in turn recognized by different bacteria. These compounds activate NodD transcription factors in specific bacteria, and the NodD proteins in different rhizobial species likely respond to compounds produced by different plants, which in part determines host specificity (232).

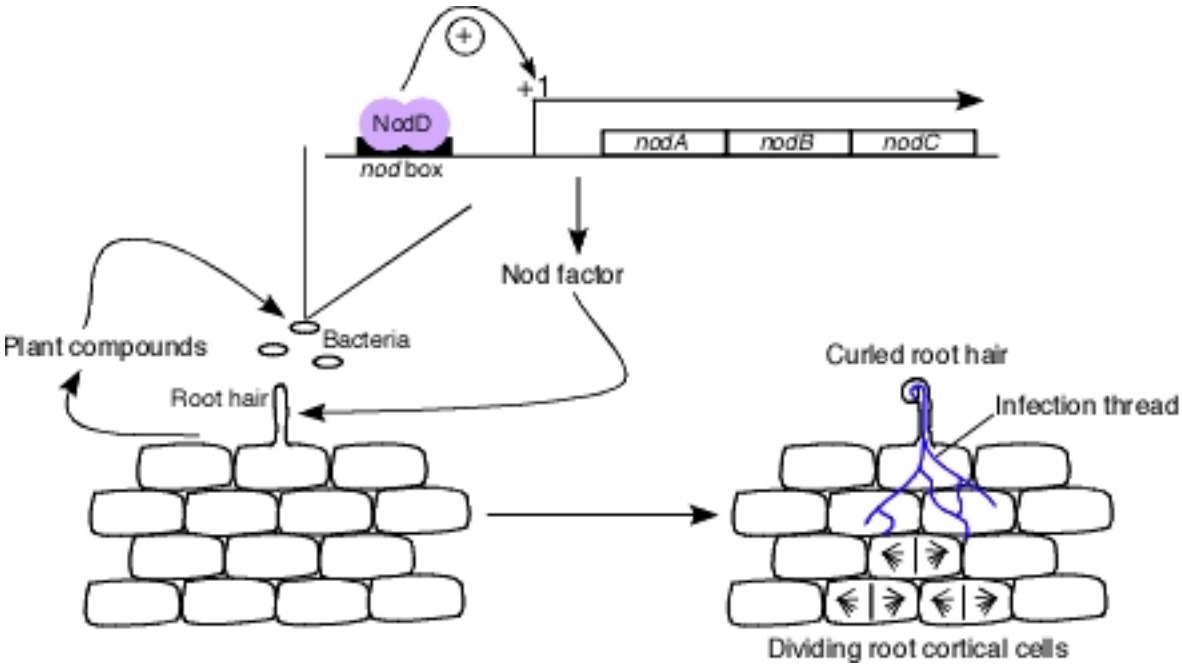

Figure 1. Exchange of plant and bacterial signals initiates symbiosis.

Plant roots secrete chemical compounds that are recognized by bacterial NodD transcription factors, which induce transcription of *nod* genes necessary for production of Nod factors. These compounds elicit several physiological responses in plant cells, including root hair curling and cortical cell division. An infection thread forms in the curled root hair to deliver bacteria to differentiated root cortical cells.

Some rhizobia have only one NodD protein, but others produce several (76, 113, 230). In *S. meliloti*, there are three NodD transcription factors (NodD1, NodD2, and NodD3) (137, 202), which respond to different plant compounds. NodD1, for example, is activated by the flavonoid luteolin (among others) (202), while NodD2 is activated by the betaines trigonelline and stachydrine (235). Because they recognize different inducers, NodD proteins might enable bacteria to form symbiotic relationships with different plants (130, 132). For example, *S. meliloti* can form symbiotic associations with legumes from the genera *Medicago*, *Melilotus*, and *Trigonella*, and these plants likely produce different chemical compounds. Multiple NodD proteins might permit *S. meliloti* to respond to several plant species (130).

The NodD transcription factors activate expression of bacterial *nod* genes by binding to promoters that contain a *nod*-box (Fig. 1) (54, 85, 87). Some of these genes encode enzymes required for the production of a lipochitooligosaccharide compound called Nod factor, which acts as a plant morphogen. The *nodABC* operon encodes the enzymes that synthesize the Nod factor core, which is composed of an *N*-acetyl-D-glucosamine backbone (12). Other genes encode enzymes that make specific modifications to the backbone, which determine the host plant range. In *S. meliloti*, these modifications include unsaturated fatty acids and acetyl and sulfate groups (8, 173). Changes in Nod factor composition can alter host plant specificity. For example, wild-type *S. meliloti* interacts with *Medicago sativa* (alfalfa) but not *Vicia sativa*. If *S. meliloti* Nod factors lack sulfate modifications, *M. sativa* can no longer respond to the bacteria, but sulfate-free Nod factors do stimulate the formation of nodules on *V. sativa* (173, 251). Each bacterial species can synthesize multiple Nod factors, which probably allows them to interact with different host plant species.

1.1.2 Nodule organogenesis and infection of plant cells

Once produced and secreted by the bacteria, Nod factor is perceived by multiple receptors on the surface of plant root hair cells (3, 176) and induces several physiological changes in plant root cells that initiate nodule organogenesis. Changes in root hair cells include increase of intracellular calcium, calcium spiking, and alterations to the cytoskeleton, which ultimately result in a morphological change such that the root hair curls (Fig. 1) (36, 37, 67, 68, 271, 285). Almost simultaneously, Nod factor stimulates differentiated cells within the root cortex to resume cell division (Fig. 1) (133, 285, 294, 306, 315). Some of these cells form the nodule meristem (Fig. 2), where plant cells divide throughout nodule development and which grows outward from the root surface. Plant cells located behind meristem tissue cease dividing and differentiate such that they increase genome copy number and cell size (38, 88). These differentiated cells are competent to receive invading bacteria.

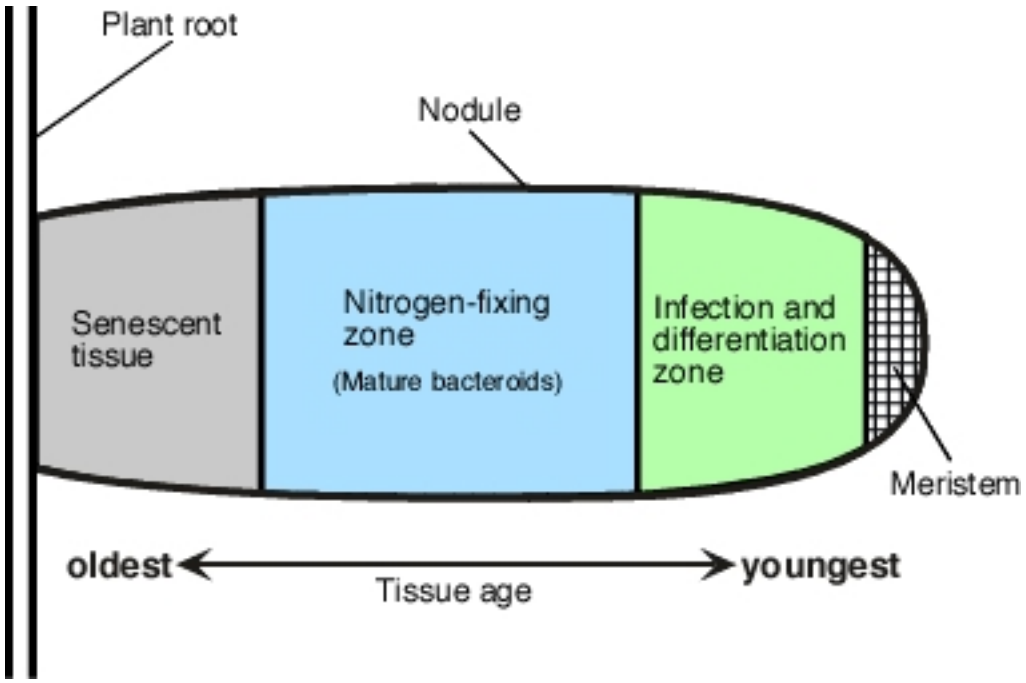

Figure 2. Zones of development in indeterminate nodules.

The persistent meristem grows outward from the surface of the root as plant cells continuously divide to generate new nodule tissue. The infection zone, where bacterial cells enter into plant cells, is located in older nodule tissue where plant cells have ceased to divide and have differentiated. Once inside the plant cells, the bacteria differentiate into bacteroids, and the nitrogen fixation zone contains mature nitrogen-fixing bacteroids. In the oldest nodule tissue, both plant cells and bacteroids undergo senescence.

At the plant root surface, root hair curling traps the bacteria near the root hair tip, where production of Nod factor (8, 301) and the bacterial exopolysaccharide (EPS) succinoglycan (82, 171, 219) induces formation of an infection thread that transports the bacteria to inner root tissues (Fig. 1). Although the molecular requirements are not well-defined, infection threads are believed to form by inversion of root hair cell tip growth (93). Infection thread growth involves deposition of new cell wall material along the sides (239), and it is likely mediated in part by the plant cell cytoskeleton (285). The infection thread must penetrate down through the root hair cell and beyond the layer of cells dividing in the meristem to deliver the bacteria to differentiated cortical cells (Fig. 2) (192).

The bacteria inside of infection threads replicate at the tip of the thread as it grows (92, 94) and are ultimately endocytosed by root cortical cells from the infection thread (26). Inside the plant cell, a symbiosome consists of a bacterium surrounded by a plant cell-derived membrane called the peribacteroid membrane. This membrane becomes specialized during symbiosis as its lipid and protein content are altered (249). Symbiosomes, including the bacteria and the peribacteroid membrane, divide once or twice in indeterminate nodules before the bacteria differentiate (249).

1.1.3 Bacteroid differentiation and nitrogen fixation

During bacteroid development in indeterminate nodules, the bacteria cease cell division, undergo genomic endoreduplication, and dramatically increase cell size, presumably to increase the cellular metabolic rate to support nitrogen fixation (98, 192, 297). The plant controls the bacteroid differentiation program, but how the plant achieves this is not understood. However, close contact between the bacterial and plant membranes appears to be important for bacteroid

development (299). Recent studies have speculated that a family of nodule-specific cysteine-rich peptides might be involved in signaling to or mediating development of bacteria within nodules, but this has not been experimentally tested (191, 192).

The bacterial genes required for bacteroid differentiation have not been fully elucidated. Phenotypic analyses of mutant bacteria (220) and analysis of bacterial gene expression within nodules (225, 231) have been used to determine the genes involved. However, many, if not all, of the genes identified appear to promote bacterial adaptation to conditions within the nodule and are not necessarily specific to bacteroid development. For example, the bacteria require expression from genes that encode proteins for transport of carbon sources from the plant (70) or that encode proteins necessary for resistance to environmental stress (225). Recent global gene expression and protein analyses have identified hypothetical gene transcripts or protein products that are enriched within bacteroids in nodules (4, 16, 18, 58, 59, 215). Whether these novel factors participate specifically in bacteroid development remains to be determined.

Free-living rhizobia are not capable of fixing atmospheric nitrogen. However, once bacteria have differentiated into bacteroids within plant cells (Fig. 2) and are located in a microaerobic environment (275), they can express the genes necessary for nitrogen fixation. The microaerobic environment plays a key role in symbiosis (250); the bacteria are obligate aerobes that require oxygen for respiration, but oxygen damages the nitrogenase complex. Plant cells create a microaerobic environment by producing leghemoglobin, which is a protein that binds oxygen with high affinity and buffers the oxygen concentration around bacteroids (228).

The regulatory cascade that controls expression of nitrogen fixation genes is controlled by low oxygen tension within the bacteroid, which is sensed by a two-component regulatory system (FixL/FixJ) (48). FixJ is the master regulator that induces expression of genes encoding

the regulatory factors NifA (which activates transcription from promoters recognized by the RpoN sigma factor) and FixK that in turn activate expression of the *nif* and *fix* genes (17, 280). The proteins encoded by the *nif* genes comprise the nitrogenase complex (9, 57), while many of the proteins encoded by the *fix* genes perform respiration under microaerobic conditions (45, 65, 257).

Eventually, as nodule tissue ages, the bacteroids and plant cells in the oldest sections of the nodule will undergo senescence (Fig. 2). The cause of senescence is not entirely clear, but reactive oxygen species, perhaps generated through respiration and nitrogen fixation, are clearly present in senescing tissues and may play a role in the process (1, 255).

1.1.4 Bacteria experience environmental stress during symbiosis

Although conditions faced by bacteria during symbiosis are incompletely defined, there is much evidence to suggest that bacteria are exposed to "stressful" conditions. For instance, bacterial mutants that are sensitive to certain environmental challenges have been shown to be impaired during symbiosis, which suggests that establishing a symbiotic relationship with a host plant requires resistance or proper adaptation to particular environmental conditions.

What are the sources of stress during symbiosis? Although the *S. meliloti*-legume symbiosis is a mutualistic relationship where both partners benefit, there is evidence that rhizobial infection triggers plant defense responses (259), indicating a possible link between pathogenesis and symbiosis. When plants are exposed to pathogenic bacteria, they initiate the hypersensitive response, which is characterized by generation of reactive oxygen species (ROS), production of phenolic compounds and defense proteins, cell wall fortification, and programmed cell death (203). From the plant perspective, it would be detrimental to allow potential pathogens

instead of beneficial rhizobia to enter. Although *S. meliloti* is able to suppress the hypersensitive response, there is evidence that bacteria must be able to resist an initial challenge from the plant. In addition, there are other potential sources of stress within nodules, which are unrelated to plant defense responses and include generation of ROS and acid stress. In this section I discuss the current knowledge of the link between bacterial stress responses and symbiosis.

1.1.4.1 Oxidative stress

An oxidative burst involving the generation of ROS, including superoxide (O_2^-) and hydrogen peroxide (H_2O_2), is among the early responses to pathogen infection in plants (6). *S. meliloti* suppresses the hypersensitive response in plants in part through the production of exopolysaccharides like succinoglycan (82, 219), lipopolysaccharides (33), and even Nod factor, which represses production of H_2O_2 during the earliest stages of infection (269). However, oxidative bursts have been detected during infection, where O_2^- is detected within the developing infection thread (241, 259). Generation of O_2^- is linked to the production of Nod factor (241), which demonstrates that regulation of ROS production during infection is complex.

The initial oxidative burst related to plant defense response is not the only source of ROS within nodules. ROS are generated by bacteroids through the processes of respiration and nitrogen fixation and in plant cells by the oxidation of leghemoglobin (185). Therefore, it is not surprising that ROS have been detected within infected plant cells up to 6 weeks post-infection (255, 259) and in senescing nodule tissue (1, 255), indicating that bacterial exposure to ROS might be prolonged. It is not clear how extensive or how damaging this exposure might be, but *S. meliloti* maintains multiple genes that encode oxidative stress resistance proteins, and mutations in some of these genes are linked to symbiotic defects.

The *S. meliloti* genome contains two genes that encode superoxide dismutase (97), which converts O_2^- to O_2 and H_2O_2. An *S. meliloti sodB* mutant is not defective for symbiosis (50), although the gene is expressed during infection (4). In a closely related strain, *S. meliloti* Rm5000, deletion of the *sodA* gene results in a strain that is only moderately impaired in oxidative stress response under free-living conditions (258). During symbiosis, however, the mutant cells nodulate poorly and rapidly senesce when released into plant cells.

S. meliloti also maintains three genes that encode catalase (97), which scavenges H_2O_2. KatA and KatC are monofunctional catalases, and KatB is a bifunctional catalase-peroxidase (7). In free-living bacteria, the three genes are regulated differentially such that *katA* is induced by H_2O_2 under control of the OxyR regulator (148, 272) and by low phosphate conditions (320), *katB* is constitutively expressed (149), and *katC* is induced by heat, osmotic and ethanol stresses (272). Indicating roles in symbiosis, *katA* is expressed in bacteroids, while *katB* and *katC*, but not *katA*, are expressed within infection threads (149). *kat* gene single mutants are not impaired during symbiosis (135, 149, 272). However, a *katA katC* double mutant is defective in nitrogen fixation (272), and a *katB katC* double mutant nodulates poorly and senesces upon entry into plant cells (149). This suggests that encoded catalases perform overlapping functions during symbiosis. Functional redundancy almost certainly maximizes the likelihood that a particular bacterium can resist oxidative stress and successfully invade nodules.

In a recent study, Davies and Walker (51) performed a genetic screen to identify *S. meliloti* mutants that were both sensitive to oxidative stress and defective in symbiosis. One set of mutants failed to produce succinoglycan, indicating that EPS production might form a diffusion barrier against H_2O_2. In another rhizobial species, *Azorhizobium caulinodans*, increased EPS production reduced the amount of H_2O_2 that entered bacterial cells (46), which

supports the barrier hypothesis. Another gene, *sitA*, encodes the periplasmic binding protein of a manganese ABC transporter, but its oxidative stress and symbiotic defects likely result from the requirement for manganese by other enzymes during the response (50). Other genes identified in the oxidative stress screen include genes that encode enzymes involved in metabolism, protein biosynthesis and cytochrome *c* biogenesis (51). The metabolic enzymes might allow the cell to deal with oxidative stress by increasing the production of NADPH for reduction of ROS (21, 145, 154). The authors also confirmed that mutation of the gene encoding the OxyR master regulator is sensitive to oxidative stress but proficient in symbiosis (51, 148). This study demonstrated that resistance to oxidative stress requires more cellular pathways than previously thought and also confirmed that only a subset of oxidative stress defense systems is required for symbiosis (51).

1.1.4.2 pH and osmotic stresses

Less is known about pH and osmotic stress within nodules than oxidative stress, but it is clear that some bacterial mutants that cannot adapt to changes in both are impaired in symbiosis. In another process related to plant defense responses, Nod factors can cause alkalinization of root hair cell cytoplasm and increase in the extracellular potassium concentration around the root hair (77, 78). A mutation in the *pha* gene cluster in *S. meliloti* disrupts genes that encode a potassium efflux system (238). This mutant cannot adapt to an alkaline pH, and the bacteria are unable to proceed through the infection thread, suggesting that the pH within infection threads is alkaline. Other mutants with osmolarity defects or sensitivity to alkaline pH include *nvd* mutants that cannot synthesize cyclic β-glucan (63, 64) and *exoD* mutants that are defective in EPS production (242). Both kinds of mutant display defects in nodule infection.

In addition to alkaline pH within the infection thread, bacteria may also be subjected to acidic conditions within nodules. It has been suggested that the peribacteroid membrane derives from the lytic vacuole, which has an acidic pH (190, 254). Some of the proteins localized to the symbiosome have an acidic pH optimum, which also suggests that bacteroids are located within an acidic compartment (190). At least three regulatory circuits are involved in the response of *S. meliloti* to low pH (110). Acid tolerance requires a two-component sensor regulator system, but a mutation in the transcriptional regulator *actR* does not affect symbiosis (51, 287). Absence of a phenotype does not necessarily mean that resistance to acid is not important for symbiosis; other regulatory circuits could be involved. This could be compared to a mutation in the *oxyR* gene, which encodes a global regulator of oxidative stress response (51, 148). This mutant also behaves normally during symbiosis, but the evidence in the section above clearly indicates that a subset of oxidative stress response genes is important for successful symbiotic interaction.

In a different species, *Rhizobium tropici*, the *lpiA* gene is part of an operon that is up-regulated in response to acid shock (298). A strain containing a mutation in *lpiA* is able to form a successful symbiosis with host plants. However, when plants are inoculated with cultures containing a mixture of wild-type and *lpiA* mutant cells, the mutant strain displays a seven-fold decrease in relative nodulation competitiveness compared to the wild type. In contrast, a *Sinorhizobium medicae lpiA* mutant is both normal for symbiosis and nodulation competitiveness (243). A gross defect in symbiosis was not seen in the *S. meliloti lpiA* mutant (110), but a subtle phenotype could have been missed since this strain was not compared to a wild-type strain in a competition experiment.

1.1.4.3 Other stress responses

Bacteria defective in other stress responses can be compromised for symbiosis. For example, the stringent response is stimulated in bacteria in response to nutrient starvation and is controlled by the effector molecule ppGpp (147). An *S. meliloti relA* mutant, which cannot produce ppGpp, cannot form nodules on plants and overproduces EPS (308). Suppressor mutant strains that can form nodules also have decreased succinoglycan synthesis, suggesting that the two phenotypes could be connected.

Bacteria that are sensitive to multiple stress conditions also demonstrate defects in symbiosis. In *S. meliloti*, disruption of a gene (SMc01113) that encodes a highly conserved protein of unknown function results in a strain that cannot infect plant cells (49). This strain is sensitive to many environmental stresses including oxidative stress and chemical agents that alter DNA metabolism, protein synthesis, cell envelope integrity, and peptidoglycan synthesis, indicating a general failure in adaptation to stress. In *R. tropici*, a mutation in a gene encoding glutathione synthetase results in a strain that is impaired in response to several stresses, including organic acid, osmotic, and oxidative stresses (247). While the mutant can form effective nodules on plants, it cannot compete efficiently with a wild-type strain during nodulation.

S. meliloti maintains multiple copies of genes that encode proteins which are implicated in stress response: six putative extracytoplasmic function (ECF) sigma factors (RpoE), two heat shock sigma factors (RpoH), and five GroEL chaperonins (97). This implies that response to environmental stress is highly regulated and crucial to the survival of this organism, perhaps in its symbiotic interactions. Only one of the six ECF sigma factors has been studied. RpoE2 is activated by various stresses, including heat, salt and stationary phase, and controls expression of other stress response genes such as *katC* and *rpoH2* (261). Free-living and symbiotic phenotypes

were not identified in the *rpoE2* mutant under any conditions tested, although mutations in some of its target genes are connected to symbiotic defects (149, 227, 272). Given the number of ECF sigma factors in *S. meliloti*, it is possible that there is functional overlap.

The activities of the alternative sigma factor RpoH and the chaperonin GroEL are connected to the heat shock response. RpoH regulates transcription of genes in response to heat stress, and the *groESL* operon is typically one of its target genes. The GroEL protein is a molecular chaperone that assists protein folding under heat shock and non-stressed conditions. Mutations in *rpoH1*, *rpoH2* and *groEL1* are connected to symbiotic defects (223, 226, 227). In the following sections I discuss the functions of RpoH and GroEL and the current knowledge of their roles during symbiosis.

1.2 THE HEAT SHOCK SIGMA FACTOR RPOH

As single-celled organisms, bacteria must rapidly adapt to changing environmental conditions in order to survive. One way in which bacteria respond to changes in the environment is to alter gene transcription. Bacterial genes are transcribed by RNA polymerase, which in *E. coli* is composed of an $\alpha_2\beta\beta'\omega$ core complex that is responsible for RNA synthesis (30, 31). The core complex associates with a sigma (σ) factor to form the holoenzyme complex. The sigma factor recognizes specific DNA sequences in the promoter region of a gene and is therefore required for polymerase to associate with its targets (131). Sigma factors are grouped into two families: the σ^{70} family and the σ^{54} (RpoN) family, which share little sequence identity (193, 260). σ^{70} family members recognize the -10 and -35 sequences of promoters (131). In contrast, σ^{54} proteins recognize the -11 and -26 sequences of promoters and require an activator protein (194).

Most bacteria maintain genes that encode several σ^{70}-family transcription factors. Typically, they produce one major, or housekeeping, sigma factor (RpoD or σ^{70} in *E. coli*) that directs transcription of the majority of the genes required for cell survival under normal growth conditions (126). However, bacteria also maintain a set of alternative sigma factors (from the σ^{70} family) that bind to consensus sequences that differ from the consensus sequence recognized by the housekeeping sigma factor. Alternative sigma factors allow bacteria to respond to changes in the environment by directing transcription of genes required for specialized functions. In response to nutrient starvation, for example, *Bacillus subtilis* enters a developmental program that requires an alternative sigma factor cascade to direct transcription of genes that encode proteins necessary for spore formation (179). Other alternative sigma factors include RpoH (125), which directs transcription of genes in response to heat stress, and the extracytoplasmic function (ECF) sigma factors, also called RpoE (71), which respond to extracellular stresses.

In this section I describe first the function of the alternative sigma factor RpoH in the model bacterium *E. coli*, with emphasis on its regulation and the regulon it controls, and then the RpoH regulons that have been defined for other bacteria with a single *rpoH* gene. Lastly, I discuss the function of RpoH in the α-proteobacteria, particularly the current knowledge of the functions of multiple *rpoH* genes in these organisms.

1.2.1 RpoH in the model organism *Escherichia coli*

In *E. coli*, RpoH (σ^{32}) was identified as a positive regulator of the heat shock response (125, 169, 216, 314). The heat shock response is conserved in eukaryotes, bacteria and archaea (13, 180), and it induces the transcription of genes that encode heat shock proteins (HSPs), whose function

is to prevent stress-induced cellular damage. These proteins also play an important role in the cell under non-stressed conditions (29, 324). The classical HSPs are either molecular chaperones that assist the folding of nascent proteins, prevent aggregation of misfolded proteins, and refold misfolded proteins or are proteases that degrade aberrant proteins. RpoH up-regulates expression from heat shock promoters in response to heat and other stresses.

It is clear that the function of RpoH is important under non-stress conditions in *E. coli*, because the *rpoH* mutant is unable to grow above 20°C (324). The requirement for RpoH during growth is tied to expression of genes that encode chaperones, in particular the *groESL* operon that encodes the subunits of the GroEL-GroES molecular chaperone machine, since expression of *groESL* from a multicopy plasmid restores growth up to 40°C in the mutant strain (167). Restoration of growth up to 42°C requires expression of *dnaK*, which encodes the DnaK (Hsp70) chaperone, in addition to expression of *groESL*. Because constitutively high activity of RpoH is costly, production and activity of RpoH are controlled at multiple levels, including transcription, translation, protein stability and activity.

1.2.1.1 Transcriptional, translational and posttranslational regulation

Under non-stress conditions, fewer than 50 molecules of RpoH are present in each cell (278). In response to heat shock, levels of RpoH rise rapidly and transiently, and the level and kinetics of RpoH induction correspond well to the activity of heat shock promoters during heat stress. Given the importance of RpoH levels for growth under normal conditions and for survival in response to stress (324), it is not surprising that expression of the *rpoH* gene and the levels/activity of the protein are exquisitely regulated.

The *rpoH* gene is transcribed from at least five promoters (P1, P3, P4, P5 and P6), three of which are regulated by the housekeeping sigma factor RpoD (72, 204). The extracytoplasmic stress response (ECF) sigma factor RpoE (σ^{24}) regulates expression from P3 (71, 304), and RpoN (σ^{54}) regulates expression from P6 (150). DnaA negatively regulates transcription of *rpoH* from P3 and P4 (195, 305), while the cAMP-CRP/CytR nucleoprotein complex negatively regulates transcription from promoters P4 and P5 (155, 204).

It has been suggested that increase in RpoH levels in response to heat shock is controlled largely at the level of translation. Early studies showed that expression of an *rpoH-lacZ* translational reporter fusion is induced by heat shock, while expression of a transcriptional reporter fusion is not (206). Additionally, translation of the reporter fusion is induced by heat shock even in the absence of RNA synthesis. A combination of *in vivo* and *in vitro* studies has demonstrated that translational repression is mediated by RNA secondary structure that forms at the 5' end of the *rpoH* mRNA transcript (199-201). This structure prevents binding of the ribosome to the Shine-Dalgarno sequence and thus inhibits protein translation. It is believed that melting of the secondary structure in response to an increase in temperature is sufficient to relieve translational repression (201).

The RpoH protein is rapidly degraded under non-stress conditions (37°C) such that its half-life is approximately one minute, and the protein is rapidly stabilized upon heat shock (278). Mutations in the genes that encode the DnaK-DnaJ-GrpE chaperone machine stabilize RpoH, suggesting that this complex might deliver RpoH to protein degradation machinery (277, 282, 284). Several proteases have been implicated in RpoH protein turnover, including the membrane-associated metalloprotease FtsH (134, 288) and the cytosolic proteases Lon, HslVU, and ClpP (158, 159, 307). It is postulated that the misfolded proteins that accumulate upon heat

shock are preferentially bound by chaperones and proteases, and RpoH is rapidly stabilized when the proteases begin to degrade other proteins (159). Because the genes encoding these chaperones and proteases are regulated by RpoH, their protein levels also increase upon RpoH stabilization. When an excess of these proteins is produced, they will again bind and degrade RpoH.

RpoH is also regulated at the level of activity, such that chaperone complexes bind and sequester RpoH to prevent it from associating with RNA polymerase and its target promoters. Heat shock regulation is similar in bacteria and eukaryotes in this respect, because chaperones have been implicated in the activity of heat shock transcription factor 1 (HSF1) (262, 270, 325). All RpoH proteins contain an "RpoH box," a conserved stretch of 9 amino acids with the sequence Q(R/K)(K/R)LFFNLR that is not conserved in other sigma factors (207, 209). DnaK physically interacts with RpoH (99, 100, 174, 175) through a binding site that contains the RpoH box (186, 205, 209). DnaJ and GrpE function together with DnaK to stimulate binding and release of RpoH, which affects activity, and presumably stability, of this transcription factor *in vivo* (99, 100, 174, 175). Interestingly, a mutation in the RpoH box alters binding of RpoH to RNA polymerase (153), which suggests that DnaK and RNA polymerase compete for the same or overlapping binding sites. The GroEL-GroES chaperone machine has recently been shown to negatively regulate activity of RpoH *in vivo* and to bind to RpoH *in vitro* (127). The authors suggest that GroEL regulates activity first by binding to and sequestering RpoH and then releasing the protein in an altered conformation that is unable to associate with RNA polymerase.

1.2.1.2 The *E. coli* RpoH regulon

Early studies identified RpoH gene targets by the analysis of individual proteins induced upon heat shock, as determined by two-dimensional gel electrophoresis (172) or of mRNA

species induced by heat shock, as determined by cDNA hybridization to membrane filters containing a genomic library (43, 44). Three recent studies have sought to completely define the RpoH regulon using global gene analyses with modern microarray technology. Two studies used transcriptional profiling to monitor increases in global gene expression over time when transcription of *rpoH* was induced under normal growth conditions (221, 323). The third study employed chromatin precipitation in combination with microarray analysis (ChIP-chip) to identify RpoH-dependent promoters under heat shock conditions (300). The advantage of the transcriptional profiling experiments was that it enabled identification of expression of genes within operons as being RpoH-dependent. Although ChIP-chip analysis would not necessarily identify genes within an operon because RpoH appears to stochastically, yet rapidly, disassociate from elongating RNA polymerase complexes (240), it eliminates the influence of indirect effects that alter transcription by directly determining the location of bound RpoH. Together, these studies have demonstrated that RpoH controls expression of more than just the classical heat shock genes and impinges upon many different cellular processes beyond protein homeostasis. The main findings of each study are discussed below.

Zhao *et al.* (323) first used global transcriptional profiling to monitor changes in expression when *rpoH* was moderately induced from a multicopy plasmid under normal growth conditions. Their strategy was to perturb RpoH levels without subjecting cells to heat stress because the heat shock stimulon is controlled by a complex network of global regulators that would also affect gene expression. They found that mRNA levels for RpoH-controlled genes increased by 5 minutes after induction and declined about 10 minutes after induction. Most known RpoH targets were identified in their analysis, including genes encoding the chaperone machines GroEL-GroES and DnaK-DnaJ-GrpE and the proteases Lon and ClpB, although

expression changes in some known targets, like *rpoD*, were below the arbitrary cutoff. Most of the genes encoded proteins that could be classified as chaperones, proteases, or proteins involved in adaptation to atypical conditions. They also identified a number of new targets not previously identified as RpoH-controlled genes. Bioinformatic analysis was used to define an RpoH consensus sequence: ggcTTGa-N_{12-20}-cCCCAT, where lowercase letters are less highly conserved. Finally, the authors compared the RpoH regulon to the heat shock stimulon to show that RpoH controls expression of only a small fraction of the genes that are induced in response to heat stress.

In the second study, Nonaka *et al.* (221) took a similar approach to the one above. After microarray analysis, they confirmed targets and identified transcriptional start sites using 5' RACE (rapid amplification of cDNA ends), which enabled inference of the -10 and -35 promoter sequences for each putative RpoH target. In total, they identified 49 validated RpoH-controlled promoters, and their results included 28 of the 32 promoters identified in the previous analysis. From the validated promoters they determined an RpoH consensus promoter: TTGAAA-N_{10-14}-CCCATAT. The targets identified in their analysis included many genes that encode the classical HSPs and other proteins involved in protein homeostasis. Additionally, they identified genes that encode proteins that maintain genomic integrity by protecting DNA and RNA or that function as effectors of transcription or translation. RpoH was shown to act as a master regulator by targeting transcription of several genes that encode transcription factors, some of which respond to extracellular conditions. Finally, nearly 25% of validated RpoH genes encoded proteins that participate in membrane homeostasis with roles in membrane protein quality control, increasing fatty acid synthesis, altering transport properties, maintaining disulfide bond formation and lipoprotein maturation, altering the composition of the lipid bilayer, and sensing

environmental conditions through the membrane. Further validating the connection of RpoH to membrane integrity, overproduction of 14 inner membrane proteins induced RpoH activity. The function of RpoH in membrane homeostasis explains why transcription of the *rpoH* gene (71, 304) and RpoH protein stability (134, 288) are in part regulated by proteins associated with the cytoplasmic membrane.

In the third study, Wade *et al.* (300) used ChIP-chip analysis to determine which promoters are bound by RpoH during heat shock. Their conservative value of 87 RpoH-dependent promoters (using a 1% false discovery rate) is much greater than the number of promoters identified in the transcriptional profiling analyses. There is significant but not complete overlap with the previous two studies: 29 of 32 from the first study (323) and 37 of 49 from the second study (221). To account for the great difference in total number of RpoH-regulated promoters, the authors suggest that RNA polymerase is recruited to certain promoters only after a transcriptional regulator is activated by heat shock (300). Therefore, this study may have missed RpoH-dependent promoters that require the function of other transcription factors induced by other conditions. A consensus binding sequence was also identified (CTTGAA-N_{13-15}-CCATAT), which is similar to the previous two consensus sequences. The most intriguing finding from this study is that 56% of RpoH-controlled promoters were also recognized by the housekeeping sigma factor RpoD. *In vitro* transcription reactions suggest that these sigma factors are capable of recognizing similar binding sites such that transcription occurred at the same start site. The authors also showed that there is some overlap between RpoE- and RpoD-controlled genes. Taken together, these results suggest that alternative sigma factors may have evolved to augment transcription from RpoD-controlled promoters in response to different growth conditions.

1.2.2 RpoH regulons in other bacteria with a single *rpoH* gene

RpoH regulons have been defined in two additional bacterial species, the pathogens *Vibrio cholerae* and *Neisseria gonorrhoeae*. The amino acid sequence of *Vibrio cholerae* RpoH is 70% identical to *E. coli* RpoH (273), so it is not surprising that the proteins regulate similar sets of stress response-related genes and recognize similar consensus binding sites (221, 300, 323). Slamti *et al.* (273) identified 49 genes that were down-regulated in *rpoH* mutant versus wild-type cells in response to heat shock, although only a third of these putative targets possessed a putative RpoH consensus sequence in the promoter region. The majority of genes identified in their analysis encodes proteins predicted to function in protein folding and degradation or encodes conserved hypothetical proteins. However, many of the targets identified are unique to the *V. cholerae* regulon, and some of these are predicted to play roles during pathogenesis. In fact, the *rpoH* mutant was attenuated for growth in the intestines of suckling mice, suggesting that the function of RpoH is important for infection.

The RpoH regulon was defined in *N. gonorrhoeae* by overexpressing the *rpoH* gene and monitoring global changes in gene expression (128). Twelve genes were identified as up-regulated upon *rpoH* overexpression, many of these encoding chaperones and proteases, and all but three of these genes were shown to be induced by heat shock. The RpoH-regulated genes represented less than half of the total number of heat-shock induced genes, demonstrating that other mechanisms must contribute to the heat shock response. Although the RpoH regulon in *N. gonorrhoeae* is much smaller than in *E. coli*, perhaps reflecting their relative genome sizes, there was significant overlap between the two. Interestingly, two genes induced by adherence to epithelial cells were shown to require RpoH for their expression but were not identified in the microarray experiment discussed above (61). These genes are believed to encode proteins that

function in pathogenesis, suggesting that RpoH has been co-opted by *Neisseria* to control some virulence factors. Interestingly, the *rpoH* mutant was shown to be defective in the ability to invade epithelial cells (62).

The above studies demonstrate that the stress response sigma factor RpoH is important for certain pathogens to invade and survive within host cells. It will be interesting to see whether the stress response function of RpoH is important for these effects or whether more of the organism-specific genes encode virulence factors.

1.2.3 RpoH in the α-proteobacteria: multiple *rpoH* genes

The heat shock stimulon is controlled by at least two mechanisms in the α-proteobacteria: the alternative sigma factor RpoH and HrcA/CIRCE. HrcA is a transcriptional repressor protein (264), which binds to an inverted repeat DNA element called CIRCE (controlling inverted repeat of chaperone expression) that is located between the promoter and the translational start site of certain heat shock genes (326). RpoH function in α-proteobacteria has been studied primarily in the plant pathogen *Agrobacterium tumefaciens* and the aquatic bacterium *Caulobacter crescentus*, which maintain single copies of the *rpoH* gene (209, 311).

The *A. tumefaciens rpoH* mutant is moderately temperature sensitive compared to the *E. coli* mutant (207); phenotypes of a *C. crescentus rpoH* mutant have not been reported. In contrast to *E. coli*, the transient increase of RpoH levels in response to heat shock appears to be regulated at the level of transcription rather than translation in the α-proteobacteria. The mRNA regions suggested to participate in translational control are not conserved in *A. tumefaciens* (209), and the *rpoH* genes in both *A. tumefaciens* and *C. crescentus* are autoregulated from

RpoH-dependent promoters (244, 312, 321). Like in *E. coli*, RpoH induces transcription of the *dnaKJ* and *groESL* operons, although HrcA/CIRCE plays a minor role in expression of *groESL* (207, 244, 266). Also unlike in *E. coli*, the RpoH protein is stable during normal growth conditions, and activation of RpoH, rather than increase in RpoH protein levels, is responsible for increased transcription at heat shock promoters (47, 208). DnaK appears to play a role in this activation, which parallels its interaction with RpoH in *E. coli*.

As indicated above, there are subtle differences in which RpoH levels and activity are regulated in the α-proteobacteria compared to the model bacterium *E. coli*. However, one of the most interesting differences is that several members of the α-proteobacteria maintain multiple copies of genes that encode RpoH sigma factors. Why might bacteria require multiple copies of *rpoH*? Possibly these genes encode factors that control similar regulons but are regulated differentially such that they are induced by different environmental conditions. Alternatively, the RpoH proteins could regulate different sets of genes and are perhaps specialized for different stress responses. It should be noted that these possibilities are not mutually exclusive. Below I review the characterization of multiple *rpoH* genes in the few organisms where they have been studied, beginning with the root-nodulating rhizobia. Where possible, I have noted differences in regulation, promoter specificity and gene targets.

1.2.3.1 *Rhizobium* species

Multiple *rpoH* genes have been found in all of the following root-nodulating members of the α-proteobacteria whose genomes have been completely sequenced: *Bradyrhizobium japonicum* (157, 213, 214), *Mesorhizobium loti* (156), *Rhizobium etli* (112), *Rhizobium leguminosarum* bv. *viceae* (319), and *S. meliloti* (97, 226, 227). All of these species have two

rpoH genes except *B. japonicum*, which has three. The numbering of the *B. japonicum rpoH* genes differs from the other rhizobia, and $rpoH_3$ encodes a protein similar to the RpoH2 proteins of other species, while $rpoH_2$ is more similar to other *rpoH1* genes (121). Characterization of multiple rhizobial *rpoH* genes has only been performed in *B. japonicum* and *S. meliloti* thus far.

B. japonicum, the nitrogen-fixing symbiont of soybean, maintains three *rpoH* genes that are differentially regulated (213, 214) and that may recognize different subsets of stress response promoters (212). $rpoH_1$ is located in a heat shock gene cluster that is transcribed as three bicistronic operons (214) and is regulated by σ^{70} and a unique DNA element called ROSE (for repression of heat shock promoters), which may act as a binding cite for a repressor protein (211, 213). $rpoH_1$ is induced by heat shock. $rpoH_2$ is regulated by two promoters: a σ^{70} promoter and a heat-inducible RpoE promoter (213). $rpoH_3$, is located in a cluster with genes that encode a two-component regulatory system and a putative efflux pump (166, 213). Taken together, this suggests that $rpoH_2$ functions under both normal and stress conditions and that $rpoH_1$ functions during heat stress. The conditions during which $rpoH_3$ functions are unknown.

All three genes are capable of initiating transcription of the *groESL* promoter in *E. coli* (213). However, they differ in their abilities to complement the temperature-sensitive growth phenotype of an *E. coli rpoH* mutant, with $rpoH_2$ being the most effective. Given that $RpoH_1$ and $RpoH_2$ display subtle differences in their affinities for representative *E. coli* and *B. japonicum* heat shock promoters *in vivo* and *in vitro* (212), the *B. japonicum* RpoHs likely have overlapping regulons but may also regulate a set of genes unique to each factor. $rpoH_1$, $rpoH_3$, and $rpoH_1$ $rpoH_3$ mutant strains were easily constructed, and the strains displayed essentially wild-type phenotypes under the growth and symbiotic conditions tested (213). An $rpoH_2$ mutant could not be constructed, suggesting that $rpoH_2$ is an essential gene in *B. japonicum*.

S. meliloti maintains two copies of *rpoH* in its genome, both of which are located on the chromosome (97, 226, 227). The predicted proteins are 42% identical to each other, and both contain a conserved RpoH box (Fig. 3) (226, 227). Despite limited identity to *E. coli* RpoH (36% for RpoH1 and 37% for RpoH2), both genes are able to complement the temperature-sensitive phenotype of an *rpoH* mutant, although RpoH1 directs transcription from the *E. coli groESL* promoter more effectively (227). Under free-living conditions, experiments using transcriptional reporter fusions demonstrated that *rpoH1* was strongly expressed during late log and exponential phases of growth in rich medium and that *rpoH2* was induced during late stationary phase in minimal medium (226). During symbiosis, only *rpoH1* was expressed within nodules. Taken together, these results suggest that *rpoH1* and *rpoH2* are differentially regulated. A recent study determined that *rpoH2* is a member of the RpoE2 regulon, which is activated by heat and salt stress and nutrient starvation (261). It is not known how *rpoH1* is regulated.

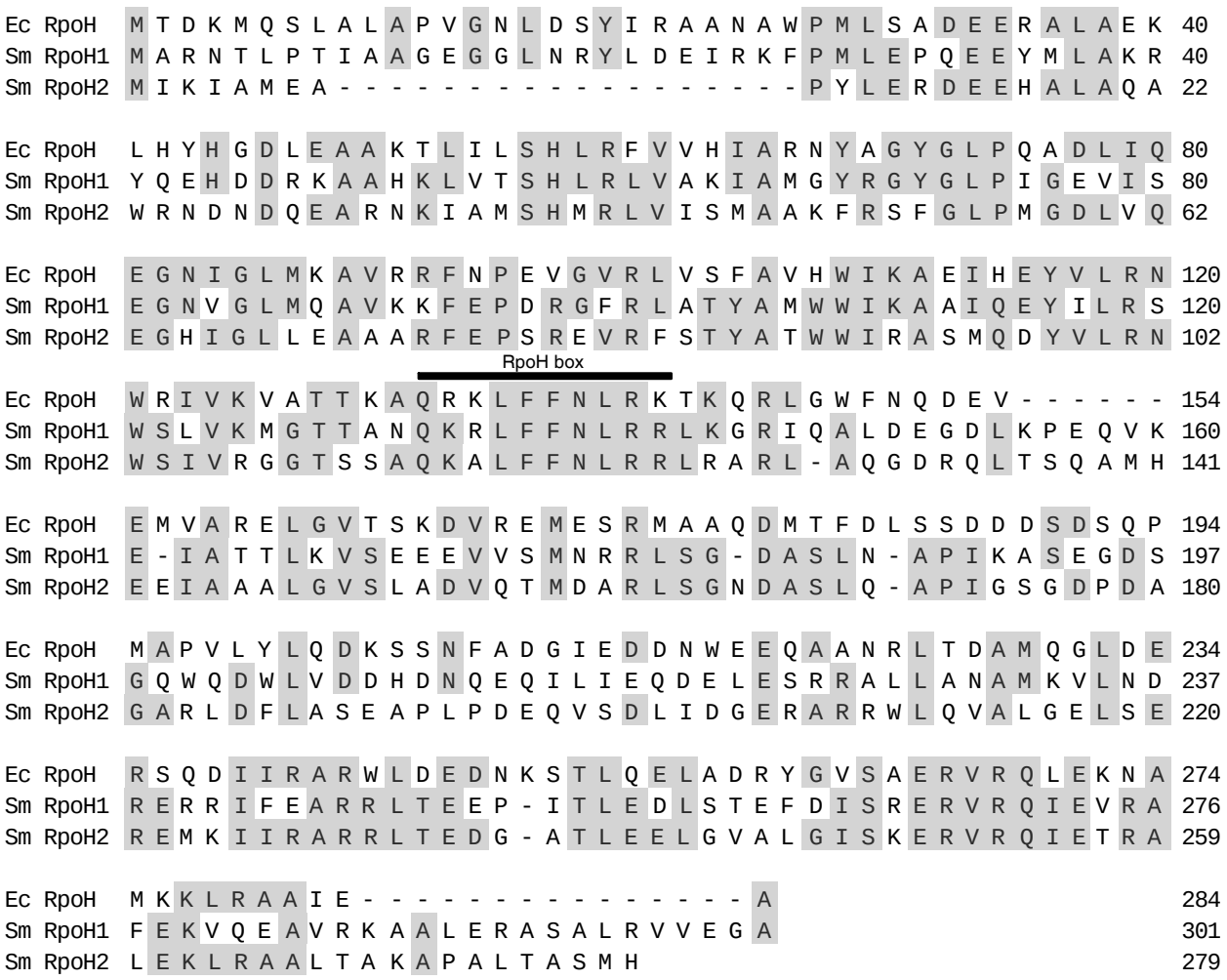

Figure 3. Alignment of *E. coli* and *S. meliloti* RpoH proteins.

The alignment was generated using the CLUSTAL V method (136) in the MegaAlign program from LASERGENE (DNASTAR, Inc.). Residues in gray match the consensus sequence. The RpoH box is indicated below a black line. Abbreviations are *E. coli* (Ec) and *S. meliloti* (Sm).

In contrast to *E. coli*, the *rpoH1* and *rpoH2* single mutants and the *rpoH1 rpoH2* double mutant were easily constructed at normal temperatures, although strains containing an *rpoH1* mutation display a slight growth defect (226, 227). During free-living growth, strains containing an *rpoH1* mutation are sensitive to various stress conditions including high temperature (226), acid pH (227), detergents, hydrophobic dye and ethanol, suggesting inability to mount a proper stress response and possible cell envelope defects. Synthesis of several, though not all, HSPs was decreased in the *rpoH1* mutant during heat shock, and synthesis of at least one additional HSP was decreased in the *rpoH1 rpoH2* double mutant (227). This suggests that RpoH1 and RpoH2 have at least one common target under the conditions tested. Another study showed that the RpoH1 regulon at least partially overlaps with the *E. coli* RpoH regulon because it controls expression of *clpB*, *groESL5*, and *lon*, but not *dnaK* and *clpA*, during heat shock (196). They used the *clpB*, *groESL5*, and *lon* promoters to determine a consensus promoter sequence for RpoH1 (CNCTTGAA-N_{17}-CCANAT). RpoH2 did not control the expression of any of these genes under the conditions tested.

During symbiosis, the *S. meliloti rpoH1* mutant forms nodules but is not able to fix nitrogen (226, 227). Electron micrographs show that *rpoH1* mutant cells senesce shortly after infecting plant cells, indicating that the phenotype results from failure to survive during symbiosis and is not a specific defect in nitrogen fixation (196). The *rpoH2* mutant forms effective nodules, but the *rpoH1 rpoH2* double mutant is unable to form nodules on plants, suggesting that the functions of RpoH1 and RpoH2 during symbiosis at least partially overlap (227).

1.2.3.2 Non-*Rhizobium* species

Two genes encoding RpoH sigma factors have been identified in the following non-root-nodulating α-proteobacteria: *Bartonella hensela* (2), *Bartonella quintana* (2), *Brucella melitensis* (56), *Rhodobacter sphaeroides* (121, 160, 181), *Rhodospirillum rubrum* (152), and *Silicibacter pomeroyi* (198). However, the possible functions of these *rpoH* genes have not been extensively studied outside of the root-nodulating bacteria, with the only exceptions being *rpoH* mutant analysis in *R. sphaeroides* and *B. melitensis*.

R. sphaeroides is an aquatic bacterium that is capable of photosynthesis and nitrogen fixation, among other metabolic activities. Both *R. sphaeroides rpoH* genes can complement an *E. coli rpoH* mutant (121, 160). An *R. sphaeroides* $rpoH_I$ mutant is sensitive to the toxic oxyanion tellurite but is not temperature sensitive, although the mutant exhibits slightly altered induction kinetics of several HSPs (160). Both RpoH proteins are capable of recognizing an overlapping but not identical subset of heat shock promoters *in vitro*, suggesting they might recognize an overlapping set of promoters *in vivo* (121). The *rpoH* genes might be expressed in response to different stresses, since $rpoH_{II}$ is regulated by RpoE but $rpoH_I$ is not (5).

B. melitensis is an intracellular mammalian pathogen, which, much like root-nodulating bacteria, is endocytosed by host cells where it likely persists within an acidic, membrane-bound compartment derived from the host cell membrane (237). Both *rpoH* genes have been mutated in *Brucella melitensis* 16M (55). An *rpoH2* mutant is sensitive to heat, cold, and oxidative stress. RpoH2 participates in the regulation of the type IV secretion system and the flagellum, both virulence factors, and the *rpoH2* mutant was impaired in all virulence models tested. The *rpoH1* mutant is impaired for survival in mice one month post-infection but behaves similarly to the wild type under all other conditions tested. Interestingly, this study suggests that RpoH2 has

been co-opted to regulate genes important for virulence that are not members of the model *E. coli* RpoH regulon. Whether both the stress response function of RpoH and the virulence factors it targets are important for infection remain to be seen.

There is no single unifying characteristic for the 11 species that maintain multiple *rpoH* genes. Several are root-nodulating symbionts (*Rhizobium* species), while others are free-living, nitrogen-fixing bacteria (*R. sphaeroides*, *R. rubrum*) or pathogens (*Bartonella* species, *B. melitensis*). *S. pomeroyi* is an aquatic bacterium that neither interacts with a host nor fixes nitrogen. In a phylogenetic analysis of RpoH proteins, all RpoHs from the α-proteobacteria cluster together (121). RpoH1-like proteins mostly cluster together, as do the RpoH2-like proteins (Fig. 4), suggesting that the *rpoH* gene was duplicated once in this lineage. Still, it is not clear whether a second RpoH was recruited once in the α-proteobacterial lineage to direct transcription of similar sets of genes in all of these bacteria or whether a second RpoH was recruited to direct transcription of different sets of genes in different bacteria. However, given that several of these bacterial species interact with eukaryotic hosts, an interesting possibility is that these heat shock sigma factors have been co-opted to transcribe genes essential for their complex lifestyles.

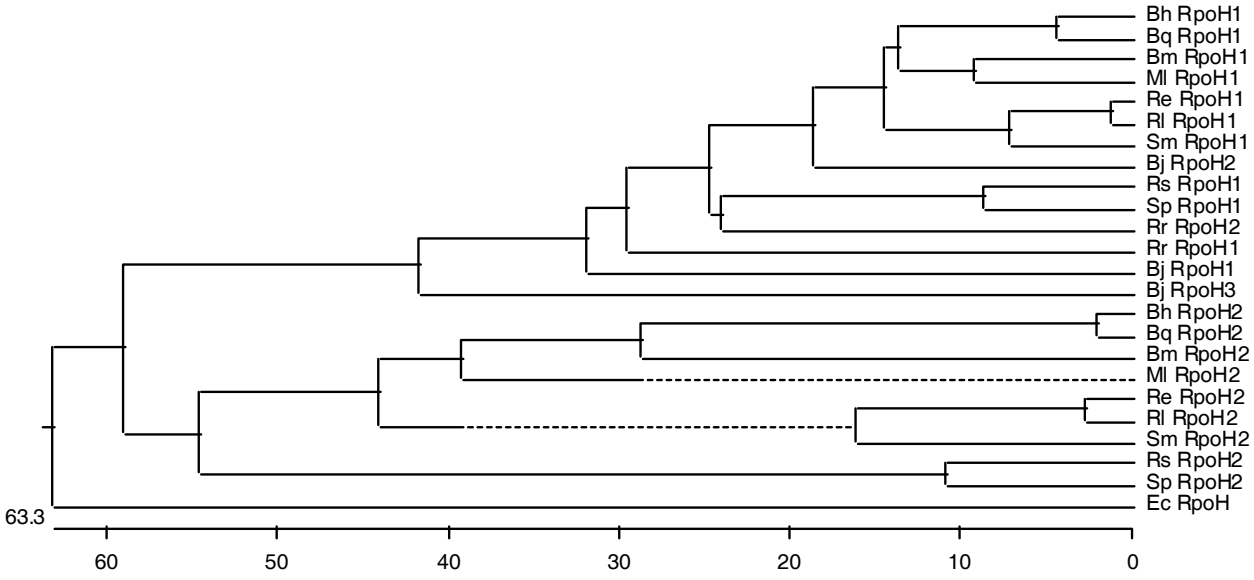

Figure 4. Phylogenetic tree showing relationships between multiple RpoH proteins in the α-proteobacteria.

E. coli RpoH was used as an outgroup. The sequence alignment was generated using the CLUSTAL V method (136) in the MegaAlign program from LASERGENE (DNASTAR, Inc.), and the phylogeny is rooted with the assumption of a biological clock. Abbreviations are *B. hensela* (Bh), *B. quintana* (Bq), *B. japonicum* (Bj), *B. melitensis* (Bm), *E. coli* (Ec), *M. loti* (Ml), *R. etli* (Re), *R. leguminosarum* (Rl), *R. sphaeroides* (Rs), *R. rubrum* (Rr), *S. pomeroyi* (Sp), and *S. meliloti* (Sm). The numbering of the three RpoH proteins in *B. japonicum* is inconsistent with the numbering of the two RpoH proteins in the other species, such that Bj RpoH2 is most similar to RpoH1 of other α-proteobacteria.

1.3 THE CHAPERONIN GROEL

A key target of the RpoH sigma factor in *E. coli* is the gene that encodes the chaperonin GroEL (167), which functions in protein folding. Protein folding is a complex problem for cells under both stressed and non-stressed conditions. Many proteins do not fold spontaneously upon synthesis, and proteins can become misfolded in response to environmental stress, like heat shock. The exposed hydrophobic patches on misfolded proteins are prone to aggregation, which is toxic to the cell. All organisms produce proteins, called molecular chaperones, which bind to the hydrophobic patches on misfolded proteins to prevent aggregation and can assist proteins in folding properly. The chaperonin GroEL with its co-chaperonin GroES comprise a barrel-shaped molecular chaperone machine. In the following section I discuss the cellular functions of the GroEL-GroES complex in the model organism *E. coli*. I then describe the regulation and function of multiple *groEL* genes in *Rhizobium* species.

1.3.1 GroEL in the model organism *Escherichia coli*

The *groE* locus was originally identified in genetic screens for bacteriophage λ-resistant mutants in *E. coli* (105, 106, 276). It was shown that the *groE* mutants failed to properly assemble bacteriophage λ (90, 104, 105, 283) and T4 (90, 104, 246, 281) capsids and T5 tails (105, 327), suggesting a link to macromolecular protein assembly. However, GroEL was subsequently shown to interact with monomeric proteins (23, 111) and therefore likely assists the proper folding of individual subunits of macromolecular complexes. The *groEL* gene was also

identified, independently, as being up-regulated in response to heat shock (125), where its protein product functions to prevent protein aggregation. The GroEL-GroES complex has now been shown to be involved in the folding reactions of many cytosolic proteins (73, 142, 164).

In *E. coli*, the *groEL* gene is located in an operon with the gene encoding its co-chaperonin, *groES*. The *groESL* operon is essential for growth at all temperatures in *E. coli* (75), underlining its importance in protein folding. In the following sections I describe the transcriptional regulation of the *groESL* operon and the folding functions of the GroEL-GroES complex.

1.3.1.1 Operon regulation

In *E. coli*, the *groESL* operon is regulated from two promoters. Under normal growth conditions and in response to temperature increases, the *groESL* operon is transcribed from an RpoH-dependent promoter. At low temperatures (below 20°C), it is transcribed from a promoter that is recognized by the housekeeping sigma factor, which presumably supplies a basal level of transcription (324). In Gram-positive and some Gram-negative organisms, expression of *groESL* is not regulated by RpoH. Rather, expression of *groESL* is negatively regulated by a DNA element called CIRCE (326), which binds the HrcA repressor protein under non-stress conditions (264).

Interestingly, the GroEL protein has been implicated in regulation of the *groESL* operon through its interactions with RpoH and HrcA. In *E. coli*, GroEL binds and sequesters RpoH, and it has been suggested that it may inactivate RpoH by altering its structure so that it cannot bind to RNA polymerase (127). Therefore, RpoH would not be able to direct transcription of its target genes, including *groESL*. In *Bacillus subtilis*, the HrcA repressor protein has been shown to be a GroEL substrate, and GroEL enhances the ability of HrcA to bind the CIRCE element and thus

repress expression of CIRCE-controlled genes (197). In this way, GroEL acts as a sensor for misfolded proteins: when the amount of unfolded proteins increases, levels of GroEL protein also increase because its interaction with unfolded proteins prevents it from repressing transcription of the *groESL* operon.

1.3.1.2 Protein folding activity of the GroEL-GroES chaperone machine

The *groES* and *groEL* genes encode 10 kDa and 60 kDa proteins, respectively. The GroES protein forms a seven-membered ring that acts as the cap on the GroEL-GroES complex (Fig. 5a) (144). The GroEL complex consists of two stacked seven-membered rings, which form a barrel-like structure (24, 25). GroEL subunits have three distinct domains (Fig. 5b). The apical domain is located at the top of the structure, is loosely structured, and contains the substrate and GroES binding sites. The non-polar amino acids on the apical surface (24, 25) bind to unfolded proteins with exposed hydrophobic patches (74, 79). The flexible binding domain might enable GroEL to bind to a wide range of substrates (39). The equatorial domain forms the bulk of the protein, mediates contacts between the two heptameric rings, and contains an ATPase active site (24, 25). The intermediate domain links the apical domain to the equatorial domain. The complete, two-ringed complex contains two cavities separated by the GroEL equatorial domains and their C-terminal tails, and each central cavity is ~45 Å in diameter. Therefore, the ability of GroEL to fold a particular substrate in part depends on the size of the substrate and whether it can enter the central cavity.

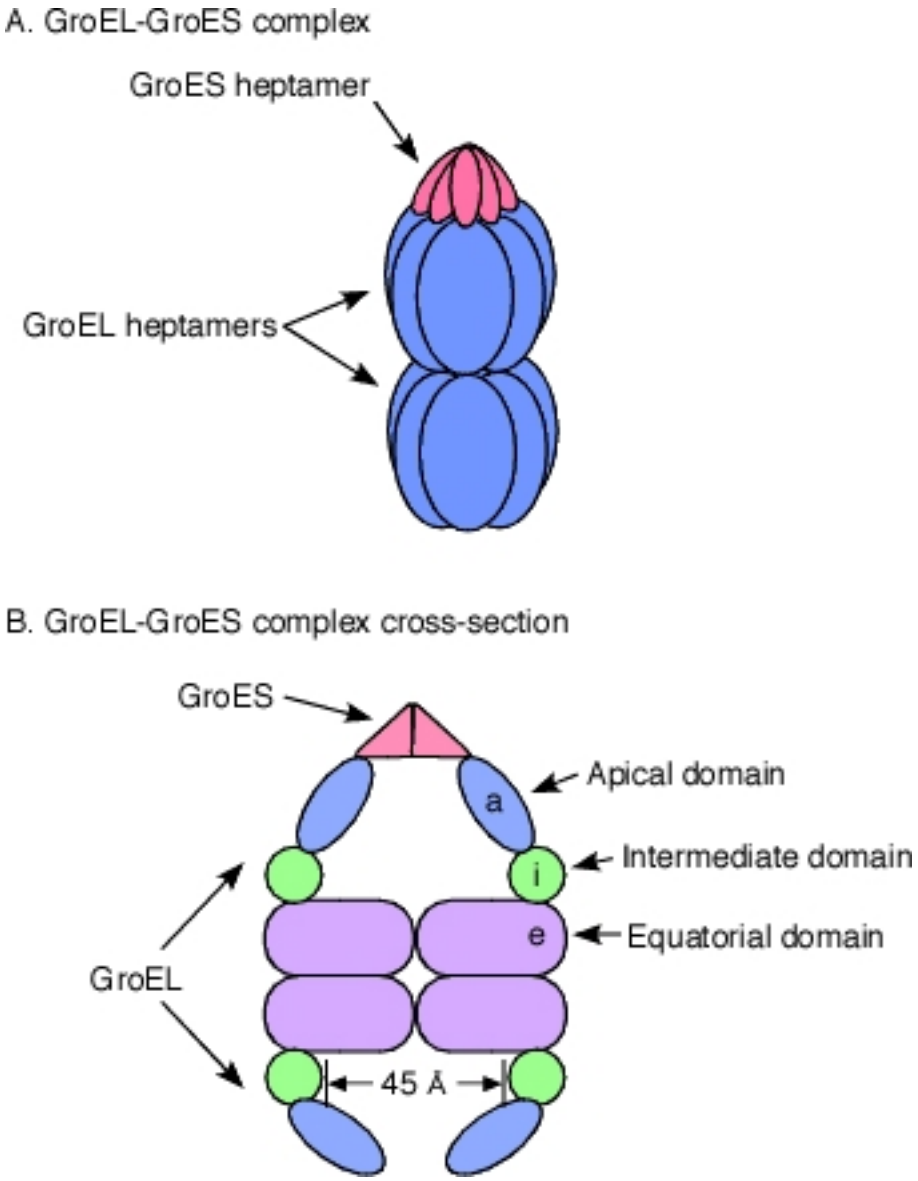

Figure 5. The GroEL-GroES molecular chaperone machine.

(A) In the GroEL-GroES complex, two heptameric rings of GroEL (blue) associate with each other and with a heptameric GroES (pink) cap. (B) The GroEL subunits are composed of three domains: apical (blue), intermediate (green), and equatorial (purple). The GroES and substrate binding sites are located in the apical domain, and the equatorial domain contains the ATPase active site and makes contacts with the second GroEL ring.

In the GroEL-GroES reaction cycle, the two GroEL rings work cooperatively such that assembly of the GroEL-substrate-GroES complex in one half is coupled to disassembly in the other half (Fig. 6). In a single reaction cycle, a non-native substrate will bind to hydrophobic residues on the GroEL apical domains (74, 79), followed by cooperative binding of ATP to the GroEL equatorial domains and GroES binding to the apical domains (22, 32, 120, 146), which strongly inhibits ATP-binding in the second GroEL ring (32, 317, 318). To finish the cycle, ATP is hydrolyzed, and the substrate and GroES cap are released as an unfolded substrate binds to the other side of the GroEL barrel. ATP-binding repositions the GroEL apical domains so that the GroES cap can bind, displacing the substrate to the center of the cavity (40, 253, 313). A second structural consequence of ATP-binding is that the amino acids lining the cavity are shifted so that hydrophilic residues are exposed on the surface.

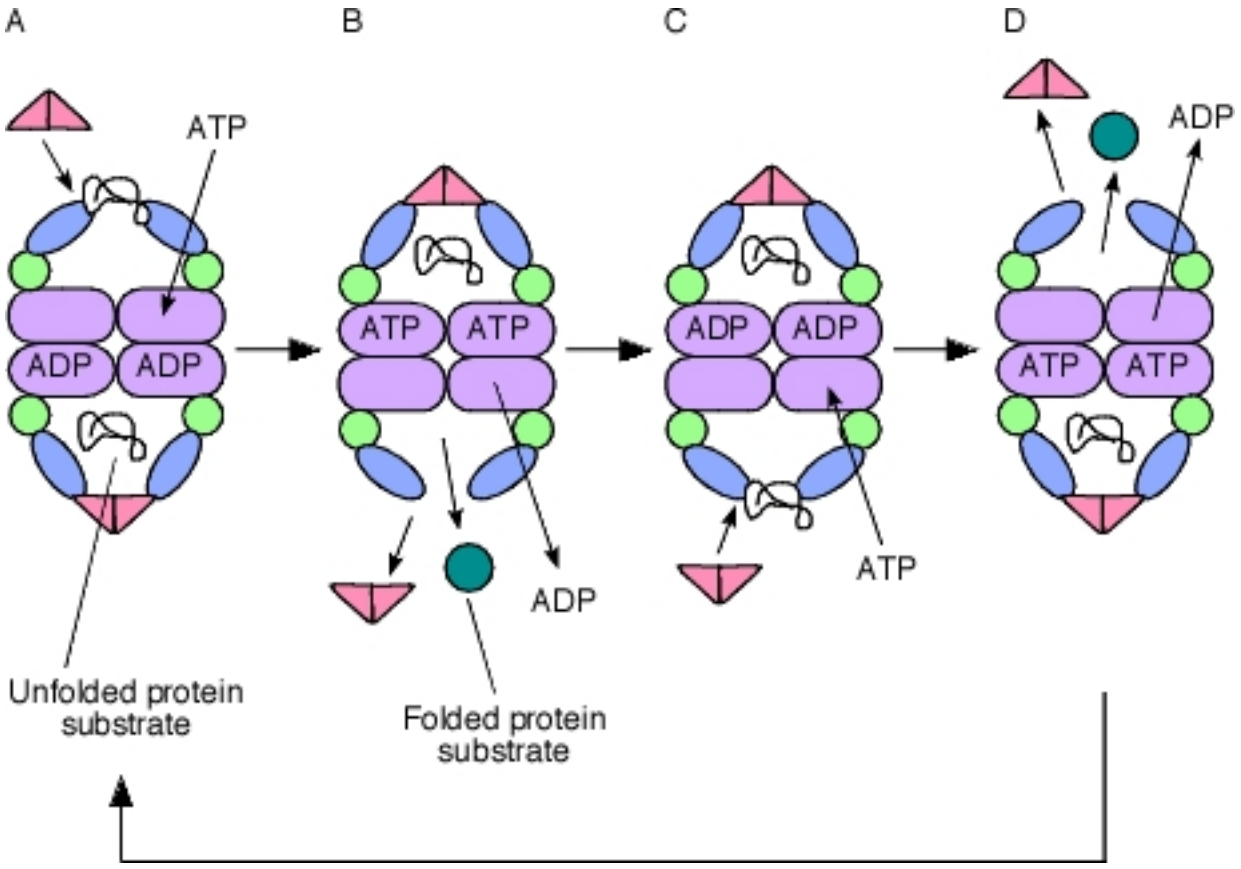

Figure 6. The GroEL-GroES reaction cycle.

(A) Once ATP has been hydrolyzed to ADP in one GroEL ring, the second GroEL ring is capable of binding an unfolded protein substrate. (B) ATP and GroES-binding to the second GroEL ring is coupled to GroES, substrate, and ADP-release from the first ring. (C) Upon ATP hydrolysis in the second GroEL ring, the first ring can bind an unfolded substrate. (D) ATP and GroES-binding to the first GroEL ring is coupled to GroES, substrate, and ADP-release from the second ring.

Currently, it is not clear whether the GroEL-GroES complex actively participates in protein folding or merely prevents protein aggregation by isolating a misfolded protein. In models of active participation, GroEL could modify protein substrates by unfolding trapped intermediates (177) or by accelerating the folding rate by lowering thermodynamic barriers (178). In passive models, the formation of protein aggregates primarily inhibits protein folding, and GroEL acts by preventing this aggregation (178). The active and passive models of GroEL-mediated protein folding are not mutually exclusive. For example, it would be possible for GroEL-GroES to actively assist certain misfolded proteins while simply preventing aggregation of others.

1.3.1.3 Substrates and specificity

Molecular chaperones in bacteria have long been considered promiscuous proteins because they bind and fold many different protein substrates. However, all chaperones do not bind to all nascent or misfolded proteins. In *E. coli*, the GroEL-GroES and DnaK-DnaJ-GrpE chaperone machines perform overlapping roles in preventing protein aggregation (119). However, only the GroEL-GroES complex is essential for growth (75), suggesting that this complex recognizes and/or folds substrates that DnaK-DnaJ-GrpE cannot.

Several studies have sought to define the substrates of GroEL-GroES in *E. coli in vivo* using proteomics approaches and have demonstrated that the complex associates with 10-15% of cytosolic proteins under normal conditions and ~30% under heat shock conditions (73, 140, 142, 164). Three classes of chaperonin substrates have been identified: class I proteins interact with GroEL-GroES complexes but fold independently, class II proteins display an intermediate dependence on GroEL-GroES for folding, and class III proteins absolutely require GroEL-GroES for folding (73, 164). Of approximately 250 GroEL-GroES-interacting proteins, about 85 are

class III substrates, and 13 of these are essential proteins, which may explain the requirement for GroEL-GroES during growth (164). The GroEL-GroES substrates are enriched for proteins that contain several αβ domains, which contain hydrophobic β sheets and could expose substantial hydrophobic patches during protein folding (141).

Because the GroEL-GroES complex is limited in the range of substrates it can productively bind and fold, some studies have sought to manipulate substrate specificity. In one study, researchers used directed evolution experiments to select for mutations in *groESL* that resulted in GroEL-GroES proteins better able to fold an exogenous substrate, green fluorescent protein (GFP) (303). The selected mutations altered the ATPase activity of the of the GroEL protein and increased the polarity of the GroEL-GroES complex cavity. The authors demonstrated a conflict between specificity and promiscuity, as the adapted complexes were less effective at folding endogenous substrates.

In a separate study, researchers attempted to engineer GroEL protein with altered substrate specificity by changing amino acids in the substrate-binding site (163). They showed that mutations in the apical domain altered both substrate and GroES binding. Therefore, the ability to modulate GroEL specificity through the apical domain is limited because the resulting proteins are less functional during folding when interaction with the GroES co-factor is disrupted.

1.3.1.4 Non-folding functions

Molecular chaperones like GroEL sometimes perform functions not specifically linked to protein folding, such as binding to a mature protein to modulate its activity. One of the best-characterized examples is the eukaryotic glucocorticoid receptor (GR), a transcription factor that regulates transcription of its target genes once it is activated by glucocorticoid-binding. Folding

of the nascent GR protein is assisted by the Hsp70 complex (274). After folding, the GR protein has low affinity for glucocorticoid and is transferred from Hsp70 to Hsp90 (41, 217). It has been suggested that GR binding to Hsp90 activates the receptor by opening its ligand-binding cleft (122). In another eukaryotic example, activity of the heat shock transcription factor HSF1 is negatively regulated by Hsp70 and Hsp90, which are thought to sequester the protein (262, 270, 325).

Non-folding functions of molecular chaperones have been less well characterized in bacteria. In *E. coli*, the DnaK-DnaJ-GrpE chaperone machine has been shown to negatively regulate the heat shock sigma factor RpoH by sequestering the protein (99, 100, 174, 175) and altering its stability (277, 282, 284). Because RpoH levels are stabilized in *dnaK*, *dnaJ*, or *grpE* mutant cells, DnaK-DnaJ-GrpE may also be involved in targeting the protein for degradation (277, 282, 284). Additionally, chaperones have been demonstrated to alter the properties of membranes or lipid bilayers through interactions with lipids rather than proteins (139). For example, HSP17 in *Synechocystis* PCC 6803 has been shown to stabilize the lipid phase of thylakoid membranes and thus participates in regulation of membrane fluidity (138, 290, 292).

GroEL has recently been shown to perform a non-folding function in *E. coli* by sequestering RpoH (127). In addition to briefly sequestering the protein, it is thought that GroEL might alter the conformation of RpoH so that it is unable to bind to RNA polymerase but is still capable of being degraded. GroEL has also been shown to interact with membrane proteins in what appear to be non-folding activities. For example, GroEL binds to and solubilizes bacteriorhodopsin and λ holin proteins to deliver them to lipid bilayers (52, 53). However, these studies have not demonstrated that GroEL delivers the proteins to cytoplasmic membranes *in vivo*. Intriguingly, a subset of the GroEL population has been shown to be associated with

membranes in *E. coli* (218), which suggests that GroEL may interact with membrane proteins, lipids, or both. One study demonstrated that *E. coli* GroEL can associate *in vitro* with model lipid membranes to stabilize the lipid bilayer, suggesting that GroEL might perform a similar function during heat stress *in vivo* (291).

Non-folding functions for GroEL have also been studied in bacterial pathogens, where it appears that GroEL functions as an intercellular signaling molecule. GroEL proteins from various pathogens have been shown to stimulate proinflammatory cytokine production by immune cells (89, 229, 245). Although it is not known how GroEL is secreted, it has been reported to be found on the cell surfaces of some bacteria (34, 91, 101, 102, 165, 265) and is suggested to play a role in bacterial adhesion to (91) or invasion of (102) host cells.

1.3.2 Multiple *groEL* genes in *Rhizobium* species

Like *E. coli*, many bacteria maintain a single essential *groESL* operon. However, nearly 20% of sequenced bacterial genomes contain multiple copies of the operon or additional *groEL* genes (114). Unlike multiple *rpoH* genes, which are restricted to several members of the α-proteobacteria, bacteria that maintain multiple *groEL* genes are from many bacterial subdivisions and include both Gram-positive and Gram-negative organisms. Much like *E. coli*, however, the species that maintain multiple copies of *groEL* seem to require at least one *groEL* gene for survival (170, 224, 252, 268). Why do bacteria maintain multiple copies of *groEL*? One possibility is that the genes are regulated differentially, but the protein products perform similar functions. Therefore, the bacteria could augment GroEL levels in response to different environmental conditions. Differential regulation has been observed in several bacterial species with multiple *groEL* genes (108, 161, 170). For example, three *groEL* genes in *Chlamydiae*

trachomatis are differentially expressed at different stages during the bacterial infection cycle (107).

A second possibility is that the encoded proteins perform different functions, either through recognizing and folding different protein substrates or performing novel non-folding functions. Bioinformatic analyses involving multiple GroEL proteins from many different species suggest that the proteins have functionally diverged because they contain differences in the domains of the protein that participate in substrate and GroES binding, ATP binding and hydrolysis, or subunit interactions (118, 187). The substrates or functions of GroEL-GroES complexes in bacteria with multiple *groEL* genes have not been fully determined for any species, although some specific functions have been identified. For example, the non-essential GroEL1 protein, but not the essential GroEL2 protein, in *Mycobacterium smegmatis* is involved in biofilm formation when it is required to associate with a component of the mycolic acid biosynthesis pathway (224). In an extreme example of a protein tailored to a particular substrate, bacteriophage T4 encodes a protein with little sequence similarity to GroES that can substitute for the host GroES to allow folding of its major capsid protein by the host GroEL (11, 295).

All of the root-nodulating rhizobia whose genomes have been completely sequenced maintain multiple copies of *groESL*, which suggests that the function of GroEL may be important for their complex lifestyle, particularly the symbiotic associations they form with host leguminous plants. The multiple *groEL* genes in these species encode similar proteins and likely arose from a combination of gene duplications, speciation, and horizontal gene transfer events (Fig. 7) (114, 118). In the following section, I review the current knowledge of multiple *groEL* genes in the three *Rhizobium* species (*B. japonicum*, *R. leguminosarum*, and *S. meliloti*) where they have been studied. Where possible, I discuss regulation, functional differences, and

connections to symbiosis. Multiple *groEL* copies are also found in *Rhizobium etli* (four *groESL* operons) (112) and *Mesorhizobium loti* (five *groESL* operons) (156), but the multiple *groEL* genes have not been studied in these bacteria.

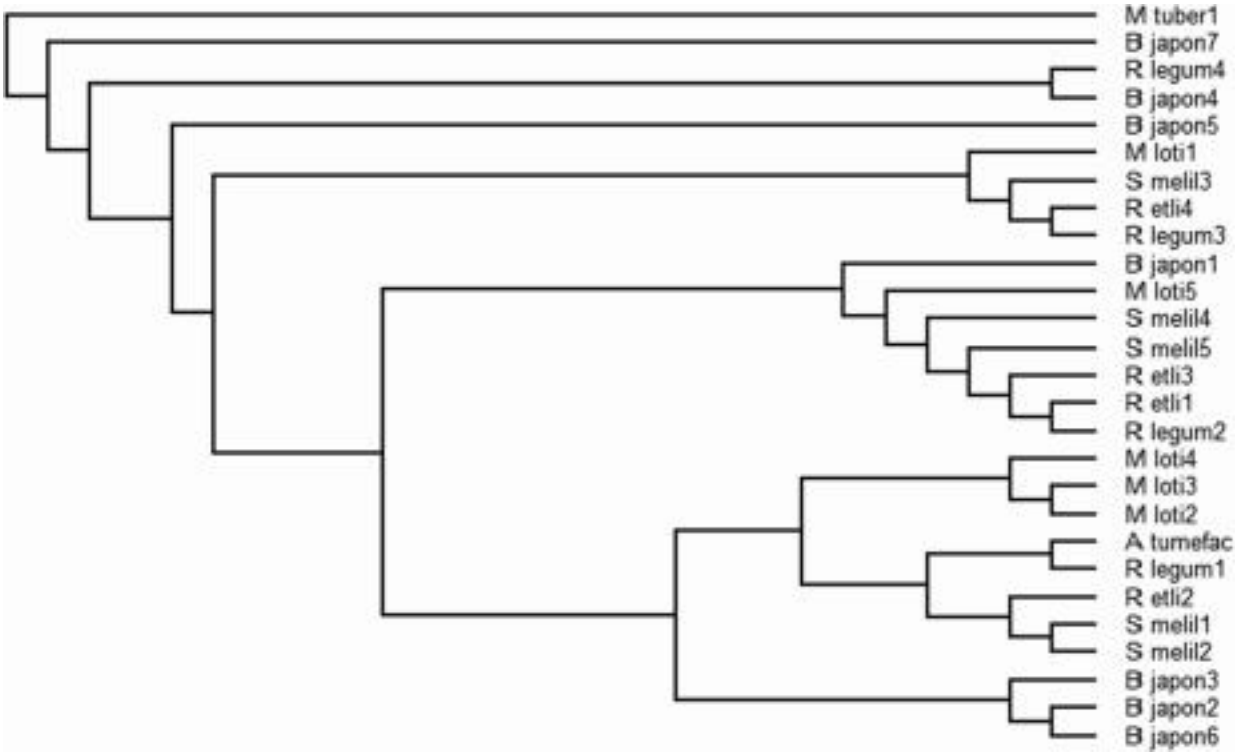

Figure 7. Cladogram showing relationships between rhizobial GroEL proteins.

Mycobacterium tuberculosis GroEL was used as an outgroup. The sequence alignment was generated by ClustalW, with analysis from seqboot and protpars in the phylip program package. From Gould, P. S. *et al.* (2007) *Cell Stress & Chaperones*. **12**: 123-31.

1.3.2.1 *Bradyrhizobium japonicum*

The genome of *B. japonicum* contains five *groESL* operons (83), and two single *groEL* genes (157). Some of these genes are differentially regulated. At least two operons are induced by heat shock (10, 83): $groESL_1$ is expressed from an RpoH-dependent promoter, and $groESL_4$ is negatively regulated by a CIRCE element. In contrast, $groESL_3$ is regulated with nitrogen fixation genes by the NifA regulatory protein from an RpoN (σ^{54})-dependent promoter (83).

The regulation of $groESL_3$ by NifA suggests that GroEL might be important for symbiosis. It was also shown that symbiotic bacteroids contained seven times the amount of GroEL protein found in free-living cells (42). However, when the *groEL* genes of the five *groESL* operons were mutated, none of the single mutants displayed a symbiotic phenotype (83), and only the $groEL_4$ mutant displayed a temperature sensitive growth phenotype. Several double mutant strains were created and tested for symbiotic phenotypes, and a $groEL_3$ $groEL_4$ mutant is unable to fix nitrogen during symbiosis (84). The effect of these mutations was to decrease the levels of nitrogenase proteins in bacteroids and anaerobically growing cells, which corresponds nicely to NifA regulation of $groESL_3$. The $groEL_3$ $groEL_4$ mutant was complemented by expression of all of the *groESL* operons, suggesting that each can function in place of $groEL_3$ and $groEL_4$ to allow production of nitrogenase. As further evidence of functional equivalence, all five *groESL* operons are able to partially complement the growth defect of an *E. coli groESL* mutant.

1.3.2.2 *Rhizobium leguminosarum*

R. leguminosarum bv. *viciae* was sequenced and shown to have three copies of the *groESL* operon and a single *groEL* gene (319). However, all of the GroEL (Cpn60) studies in

this organism have been performed in the *R. leguminosarum* strain A34, which has three *groESL* operons (302). Several studies have shown that these operons are differentially regulated. Expression of *cpn.1* (or *groEL1*) is heat inducible and negatively regulated by a CIRCE element (115, 252), which corresponds to an increase of the Cpn60.1 (or GroEL1) protein upon heat shock (302). Expression of *cpn.2* is also heat inducible but positively regulated by RpoH (115, 252). *cpn.3* is weakly expressed and is only detectable in anaerobically grown cells (252). Expression of *cpn.3* is not detected in a *nifA* mutant, suggesting that it is regulated with nitrogen fixation genes.

The *cpn.1* gene is essential for growth, but the *cpn.2* and *cpn.3* single mutants and the *cpn.2 cpn.3* double mutant can be constructed (252). However, no growth or symbiotic phenotypes were observed for any of the mutants. All three genes are expressed in bacteroids, with Cpn60.1 comprising the dominant Cpn60 protein detected. Regulation of *cpn.3* with nitrogen fixation genes suggested a possible role during symbiosis, but the Cpn60.3 protein could not be detected in bacteroids.

Purified Cpn60 proteins display different biochemical properties *in vitro* (103). For example, the proteins display different stabilities in response to increasing temperature. Each is less stable than *E. coli* GroEL, which could reflect the differences in optimal growth temperature between the two organisms (37°C for *E. coli* and 28°C for *R. leguminosarum*). However, the stability transitions occur at temperatures well above the optimal growth temperatures of these organisms. All three Cpn60 proteins are able to interact with a denatured GroES-independent substrate, LDH (bovine lactate dehydrogenase), but Cpn60.3 is the least effective at refolding denatured LDH. Although the Cpn60 proteins demonstrate different properties *in vitro*, their functions *in vivo* could still be similar. A separate study sought to determine whether the *groEL*

genes encoded functionally equivalent proteins *in vivo* (114). The strategy was to replace the essential *cpn.1* gene with expression from the *cpn.2* or *cpn.3* genes. *cpn.2* is not well-expressed under any of the conditions tried, so its inability to function in place of *cpn.1* might be due to protein level. Expression of *cpn.3* enables creation of a *cpn.1* mutant strain only when Cpn60.3 is produced at levels 4 times greater than the level of endogenous Cpn60.1, and this mutant strain is temperature sensitive. This indicates that Cpn60.3 can perform some of the same functions of Cpn60.1, although not as effectively. Nitrogenase activity is reduced in bacteroids that lack Cpn60.3, but this cannot be the only activity of which it is capable, given it can partially function for Cpn60.1. Although this study suggests that these proteins are specialized for different functions *in vivo*, the precise identity of different substrates or functions remains elusive.

1.3.2.3 *Sinorhizobium meliloti*

S. meliloti maintains five *groEL* loci in the genome: *groESL1* and *groEL4* are located on the chromosome, *groESL2* and *groESL3* are located on the pSyma megaplasmid, and *groESL5* is located on the pSymb megaplasmid (*groES5* is not annotated) (97). The proteins encoded by the *groEL* genes are very similar (Fig. 8). GroEL1 is 99% identical to GroEL2, with only a single amino acid different between the two (serine to threonine substitution near the N-terminus) (Table 1). The most divergent GroEL is GroEL3, at 72-74% identity to the other GroEL proteins.

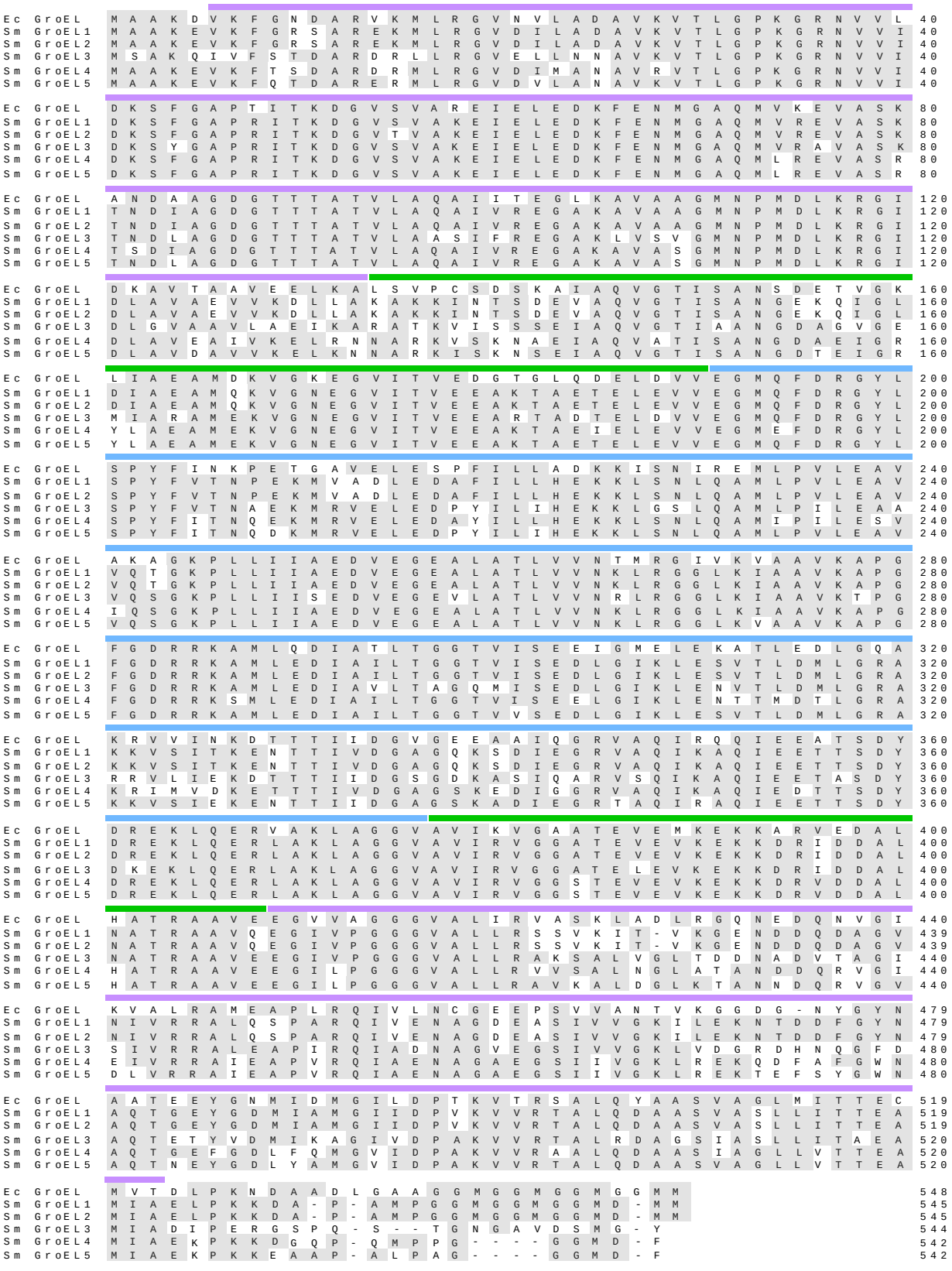
Ec GroEL M A A K D V K F G N D A R V K M L R G V N V L A D A V K V T L G P K G R N V V L 40
Sm GroEL1 M A A K E V K F G R S A R E K M L R G V D I L A D A V K V T L G P K G R N V V I 40
Sm GroEL2 M A A K E V K F G R S A R E K M L R G V D I L A D A V K V T L G P K G R N V V I 40
Sm GroEL3 M S A K Q I V F S T D A R D R L L R G V E L L N N A V K V T L G P K G R N V V I 40
Sm GroEL4 M A A K E V K F T S D A R D R M L R G V D I M A N A V R V T L G P K G R N V V I 40
Sm GroEL5 M A A K E V K F Q T D A R E R M L R G V D V L A N A V K V T L G P K G R N V V I 40
Ec GroEL D K S F G A P T I T K D G V S V A R E I E L E D K F E N M G A Q M V K E V A S K 80
Sm GroEL1 D K S F G A P R I T K D G V S V A K E I E L E D K F E N M G A Q M V R E V A S K 80
Sm GroEL2 D K S F G A P R I T K D G V T V A K E I E L E D K F E N M G A Q M V R E V A S K 80
Sm GroEL3 D K S Y G A P R I T K D G V S V A K E I E L E D K F E N M G A Q M V R A V A S K 80
Sm GroEL4 D K S F G A P R I T K D G V S V A K E I E L E D K F E N M G A Q M L R E V A S R 80
Sm GroEL5 D K S F G A P R I T K D G V S V A K E I E L E D K F E N M G A Q M L R E V A S R 80
Ec GroEL A N D A A G D G T T T A T V L A Q A I I T E G L K A V A A G M N P M D L K R G I 120
Sm GroEL1 T N D I A G D G T T T A T V L A Q A I V R E G A K A V A A G M N P M D L K R G I 120
Sm GroEL2 T N D I A G D G T T T A T V L A Q A I V R E G A K A V A A G M N P M D L K R G I 120
Sm GroEL3 T N D L A G D G T T T A T V L A A S I F R E G A K L V S V G M N P M D L K R G I 120
Sm GroEL4 T S D I A G D G T T T A T V L A Q A I V R E G A K A V A S G M N P M D L K R G I 120
Sm GroEL5 T N D L A G D G T T T A T V L A Q A I V R E G A K A V A S G M N P M D L K R G I 120
Ec GroEL D K A V T A A V E E L K A L S V P C S D S K A I A Q V G T I S A N S D E T V G K 160
Sm GroEL1 D L A V A E V V K D L L A K A K K I N T S D E V A Q V G T I S A N G E K Q I G L 160
Sm GroEL2 D L A V A E V V K D L L A K A K K I N T S D E V A Q V G T I S A N G E K Q I G L 160
Sm GroEL3 D L G V A A V L A E I K A R A T K V I S S S E I A Q V G T I A A N G D A G V G E 160
Sm GroEL4 D L A V E A I V K E L R N N A R K V S K N A E I A Q V A T I S A N G D A E I G R 160
Sm GroEL5 D L A V D A V V K E L K N N A R K I S K N S E I A Q V G T I S A N G D T E I G R 160
Ec GroEL L I A E A M D K V G K E G V I T V E D G T G L Q D E L D V V E G M Q F D R G Y L 200
Sm GroEL1 D I A E A M Q K V G N E G V I T V E E A K T A E T E L E V V E G M Q F D R G Y L 200
Sm GroEL2 D I A E A M Q K V G N E G V I T V E E A K T A E T E L E V V E G M Q F D R G Y L 200
Sm GroEL3 M I A R A M E K V G N E G V I T V E E A R T A D T E L D V V E G M Q F D R G Y L 200
Sm GroEL4 Y L A E A M E K V G N E G V I T V E E A K T A E I E L E V V E G M E F D R G Y L 200
Sm GroEL5 Y L A E A M E K V G N E G V I T V E E A K T A E T E L E V V E G M Q F D R G Y L 200
Ec GroEL S P Y F I N K P E T G A V E L E S P F I L L A D K K I S N I R E M L P V L E A V 240
Sm GroEL1 S P Y F V T N P E K M V A D L E D A F I L L H E K K L S N L Q A M L P V L E A V 240
Sm GroEL2 S P Y F V T N P E K M V A D L E D A F I L L H E K K L S N L Q A M L P V L E A V 240
Sm GroEL3 S P Y F V T N A E K M R V E L E D P Y I L I H E K K L G S L Q A M L P I L E A A 240
Sm GroEL4 S P Y F I T N Q E K M R V E L E D A Y I L L H E K K L S N L Q A M I P I L E S V 240
Sm GroEL5 S P Y F I T N Q D K M R V E L E D P Y I L I H E K K L S N L Q A M L P V L E A V 240
Ec GroEL A K A G K P L L I I A E D V E G E A L A T L V V N T M R G I V K V A A V K A P G 280
Sm GroEL1 V Q T G K P L L I I A E D V E G E A L A T L V V N K L R G G L K I A A V K A P G 280
Sm GroEL2 V Q T G K P L L I I A E D V E G E A L A T L V V N K L R G G L K I A A V K A P G 280
Sm GroEL3 V Q S G K P L L I I S E D V E G E V L A T L V V N R L R G G L K I A A V K T P G 280
Sm GroEL4 I Q S G K P L L I I A E D V E G E A L A T L V V N K L R G G L K I A A V K A P G 280
Sm GroEL5 V Q S G K P L L I I A E D V E G E A L A T L V V N K L R G G L K V A A V K A P G 280
Ec GroEL F G D R R K A M L Q D I A T L T G G T V I S E E I G M E L E K A T L E D L G Q A 320
Sm GroEL1 F G D R R K A M L E D I A I L T G G T V I S E D L G I K L E S V T L D M L G R A 320
Sm GroEL2 F G D R R K A M L E D I A I L T G G T V I S E D L G I K L E S V T L D M L G R A 320
Sm GroEL3 F G D R R K A M L E D I A V L T A G Q M I S E D L G I K L E N V T L D M L G R A 320
Sm GroEL4 F G D R R K S M L E D I A I L T G G T V I S E E L G I K L E N T T M D T L G R A 320
Sm GroEL5 F G D R R K A M L E D I A I L T G G T V V S E D L G I K L E S V T L D M L G R A 320
Ec GroEL K R V V I N K D T T T I I D G V G E E A A I Q G R V A Q I R Q Q I E E A T S D Y 360
Sm GroEL1 K K V S I T K E N T T I V D G A G Q K S D I E G R V A Q I K A Q I E E T T S D Y 360
Sm GroEL2 K K V S I T K E N T T I V D G A G Q K S D I E G R V A Q I K A Q I E E T T S D Y 360
Sm GroEL3 R R V L I E K D T T T I I D G S G D K A S I Q A R V S Q I K A Q I E E T A S D Y 360
Sm GroEL4 K R I M V D K E T T T I V D G A G S K E D I G G R V A Q I K A Q I E D T T S D Y 360
Sm GroEL5 K K V S I E K E N T T I I D G A G S K A D I E G R T A Q I R A Q I E E T T S D Y 360
Ec GroEL D R E K L Q E R V A K L A G G V A V I K V G A A T E V E M K E K K A R V E D A L 400
Sm GroEL1 D R E K L Q E R L A K L A G G V A V I R V G G A T E V E V K E K K D R I D D A L 400
Sm GroEL2 D R E K L Q E R L A K L A G G V A V I R V G G A T E V E V K E K K D R I D D A L 400
Sm GroEL3 D K E K L Q E R L A K L A G G V A V I R V G G A T E L E V K E K K D R I D D A L 400
Sm GroEL4 D R E K L Q E R L A K L A G G V A V I R V G G S T E V E V K E K K D R V D D A L 400
Sm GroEL5 D R E K L Q E R L A K L A G G V A V I R V G G S T E V E V K E K K D R V D D A L 400
Ec GroEL H A T R A A V E E G V V A G G G V A L I R V A S K L A D L R G Q N E D Q N V G I 440
Sm GroEL1 N A T R A A V Q E G I V P G G G V A L L R S S V K I T - V K G E N D D Q D A G V 439
Sm GroEL2 N A T R A A V Q E G I V P G G G V A L L R S S V K I T - V K G E N D D Q D A G V 439
Sm GroEL3 N A T R A A V E E G I V P G G G V A L L R A K S A L V G L T D D N A D V T A G I 440
Sm GroEL4 H A T R A A V E E G I L P G G G V A L L R V V S A L N G L A T A N D D Q R V G I 440
Sm GroEL5 H A T R A A V E E G I L P G G G V A L L R A V K A L D G L K T A N N D Q R V G V 440
Ec GroEL K V A L R A M E E A P L R Q I V L N C G E E P S V V A N T V K G G D G - N Y G Y N 479
Sm GroEL1 N I V R R A L Q S P A R Q I V E N A G D E A S I V V G G K I L E K N T D D F G Y N 479
Sm GroEL2 N I V R R A L Q S P A R Q I V E N A G D E A S I V V G G K I L E K N T D D F G Y N 479
Sm GroEL3 S I V R R A L E A P I R Q I A D N A G V E G S I V V G G K L V D G R D H N Q G F D 480
Sm GroEL4 E I V R R A I E A P V R Q I A E N A G A E G S I I V G G K L R E K Q D F A F G W N 480
Sm GroEL5 D L V R R A I E A P V R Q I A E N A G A E G S I I V G G K L R E K T E F S Y G W N 480
Ec GroEL A A T E E Y G N M I D M G G I L D P T K V T R S A L Q Y A A S V A G L M I T T E C 519
Sm GroEL1 A Q T G E Y G D M I A M G G I I D P V K V V R T A L Q D A A S V A S L L I T T E A 519
Sm GroEL2 A Q T G E Y G D M I A M G G I I D P V K V V R T A L Q D A A S V A S L L I T T E A 519
Sm GroEL3 A Q T E T Y V D M I K A G I V D P A K V V R T A L R D A G S I A S L L I T A E A 520
Sm GroEL4 A Q T G E F G D L F Q M G V I D P A K V V R A A L Q D A A S I A G L L V T T E A 520
Sm GroEL5 A Q T N E Y G D L Y A M G V I D P A K V V R T A L Q D A A S V A G L L V T T E A 520
Ec GroEL M V T D L P K N D A A D L G A A G G M G G M G G M G G M G G M M 548
Sm GroEL1 M I A E L P K K D A - P - A M P G G M G G M D - M M 545
Sm GroEL2 M I A E L P K K D A - P - A M P G G M G G M D - M M 545
Sm GroEL3 M I A D I P E R G S P Q - S - - T G N G A V D S M G - Y 544
Sm GroEL4 M I A E K P K K D G Q P - Q M P P G - - - - G G M D - F 542
Sm GroEL5 M I A E K P K K E A A P - A L P A G - - - - G G M D - F 542

Figure 8. Alignment of *E. coli* and *S. meliloti* GroEL proteins.

The alignment was generated using the CLUSTAL V method (136) in the MegaAlign program from LASERGENE (DNASTAR, Inc.). Residues in gray match the consensus sequence. Domains of the GroEL protein are indicated as follows: apical domain (blue line), intermediate domain (green line), and equatorial domain (purple line). Abbreviations are *E. coli* (Ec) and *S. meliloti* (Sm).

Table 1. Pairwise identity between GroEL proteins in *S. meliloti*.

	GroEL2	GroEL3	GroEL4	GroEL5
GroEL1	99	74	79	82
GroEL2		74	79	82
GroEL3			72	74
GroEL4				88

Some of the *groEL* genes are differentially regulated. Experiments using transcriptional reporter fusions demonstrated that *groESL1* is more highly expressed than *groESL2* during free-living growth (222). Expression of both operons appears to be repressed by the *groEL1* gene product. Transcription of *groESL1* and *groESL5* is induced upon heat shock, although the mechanism of induction differs between the two (196). Induction of *groESL5* is RpoH1-dependent, and *groESL1* might be controlled by the putative CIRCE sequence located upstream of the open reading frame. Although *groESL2* also has a putative CIRCE element, its expression is not induced by heat shock. Expression levels during symbiosis have only been determined for *groESL1* and *groESL2*, where *groESL1* is more highly expressed than *groESL2* (222). However, expression of each of the *groE* loci within nodules on a variety of hosts has been detected in one or more of the global transcript or protein analyses (4, 16, 18, 58, 59, 215). Unfortunately, these analyses do not enable direct comparison of expression levels from the *groEL* genes.

The *groEL1* locus was originally identified in a genetic screen for reduced *nod* gene expression (223). The *nod* genes encode enzymes that produce Nod factor, which elicits nodule formation during symbiosis. The genes are controlled by several related transcription factors (NodD1, NodD2, and NodD3), some of which require plant inducers for activity. In cell free extracts, binding of NodD1 to the *nodA* promoter is reduced in a *groEL1* mutant, suggesting that *groEL1* is specifically involved in this activity. *In vitro* analysis showed that preincubation of NodD3 with GroEL-GroES enhanced DNA binding, while the DNA-binding activity of NodD1 required both the plant flavonoid luteolin and GroEL-GroES (316). In addition, GroEL copurifies with NodD1 and NodD3, although the identity of the GroEL protein was not determined (223, 316). Taken together, these results suggest that GroEL affects *nod* gene expression through physical interactions with the NodD transcription factors during symbiosis.

Additionally, the *groEL1* locus was identified in a screen for genes required for the production of *N*-acyl homoserine lactones used in quorum sensing (183). The authors suggest that deficiency in *N*-acyl homoserine lactone production may be due to interaction of GroEL-GroES with the TraR regulator, much like the interaction suggested for the NodD transcriptional activators.

Given its role in *nod* gene expression, it is not surprising that the *groEL1* mutant displays symbiotic phenotypes (223). This mutant is delayed in nodule formation and cannot fix nitrogen during symbiosis. Nitrogen fixation genes are expressed in the *groEL1* mutant, but possible posttranslational effects on nitrogenase proteins have not been explored. A *groEL1* mutant also displays a subtle growth phenotype under normal growth conditions. All single *groEL* mutants have been constructed (20, 196, 222, 223), but *groEL1* is the only mutant with an altered phenotype under any of the conditions tested. However, expression of *groESL2* from a multicopy plasmid was able to complement a *groEL1* mutation during symbiosis, suggesting that the virtually identical proteins encoded by these genes could be functionally equivalent (223). Differences in *groEL1* and *groEL2* mutant phenotypes likely result from differences in gene expression levels. A *groEL1 groEL2* double mutant cannot be constructed, which provides further evidence that the encoded proteins have overlapping roles that are essential to cell survival (222).

1.4 WHY DOES *SINORHIZOBIUM MELILOTI* MAINTAIN MULTIPLE COPIES OF *RPOH* AND *GROEL* GENES?

Many bacterial species (20% of sequenced genomes) maintain multiple copies of genes that encode the chaperonin GroEL, but only a few members of the α-proteobacteria maintain

multiple copies of genes that encode the stress response sigma factor RpoH. Why do bacteria have multiple copies of these genes? Possibly the genes are differentially regulated but encode proteins that perform similar functions, which would allow bacteria to modulate their expression levels in response to different stimuli. Alternatively, the genes could encode proteins that have different functions. For multiple RpoH proteins, this would involve regulating the transcription of different sets of genes. Multiple GroEL proteins could fold different substrates or perform different non-folding functions. Differential regulation and different function are not mutually exclusive, and there is evidence for both in different bacterial systems.

Intriguingly, all of the root-nodulating *Rhizobium* species whose genomes have been completely sequenced maintain multiple copies of *rpoH* and *groEL* genes. Their copies show evidence of differential regulation and different function, as suggested above. However, the *Rhizobium* species are particularly interesting because of their complex lifestyles: they exist as either free-living bacteria or in symbiosis with plant hosts. Because these nitrogen-fixing symbionts have multiple copies, it is possible that *rpoH* and *groEL* genes play important roles in symbiosis. Why might multiple copies of stress response genes be important for symbiosis? First, the bacteria are exposed to oxidative and pH stresses, at a minimum, during symbiosis, and stress response proteins might be required to withstand and to repair cellular damage. Second, multiple stress response genes might be required for production of new proteins during symbiosis. Initiation, infection, and bacteroid development almost certainly require the synthesis of many proteins that are not present during free-living growth. Third, these multiple stress response genes might be required for symbiosis-specific functions. While the RpoH regulons in several pathogens overlap with the *E. coli* RpoH regulon, they also include targets not found in *E. coli*, some of which might be involved in pathogenesis (55, 128, 273). Multiple *groEL* genes

might also be specialized to fold symbiosis-specific substrates, such as the proteins that form the nitrogenase complex. For example, GroEL has been implicated in the regulation of *nif* gene transcription and in activating nitrogenase subunits posttranslationally (84, 116, 117).

I am studying multiple *rpoH* and *groEL* genes in the model symbiont *S. meliloti*. Previous studies have demonstrated that some of these genes are regulated differentially (196, 222, 226, 261), have overlapping functions (222, 227), or are required for successful symbiosis (223, 226, 227). The goal of this thesis was to further characterize the functions of multiple *rpoH* and *groEL* genes in *S. meliloti*. In Chapter 2, I tested the hypothesis that the *rpoH* genes are required during symbiosis simply to maintain sufficient levels of GroEL, a key target of RpoH in *E. coli*. I show that this hypothesis is not correct, and that other targets must be required. To identify these targets, in Chapter 3 and Appendices A and B I describe microarray experiments to determine the RpoH1 and RpoH2 regulons during stationary phase and heat shock and under conditions where the *rpoH* genes are overexpressed. In Chapter 4, I performed a genetic analysis of multiple *groEL* mutants to uncover possible functional redundancies. I show that either *groEL1* or *groEL2* is essential to cell survival, but only *groEL1* is necessary and sufficient for symbiosis.

2.0 MULTIPLE *GROESL* OPERONS ARE NOT KEY TARGETS OF RPOH1 AND RPOH2 IN *SINORHIZOBIUM MELILOTI*

This chapter is adapted from previously published material in *Journal of Bacteriology*, Volume 188, Number 10, pages 3507-3515. It is reproduced here with permission from the American Society for Microbiology Journals Division.

2.1 ABSTRACT

Among the Rhizobia that establish nitrogen-fixing nodules on the roots of host plants, many contain multiple copies of genes encoding the sigma factor RpoH and the chaperone GroEL-GroES. In *Sinorhizobium meliloti* there are two *rpoH* genes, four *groESL* operons, and one *groEL* gene. *rpoH1* mutants are defective for growth at high temperature and form ineffective nodules, *rpoH1 rpoH2* double mutants are unable to form nodules, and *groEL1* mutants form ineffective nodules. To explore the roles of RpoH1 and RpoH2, mutants that suppress both the growth and nodulation defects were identified. These mutants do not suppress the nitrogen fixation defect. This implies that the functions of RpoH1 during growth and RpoH1/RpoH2 during the initiation of symbiosis are similar but that there is a different function of RpoH1

needed later during symbiosis. I showed that, unlike in *E. coli*, overexpression of *groESL* is not sufficient to bypass any of the RpoH defects. Under free-living conditions I determined that RpoH2 does not control expression of the *groEL* genes, and RpoH1 only controls expression of *groEL5*, which is not required for symbiosis. Taken together, these results suggest that GroEL-GroES production alone cannot explain the requirements for RpoH1 and RpoH2 in *S. meliloti* and that there must be other crucial targets.

2.2 INTRODUCTION

Sinorhizobium meliloti can be found as a free-living bacterium residing in the soil or as a nitrogen-fixing symbiont residing in nodules on the roots of leguminous host plants, such as alfalfa. The *S. meliloti* genome contains 14 genes for sigma factors (97), which are subunits of RNA polymerase that direct transcription initiation by recognizing promoters. Two of these genes, *rpoH1* and *rpoH2*, encode members of the RpoH family of secondary sigma factors. RpoH (σ^{32}) was originally identified in *E. coli* as a sigma factor that responds to heat shock. In response to a sudden increase in temperature or other stresses, the levels of RpoH rise transiently, inducing transcription of a subset of genes encoding heat shock proteins (HSPs). HSPs include chaperones involved in protein folding, such as GroEL-GroES and DnaK-DnaJ-GrpE, and proteases, such as FtsH and Lon (321). Although RpoH and the HSPs were identified as part of the heat shock response, these proteins are present at low temperature and play important roles in cellular processes under non-stress conditions, such that the *rpoH* gene in *E. coli* is essential above 20°C (324). The requirement for RpoH in *E. coli* can largely be explained as a requirement for expression of the *groESL* operon because overexpression of *groESL* is

sufficient to suppress the temperature sensitive growth defect of the *rpoH* mutant from 20-40 °C (167).

Although the *E. coli* genome only contains one *rpoH* gene and one *groESL* operon, other bacterial genomes contain multiple copies of these genes. In particular, many *Rhizobium* species have multiple *rpoH* and *groESL* genes, and mutations in some of them result in symbiotic defects. In *S. meliloti*, in addition to two *rpoH* genes (226, 227), there are four *groESL* operons (97, 223, 225, 256) and one single *groEL* gene (35). *rpoH1* and *groEL1* mutants are unable to fix nitrogen (Fix^-) (223, 226, 227), and *rpoH1 rpoH2* double mutants are unable to form nodules (Nod^-) (227). In *Bradyrhizobium japonicum*, there are three *rpoH* genes (213, 214), five *groESL* operons (83), and two single *groEL* genes (157). $groESL_3$ is regulated with nitrogen fixation genes (83), and a $groESL_3$ $groESL_4$ double mutant is unable to fix nitrogen (84). *Rhizobium* sp. strain TAL1145 has at least one *rpoH* gene, and the *rpoH* mutant exhibits reduced nodulation, resulting in stunted plant growth (162). *Rhizobium leguminosarum* has at least three *groESL* operons (252, 302), and *Mesorhizobium loti* has two *rpoH* genes and five *groESL* operons (156, 162). Interestingly, the genome of the closely related plant pathogen *Agrobacterium tumefaciens*, also a member of the *Rhizobiaceae*, only contains single copies of these genes (310). The reason for multiple *rpoH* and *groESL* genes in these plant endosymbionts is unclear. Are the genes regulated differentially but encode proteins with similar functions, or do they encode proteins with specialized functions?

The *rpoH* genes in *S. meliloti* were identified as members of the *rpoH* family by sequence analysis and by the ability to complement an *E. coli rpoH* mutation (226, 227). Under free-living conditions, *rpoH1* mutants exhibit a slight growth defect at the optimum growth temperature (30°C) and a severe defect at higher temperatures (226). During symbiosis, *rpoH1*

mutant cells invade the nodule and differentiate into bacteroids but undergo early senescence (196), resulting in a Fix$^-$ phenotype (226, 227). *rpoH2* mutants have no discernable phenotype under free-living or symbiotic conditions (226, 227). However, Ono *et al.* (227) discovered that an *rpoH1 rpoH2* double mutant is unable to form nodules.

Transcriptional reporter fusions to *rpoH1* and *rpoH2* have shown that *rpoH1* is transcribed during stationary phase in LB rich medium and M9 minimal medium and that *rpoH2* is transcribed during stationary phase only in M9 medium. During symbiosis, *rpoH1* is strongly expressed throughout the nodule whereas *rpoH2* is not expressed in the nodule except for low levels at the tip and variable punctate spots at other locations (226). The phenotypes and expression data suggest that *rpoH1* and *rpoH2* have distinct but overlapping functions.

The presence of a family of 4-5 *groEL* genes in *S. meliloti* was initially discovered by Southern blot analysis (256). Additional work by other groups and subsequent sequencing of the *S. meliloti* genome has led to a final count of four *groESL* operons and one *groEL* gene (35, 97, 223, 225, 256). The names of the *groE* genes used in this paper are those given in the genome annotation, although *groES5* was not annotated (97).

A connection between GroEL-GroES and symbiosis was uncovered when *groEL1* was identified in a genetic screen for *S. meliloti* genes required for full induction of *nod* genes, which are required for formation of a bacterial signal that initiates nodule formation by host plants (223). The *groEL1* mutation affects the activities of several related transcription factors (NodD1, NodD3, and SyrM) that activate expression of *nod* genes, and GroEL copurifies with NodD1 and NodD3 (86, 223). *In vitro* work has demonstrated that the NodD proteins are substrates for GroEL-GroES, resulting in modulation of the DNA binding activity (316). *groEL1*, *groEL2*, and *groESL5* mutants have been studied. *groEL1* mutants have a slight growth

defect, are delayed for nodulation, and form Fix$^-$ nodules (223). A *groEL2* mutant displays neither a growth nor a symbiotic defect, but the *groEL1 groEL2* double mutant is not viable (222). The *groESL5* mutant has no symbiotic defect (196). GroEL1/GroES1 and GroEL2/GroES2 are the most similar to each other (99% identical for GroEL and 97% identical for GroES), whereas GroEL3/GroES3 are the most dissimilar from any other *S. meliloti* homologs (72-74% identical for GroEL and 75-78% identical for GroES).

Mitsui *et al*. (196) tested whether RpoH1 or RpoH2 controls expression of the *groESL* genes in *S. meliloti* during heat shock. *groESL5* was the only *groESL* operon whose transcription was controlled by RpoH1, and none of the genes were controlled by RpoH2. However, this work did not explore regulation during stationary phase and within the nodule, other conditions where the *rpoH1* and *rpoH2* genes are known to be expressed (226).

Given that *groESL* is a crucial target of RpoH in *E. coli* and that *groEL1*, *rpoH1*, and *rpoH1 rpoH2 S. meliloti* mutants have symbiotic phenotypes, I hypothesized that *groESL* might also be a key target of RpoH in *S. meliloti*. However, I used suppressor mutant analysis and overexpression experiments to demonstrate that the relationships between RpoH and GroEL-GroES are different in the two organisms. Specifically, my results suggest that GroEL-GroES production is not sufficient to bypass the requirements for RpoH1 or RpoH1/RpoH2 during growth and symbiosis and that there must be other crucial targets. In addition, I showed that only *groESL5* is controlled by RpoH1 during free-living growth and stationary phase at 30°C, which agrees with results obtained by Mitsui *et al.* during growth and heat shock (196).

2.3 MATERIALS AND METHODS

2.3.1 Strains, plasmids, and growth conditions

Bacterial strains and plasmids used in this work are listed in Table 2. Bacterial cultures were grown in LB medium, LB/MC medium (109), or M9 minimal medium containing 0.2% sucrose, 0.5 μg biotin ml^{-1}, 1 mM $MgSO_4$, and 0.25 mM $CaCl_2$. Antibiotics were added to the media as follows: 100 μg ampicillin ml^{-1}, 25 μg gentamicin ml^{-1}, 5-50 μg hygromycin ml^{-1}, 25 μg kanamycin ml^{-1}, 50/200 μg neomycin ml^{-1}, 50/200 μg spectinomycin ml^{-1}, 500 μg streptomycin ml^{-1}, and 2/10 μg tetracycline ml^{-1}. *S. meliloti* cells were grown at 30°C unless otherwise indicated. Plasmids were introduced into *S. meliloti* cells by triparental conjugation (109). Chromosomally located constructs were moved between *S. meliloti* strains by generalized transduction using N3 phage (184). Although the *rpoH2::aacC1* containing strain BY294 (227) was constructed in the Rm1021 background, I transferred the mutation by transduction into our laboratory Rm1021 strain, creating AB3, to ensure isogenicity.

Table 2. Strains used in Chapter 2

Strain	Relevant characteristics	Reference
AB3	*rpoH2::aacCI*	(20)
AB4	Wild type/pAB1 (P_{lac}-*groESL1*)	(20)
AB9	*rpoH1::aadA rpoH2::aacCI*	(20)
AB16	Wild type/pAB2 (P_{lac}-*groESL3*)	(20)
AB35	*rpoH1::aadA rpoH2*::pVO101 NDS-3	(20)
AB36	*rpoH1::aadA rpoH2*::pVO101 NDS-4	(20)
AB37	*rpoH1::aadA rpoH2*::pVO101 NDS-5	(20)
AB38	*rpoH1::aadA rpoH2*::pVO101 NDS-6	(20)
AB39	*rpoH1::aadA rpoH2*::pVO101 NDS-7	(20)
AB40	*rpoH1::aadA rpoH2*::pVO101 NDS-8	(20)
AB41	*rpoH1::aadA rpoH2*::pVO101 NDS-9	(20)
AB42	*rpoH1::aadA rpoH2*::pVO101 NDS-10	(20)
AB43	*rpoH1::aadA rpoH2*::pVO101 NDS-11	(20)
AB44	*rpoH1::aadA rpoH2*::pVO101 NDS-12	(20)
AB92	Wild type/pAB7 (P_{trp}-*groESL1*)	(20)
AB103	Wild type/pAB8 (P_{trp}-*groESL3*)	(20)
AB129	*groEL2*::pAB10 (*groEL2-gfp-gus* transcriptional fusion, *groEL2* not disrupted)	(20)
AB140	*groEL1*::pAB11 (*groEL1-gfp-gus* transcriptional fusion, *groEL1* not disrupted)	(20)
AB145	*groEL3*::pAB12 (*groEL3-gfp-gus* transcriptional fusion, *groEL3* not disrupted)	(20)
AB147	*groEL4*::pAB13 (*groEL4-gfp-gus* transcriptional fusion, *groEL4* not disrupted)	(20)
AB150	*groEL5*::pAB14 (*groEL5-gfp-gus* transcriptional fusion, *groEL5* not disrupted)	(20)
B4T1	*groEL1*::Tn*5*	(223)
BY249	*rpoH2::aacCI*	(227)
Rm1021	Wild type	(188)
VO2148	*rpoH2*::pVO101 (*rpoH2* disruption)	(226)
VO3128	*rpoH1::aadA*	(226)
VO3148	*rpoH1::aadA rpoH2*::pVO101	(20)
VO3149	*rpoH1::aadA rpoH2*::pVO101	(20)
VO3150	*rpoH1::aadA rpoH2*::pVO101 NDS-1	(20)
VO3151	*rpoH1::aadA rpoH2*::pVO101 NDS-2	(20)
VO3165	*rpoH1::aadA* GDS-1	(20)
VO3166	*rpoH1::aadA rpoH2*::pVO101 GDS-2	(20)
VO3170	*rpoH1::aadA rpoH2*::pVO101 GDS-1	(20)

2.3.2 Plant assays

Alfalfa plants (*Medicago sativa* GT13R plus) were grown on nitrogen free BNM medium and inoculated with *S. meliloti* cells as previously described (225). Plant height, leaf color, and nodule color were scored at six weeks post inoculation to determine the status of nitrogen fixation. Inoculation with Fix^+ bacteria results in tall, green plants with pink nodules. Inoculation with Fix^- bacteria results in stunted, chlorotic plants with white nodules. Bacteria were isolated from nodules by surface sterilizing nodules in 20% Clorox bleach for 5 minutes, washing two times with water and one time with LB medium, crushing with forceps, and then streaking on LB medium.

2.3.3 Western blot analysis

To obtain samples for western blot analysis, cells were grown overnight at 30°C in LB/MC medium with streptomycin, diluted to OD_{595} of 0.1, grown to mid log phase ($0.6 \leq OD_{595} \leq 0.8$), harvested, and stored at -80°C. Cells were resuspended in 1X phosphate-buffered saline at 0.1 ml per OD_{595} unit. The cells were disrupted by sonication, and the resulting extracts were combined with 2X Laemmli sample buffer. Equal volumes of extract were separated by SDS-PAGE and transferred to nitrocellulose. Blots were probed with rabbit polyclonal antibodies to the *E. coli* proteins at the following dilutions: anti-GroEL (Stressgen) at 1: 5000, anti-DnaK (gift from J. Brodsky) at 1:5000 or anti-DnaK (Upstate Biotechnology) at 1:2500, and anti-DnaJ (Stressgen) at 1:1250. Blots were then probed with a 1:15000 dilution of anti-rabbit horseradish peroxidase-conjugated secondary antibody (Amersham), developed with enhanced chemiluminescence reagents (Pierce), and imaged using a Fujifilm LAS-3000 imaging system.

To quantify relative protein levels, band density was determined by ImageGauge software (Fuji). Protein concentration of cell lysates was determined by BCA protein assay (Pierce), and band intensities were then normalized to protein concentration.

2.3.4 Statistical analysis

Significance of differences in bacterial growth levels and protein levels was determined by using both the Student's *t*-test and the Wilcoxon rank sum test, which does not assume a normal distribution. Although the *P* values varied, differences were significant with both tests unless otherwise noted.

2.3.5 Construction of plasmids for overexpression of *groESL1* and *groESL3*

To place *groESL1* under the control of the *E. coli lac* promoter, a 2.1 kb DNA fragment that extends from 68 bp upstream of the *groES1* start codon to 30 bp downstream of the *groEL1* stop codon was amplified using primers that generate *Apa* I and *Xba* I restriction sites. The fragment was inserted into *Apa* I-*Spe* I digested pMB403 (14), a broad-host-range vector that contains the *lac* promoter, creating pAB1 (P_{lac}-*groESL1*). To place *groESL3* under the control of the *lac* promoter, a 2.2 kb fragment that extends from 78 bp upstream of the *groES3* start codon to 85 bp downstream of the *groEL3* stop codon was amplified with primers generating *Apa* I and *Xba* I restriction sites. The fragment was inserted into *Apa* I-*Spe* I digested pMB403, creating pAB2 (P_{lac}-*groESL3*).

To place *groESL1* and *groESL3* under the control of the *S. typhimurium trp* promoter, the *lac* promoter was removed from pAB1 and pAB2 and replaced with a fragment containing 141

bp of *S. typhimurium* DNA containing the *trp* promoter. First, a 220 bp *Eco* RV-*Acc* 65I fragment containing P_{trp} was isolated from pVO131. To delete the *lac* promoter, pAB1 and pAB2 were digested with *Nsi* I, blunted with T4 DNA polymerase, and digested with *Acc* 65I. The P_{trp} fragment was then inserted into pAB1 to create pAB7 (P_{trp}-*groESL1*) and into pAB2 to create pAB8 (P_{trp}-*groESL3*).

To confirm that expression of the *groESL1* operon from pAB1 was due to the *lac* promoter and that the *groESL1* fragment lacked its native promoter, the fragment was cut from pAB1 using *Kpn* I and *Xba* I and inserted into *Kpn* I/*Xba* I-digested pAB3 to create pAB6. pAB3 is a derivative of pMB393 (14) that lacks the *lac* promoter and contains the *trp* terminator to prevent transcription through the polylinker. Therefore, the *groESL1* fragment in pAB6 should be promoter-less if the native promoter is absent.

The expression plasmids were introduced into wild type (Rm1021), *groEL1*::Tn*5* (B4T1), *rpoH1::aadA* (VO3128), and *rpoH1::aadA rpoH2*::pVO101 (VO3148) by triparental conjugation.

2.3.6 Construction of *groEL-gus* fusions

The *groEL-gus* fusions were constructed using recombinational cloning as described in House *et al.* (143). This method is a modification of Invitrogen's Gateway Technology, such that transfer of DNA from an entry vector to a destination vector by the λ recombination system is performed *in vivo* via a pentaparental mating. In brief, each *groEL* ORF was transferred from an entry vector (pESmc00913, pESma0744, pESma0124, pESmc01758, and pESmb21566) (263) to the destination vector pMK2030 (B.K. Schroeder, B.L. House, M.W. Mortimer, and M.L. Kahn, unpublished data) during a pentaparental mating using the helper plasmid pRK2013 (81) and the

λ integrase- and excisionase-expressing plasmid pXINT129 (236). This destination vector is a suicide vector that contains *attR* recombination sites upstream of promoterless *gfp* and *gus* genes to allow the formation of transcriptional fusions. Each *groEL-gfp-gus* containing plasmid was moved into Rm1021 by triparental mating and integrated at the respective *groEL* gene by single reciprocal recombination, resulting in a P_{groE}-*groES-groEL-gfp-gus* construct. The resulting strains AB140 (*groESL1-gfp-gus*), AB129 (*groESL2-gfp-gus*), AB145 (*groESL3-gfp-gus*), AB147 (*groEL4-gfp-gus*), and AB150 (*groESL5-gfp-gus*) were confirmed by Southern analysis. The fusions were transferred into *rpoH1::aadA* (VO3128), *rpoH2::aacCI* (AB3), and *rpoH1::aadA rpoH2::aacCI* (AB9) mutant backgrounds by transduction.

2.3.7 Assay of β-glucuronidase activity

Cells were collected for β-glucuronidase assays at the indicated times and frozen at –80 °C until assayed for activity. The cells were permeabilized with lysozyme (200 μg ml^{-1}, 37 °C for 10 minutes), and β-glucuronidase activity was assayed using *p*-nitrophenyl-β-D-glucuronide as described previously (151). GUS activity is expressed in (nmol min^{-1} per OD_{595} unit) x 1000.

2.4 RESULTS

2.4.1 Suppression of the *rpoH1* and *rpoH1 rpoH2* mutant defects

The *rpoH1 rpoH2* double mutant RmHM9 was reported to be Nod^- (227). Using *rpoH1* and *rpoH2* mutant alleles (226), Valerie Oke generated two isolates of an *rpoH1 rpoH2* double

mutant (VO3148 and VO3149) (20). When Valerie Oke and I inoculated *Medicago sativa* GT13R plus alfalfa plants under our growth conditions with these strains, as well as RmHM9, we found that the double mutants varied greatly in the ability to nodulate plants from experiment to experiment (average of 36% nodulated plants with a range from 10-78%) (Fig. 9). There were two possible explanations for the variability: either the *rpoH1 rpoH2* phenotype is leaky or the nodules contain suppressor mutants. To distinguish between these possibilities, we isolated bacteria from twelve nodules elicited by the *rpoH1 rpoH2* mutants in two independent experiments and confirmed that both mutations were still present (data not shown). I used these strains to inoculate alfalfa and found that they were similar to the wild type in nodulation efficiency (Fig. 9) although still defective in nitrogen fixation (data not shown). Therefore, the nodules are due to suppressor mutants, which are called NDS for nodulation defect suppressor.

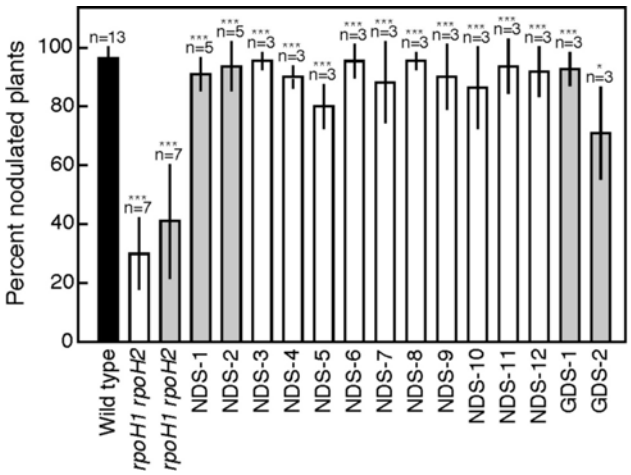

Figure 9. Nodulation by suppressor mutant strains.

Alfalfa plants were inoculated with control and suppressor mutant strains (NDS and GDS), and percent nodulated plants was determined after at least 3 weeks incubation. The graph depicts the average percent nodulation over the indicated number of experiments, and error bars represent the sample standard deviation. At least 50 plants in total were inoculated with each bacterial strain. White bars indicate strains derived from the *rpoH1 rpoH2* mutant strain VO3148, and grey bars indicate strains derived from the *rpoH1 rpoH2* mutant strain VO3149. Significance was determined using the Student's *t*-test with *** indicating $P < 0.001$ and * indicating $P < 0.05$. The strains from left to right are Rm1021, VO3148-VO3151, AB35-AB44, VO3170, and VO3166.

Cells containing an *rpoH1* mutation grow more slowly than wild-type cells in LB medium at 30°C (226). Because I was concerned about the generation of suppressor mutations, I looked for conditions in which the *rpoH1* mutant cells would grow as well as the wild-type cells. I switched to LB medium supplemented with $MgSO_4$ and $CaCl_2$ (LB/MC) (109). As shown in Fig. 10, *rpoH1 rpoH2* mutant cells grow like wild-type cells at 30°C in LB/MC but display a severe growth defect at 40°C.

To determine if the *rpoH1 rpoH2* nodulation suppressor mutations also suppressed the growth defect, I compared the growth of two independent suppressors strains (NDS-1 and NDS-3) to the *rpoH1 rpoH2* parent strains (VO3149 and VO3148, respectively) at 30°C and 40°C in LB/MC. As shown in Fig. 10, NDS-1 grows slightly poorer than the wild type at 30°C, whereas NDS-3 is indistinguishable. At 40°C neither NDS-1 nor NDS-3 cells grow as well as wild-type cells, but both appear to grow better than the *rpoH1 rpoH2* parent strain. To determine if the difference was significant, Valerie Oke compared the amount of growth as measured by OD_{595} at 48 hours and performed the Student's *t*-test (20). The OD_{595} of NDS-1 at 48 hours was significantly higher than the double mutant parent ($P < 0.05$), whereas NDS-3 was not significantly different.

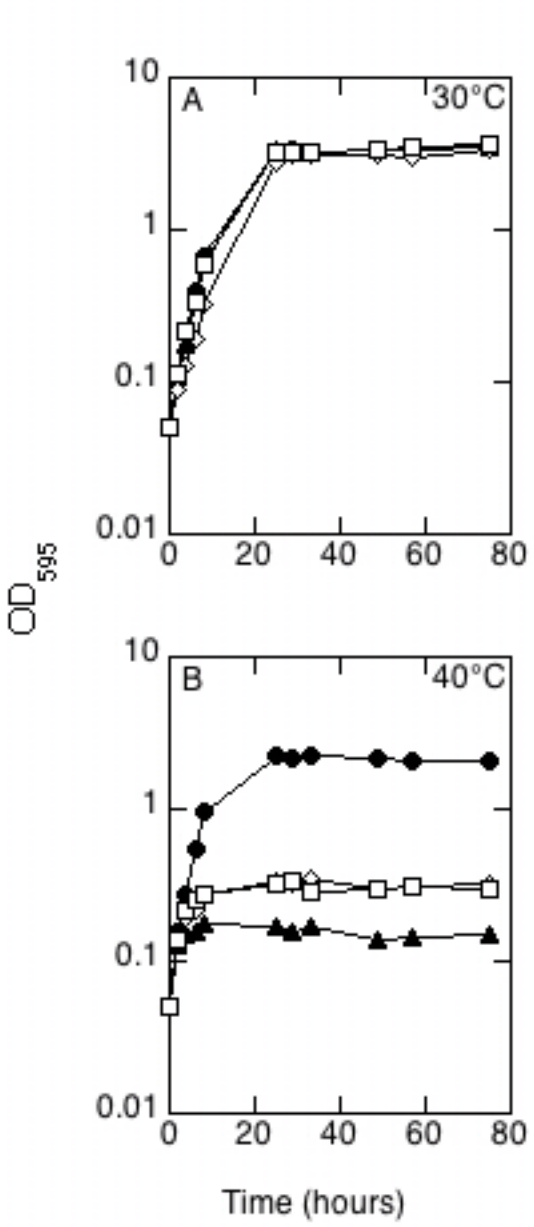

Figure 10. Comparison of the growth of suppressor mutant cells with wild-type and *rpoH1 rpoH2* double mutant cells, as measured using OD_{595}.

Cells were grown in LB/MC medium plus streptomycin at 30°C (A) or 40°C (B). The control strains are Rm1021 (wild type, filled circles) and VO3148 (*rpoH1 rpoH2*, filled triangles); and the suppressor mutant strains are VO3150 (NDS-1, open diamonds) and AB35 (NDS-3, open squares). The panels show representative data from one of four experiments.

In *E. coli*, suppressors of the *rpoH* growth defect are readily obtained by plating *rpoH* mutant cells at 30-40°C (167). By streaking for single colonies, Valerie Oke found that wild-type *S. meliloti* cells form colonies on LB/MC plates at 42°C, whereas cells containing an *rpoH1* mutation do not (20). To select for suppressor mutants, *rpoH1* and *rpoH1 rpoH2* mutant cells were plated at high density at 42°C (20). Many of the suppressor mutations were not stable, such that the ability to grow at 42°C was lost upon streaking for single colonies at 30°C or 42°C. However, by selecting for growth at 42°C multiple times, two independent, stable suppressor mutants were obtained and called GDS-1 (*rpoH1* background) and GDS-2 (*rpoH1 rpoH2* background) for growth defect suppressor. To facilitate characterization of GDS-1 for suppression of the growth defect as well as the nodulation defect, the *rpoH2* mutation was introduced into the cells by generalized transduction so that all of the suppressor mutants were in the *rpoH1 rpoH2* background.

To determine whether the growth defect suppressor mutations also suppress the nodulation and nitrogen fixation defects, Valerie Oke and I inoculated alfalfa plants with GDS-1 and GDS-2. The growth defect suppressor mutants nodulated alfalfa plants at levels significantly higher than the parent strains (Fig. 9), indicating suppression of the nodulation defect. However, the strains were still unable to fix nitrogen (data not shown).

In *E. coli*, an *rpoH* mutant cannot grow above 20°C (324). Suppressor mutants selected at 30°C to 40°C display increased expression of the *groESL* operon, and suppressor mutants selected at 42°C display increased expression of both *groESL* and *dnaK*. The increased transcription and subsequent synthesis of these HSPs in the suppressor mutants is independent of heat shock, such that high levels are observed at 30°C unlike in wild-type cells (167). To test whether the NDS or GDS suppressor mutants function by a similar mechanism, I grew cells to

mid log phase at 30°C and performed western blot analysis for GroEL and the DnaK-DnaJ chaperone complex using polyclonal antibodies generated to the *E. coli* proteins (Fig. 11). Each antibody recognized a major band of the appropriate molecular weight in *S. meliloti* cell extracts. In the case of GroEL, the polyclonal antibody recognizes the *S. meliloti* GroEL1, GroEL2, and GroEL5 proteins (19). It is likely that the antibody also recognizes GroEL3 and GroEL4 since all of the *S. meliloti* GroEL proteins are 57-62% identical to *E. coli* GroEL. Although the level of total GroEL appears lower in the *rpoH1 rpoH2* mutant as compared to the wild type and the level of DnaK appears higher in the *rpoH1* and *rpoH1 rpoH2* mutants as compared to the wild type, the differences were not significant using the Student's *t*-test and just significant using the Wilcoxon rank sum test ($P = 0.0496$).

Analysis of the suppressor mutants shows that GroEL, DnaK, and DnaJ protein levels were not significantly higher in the mutants compared to the *rpoH1* and the *rpoH1 rpoH2* parent strains using the Student's *t*-test and the Wilcoxon rank sum test. However, it is possible that a small increase of one particular GroEL protein is masked by GroEL1, since *groESL1* is expressed at the highest levels (Fig. 13). In the case of DnaJ, there was a significant decrease in protein levels relative to the *rpoH1 rpoH2* double mutant in NDS-1 ($P < 0.01$) and NSD-2 ($P < 0.05$). Since the *S. meliloti* suppressor mutants do not exhibit the increased production of GroEL and DnaK seen in the *E. coli* suppressor mutants, the suppression appears to function by a different mechanism.

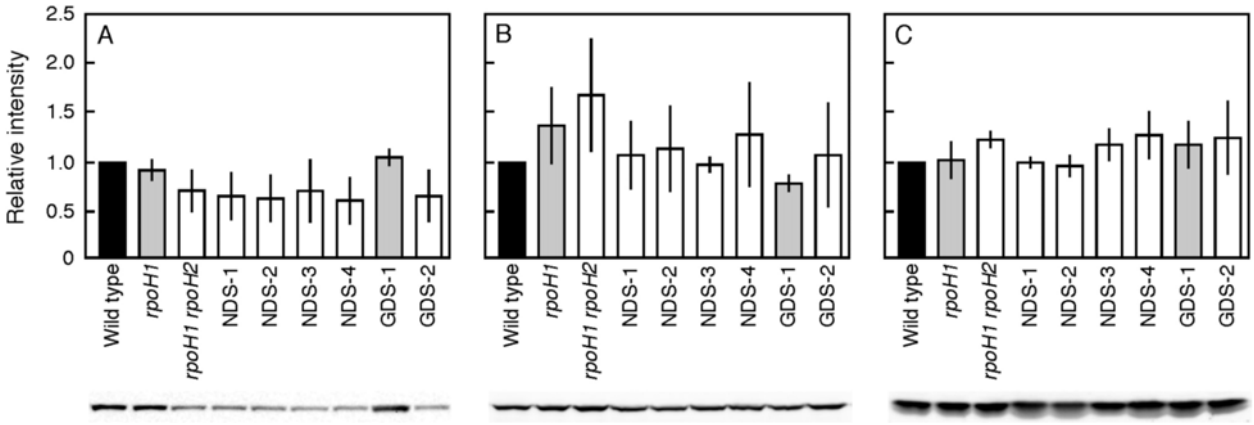

Figure 11. Western analysis of heat shock proteins in suppressor mutant strains.

Cells were grown to mid log phase in LB/MC medium plus streptomycin at 30°C. Equal numbers of cells as measured by OD_{595} were resuspended in buffer and sonicated. Equal volumes of cell extracts were subjected to SDS-PAGE and western immunoblotting using primary antibodies generated to the following *E. coli* proteins: GroEL (A), DnaK (B), and DnaJ (C). A representative immunoblot is shown. Graphs depict the mean band intensity normalized to protein concentration and relative to the wild-type signal with error bars representing the sample standard deviation (n = 3). Grey bars indicate strains with the *rpoH1* mutant background, and white bars indicate strains with the *rpoH1 rpoH2* double mutant background. The strains from left to right are Rm1021, VO3128, VO3148, VO3150, VO3151, AB35, AB36, VO3165, and VO3166.

2.4.2 Overexpression of *groESL1* and *groESL3* does not suppress the *rpoH* mutant phenotypes

In *E. coli*, expression of *groESL* from a multicopy plasmid is sufficient to suppress the temperature sensitive phenotype of the *rpoH* mutant (167). In *S. meliloti* GroEL-GroES affects NodD activity (223, 316) and *groEL1* mutants form Fix$^-$ nodules (223). I hypothesized that the defects observed for the *S. meliloti rpoH1* single mutant and *rpoH1 rpoH2* double mutant might be suppressed by overexpression of *groESL*. Therefore, I created constructs to express *groESL1* and *groESL3* independently of any possible RpoH control. I chose *groESL1* because it is highly expressed (223, Fig. 13) and the *groEL1* mutant displays a Fix$^-$ phenotype (223), and *groESL3* because it is the most divergent of the five *groE* genes. I chose the *E. coli lac* promoter and the *S. typhimurium trp* promoter because both act constitutively in *S. meliloti*, they have been successfully used to overexpress other genes in *S. meliloti* (15, 66, 86, 94), and expression from these promoters on a multicopy plasmid is stronger than expression from the endogenous *groESL1* and *groESL3* promoters (data not shown). Each construct was introduced separately into wild-type, *groEL1*, *rpoH1*, and *rpoH1 rpoH2* mutant cells.

To determine whether expression of *groESL1* or *groESL3* bypasses the symbiotic phenotypes of the *rpoH1* and *rpoH1 rpoH2* mutants, I inoculated alfalfa seedlings with wild-type and mutant bacteria containing the expression constructs. As shown in Table 3, none of the constructs altered nodulation or nitrogen fixation in the wild-type strain. Both *groESL1* constructs were able to complement the Fix$^-$ phenotype of the *groEL1* mutant. Therefore, these constructs produce active GroEL1 protein. Complementation required the *lac* promoter (data not

shown), confirming that the *groESL1* fragment does not contain the endogenous promoter. Neither *groESL1* construct was able to suppress the Fix$^-$ phenotype of the *rpoH1* mutant (Table 3). In terms of the nodulation defect of the *rpoH1 rpoH2* double mutant, P_{lac}-*groESL1* did not suppress the defect but P_{trp}-*groESL1* elicited an increase in the number of nodulated plants. However, the nodulation defect was still apparent in the low number of nodules per nodulated plant, which was similar to the *rpoH1 rpoH2* mutant. Therefore, overexpression of *groESL1* is only able to bypass partially the Nod$^-$ phenotype of the *rpoH1 rpoH2* mutant and has no effect on the Fix$^-$ phenotype of the *rpoH1* mutant.

Neither the P_{lac}-*groESL3* nor the P_{trp}-*groESL3* construct suppressed the symbiotic phenotypes of the *groEL1*, *rpoH1,* and *rpoH1 rpoH2* mutants. Since a *groESL3* mutant has no measurable phenotype (20), I could not do a genetic complementation test to prove that my *groESL3* constructs were producing active protein. Therefore, to determine whether the *groESL3* constructs were functional, I first sequenced the P_{lac}-*groESL3* construct and confirmed that no mutations were introduced during amplification of *groESL3*. I then used site-directed mutagenesis to insert codons generating a hexahistidine tag at the carboxy terminus of GroEL3. Subsequent detection using the SuperSignal® West HisProbe™ Kit (Pierce) showed that the P_{lac}-*groESL3* construct produced protein (data not shown). Therefore, my *groESL3* constructs probably produce active proteins. Thus, the results suggest that *groEL3* is not interchangeable with *groEL1* and that *groESL3* does not bypass the symbiotic phenotypes of the *rpoH1* and *rpoH1 rpoH2* mutants.

Table 3. Symbiotic phenotypes of strains overexpressing *groESL*

Strain/construct	Percent plants nodulated[a]	Nodules/nodulated plant[b]	Fixation
Wild type	98 ± 4	3.9 ± 1.9	+
Wild type/P$_{lac}$-*groESL1*	98 ± 3	3.8 ± 1.6	+
Wild type/P$_{trp}$-*groESL1*	98 ± 3	3.8 ± 2.1	+
Wild type/P$_{lac}$-*groESL3*	98 ± 3	4.0 ± 1.9	+
Wild type/P$_{trp}$-*groESL3*	100 ± 0	3.6 ± 2.0	+
groEL1	98 ± 3	3.7 ± 2.4	-
groEL1/P$_{lac}$-*groESL1*	98 ± 3	3.6 ± 1.8	+
groEL1/P$_{trp}$-*groESL1*	97 ± 6	3.6 ± 1.8	+
groEL1/P$_{lac}$-*groESL3*	90 ± 14	3.2 ± 1.8	-
groEL1/P$_{trp}$-*groESL3*	95 ± 6	3.4 ± 2.0	-
rpoH1	96 ± 4	3.7 ± 2.2	-
rpoH1/P$_{lac}$-*groESL1*	90 ± 4	3.7 ± 2.3	-
rpoH1/P$_{trp}$-*groESL1*	100 ± 0	3.7 ± 2.0	-
rpoH1/P$_{lac}$-*groESL3*	97 ± 6	3.8 ± 2.2	-
rpoH1/P$_{trp}$-*groESL3*	100 ± 0	3.1 ± 1.8	-
rpoH1 rpoH2	38 ± 15	1.8 ± 1.2	-
rpoH1 rpoH2/P$_{lac}$-*groESL1*	35 ± 13	1.4 ± 0.8	-
rpoH1 rpoH2/P$_{trp}$-*groESL1*	74 ± 23	1.8 ± 1.0	-
rpoH1 rpoH2/P$_{lac}$-*groESL3*	31 ± 30	1.3 ± 0.5	-
rpoH1 rpoH2/P$_{trp}$-*groESL3*	47 ± 15	1.7 ± 1.1	-

[a]Average and standard deviation (n ≥ 3).

[b]Average and standard deviation of total nodulated plants from all experiments (n ≥ 3).

In addition to the symbiotic phenotype, the *rpoH1* mutant displays a high temperature growth defect (Fig. 10). Given that overexpression of *groESL* bypasses the growth defect of the *rpoH* mutant in *E. coli* (167), I tested whether our *groESL1* or *groESL3* constructs could bypass the growth defect of the *rpoH1* mutant. There was no significant increase in growth of cells grown in LB/MC at 30°C or 40°C when the constructs were present (Fig. 12).

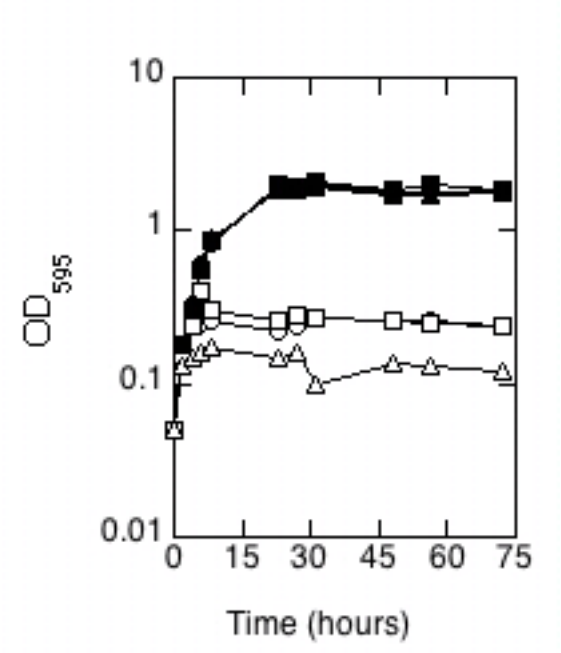

Figure 12. Comparison of the growth of wild-type or *rpoH1* cells overexpressing *groESL1* or *groESL3*, as measured using OD_{595}.

Cells were grown in LB/MC medium plus streptomycin at 40°C. The control strains are AB14 (wild type/vector, filled circles) and AB15 (*rpoH1*/vector, open circles); and the test strains are AB4 (wild type/P_{lac}-*groESL1*, filled squares), AB16 (wild type/P_{lac}-*groESL3*, filled triangles), AB7 (*rpoH1*/P_{lac}-*groESL1*, open squares), and AB17 (*rpoH1*/P_{lac}-*groESL3*, open triangles). The panels show representative data from one of three experiments.

2.4.3 Control of *groEL* gene expression by RpoH1 and RpoH2 under free-living conditions

To test if RpoH1 or RpoH2 control expression of the various *groESL* genes, I constructed a matched set of chromosomal *groEL-gus* transcriptional fusions by recombinational cloning as described in Materials and Methods. I compared expression of the *groEL-gus* fusions in wild-type, *rpoH1*, *rpoH2*, and *rpoH1 rpoH2* cells during growth in M9 sucrose medium (Fig. 13). All five *groEL-gus* fusions generated GUS activity above background levels, with *groEL1* the most highly expressed. Neither *rpoH1* nor *rpoH2* was required for expression of *groEL1*, *groEL2*, or *groEL3*. However, *rpoH1* was required for full expression of *groEL5*. Both *rpoH1* and *rpoH2* have subtle effects on *groEL4* expression. This was unexpected, given the study by Mitsui *et al.* (196), and may reflect indirect effects on gene expression. Valerie Oke obtained similar results with the *rpoH1* mutant grown in LB/MC medium, although she additionally observed a slight decrease in *groEL3* and *groEL4* expression (20).

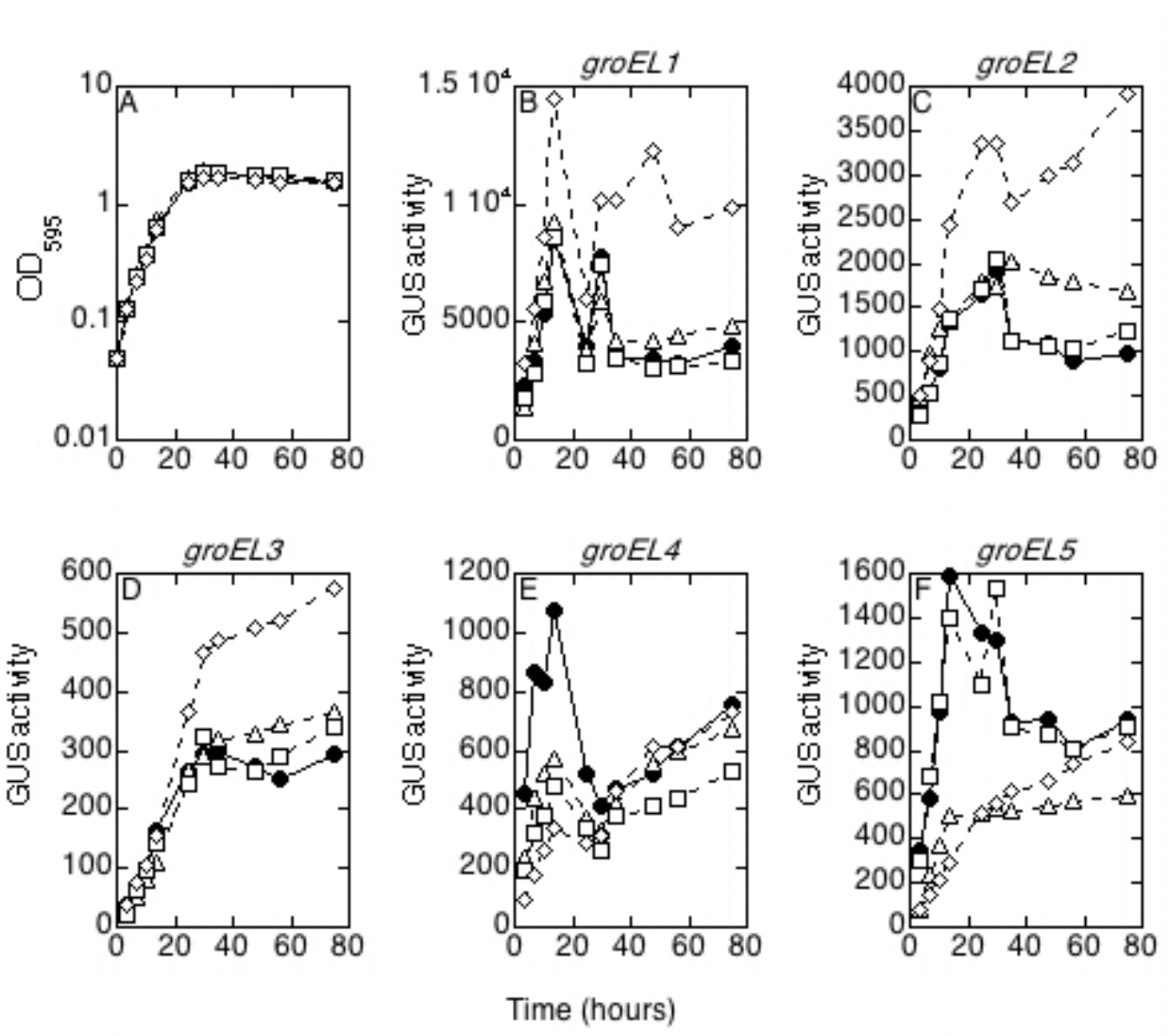
A
B
C
D
E
F
groEL1
groEL2
groEL3
groEL4
groEL5
OD595
GUS activity
Time (hours)

Figure 13. *groEL* expression in *rpoH* mutant cells grown in M9 sucrose minimal medium.

Growth as measured by OD_{595} and gene expression as monitored by β-glucuronidase (GUS) activity were determined in wild-type (filled circles), *rpoH1* (open triangles), *rpoH2* (open squares), and *rpoH1 rpoH2* (open diamonds) backgrounds. (A) Representative growth curve of cells containing the *groEL1-gus* fusion. All of the strains in the experiment showed indistinguishable growth patterns. (B-F) GUS activity of cells containing *groEL1-gus* (B), *groEL2-gus* (C), *groEL3-gus* (D), *groEL4-gus* (E), and *groEL5-gus* (F). Each panel shows the data from one representative experiment.

Although *rpoH1* is expressed within root nodules, I am not able to test directly whether RpoH1 controls expression of the *groEL* genes during symbiosis because *rpoH1* mutant cells undergo early senescence within the nodule (196). Therefore, it is possible that RpoH1 directs transcription of a different subset of these genes within the host plant.

2.5 DISCUSSION

A simple hypothesis to explain why RpoH1 is required for growth at high temperature and RpoH1 and RpoH2 are required for nodulation and nitrogen fixation during the *S. meliloti*-alfalfa symbiosis is that the transcription factors are required for the expression of one or more of the *groESL* operons and that production of GroEL-GroES is the crucial function. This hypothesis is based on two observations. First, in *E. coli groESL* is a key target of RpoH. This has been concluded because mutants that suppress the growth defect of *rpoH* overexpress *groESL,* and expression of *groESL* from a multicopy plasmid is sufficient to allow *rpoH* mutant cells to grow up to 40°C (167). Second, in *S. meliloti groEL1* mutants are delayed in nodulation and form Fix^- nodules (223). However, several lines of evidence suggest that this hypothesis is not correct. First, suppressor mutants of the high temperature growth defect and the nodulation defect do not exhibit increased production of total GroEL protein. Second, overexpression of *groESL1* or *groESL3* from constitutive promoters does not bypass the defects of the *rpoH* mutants. Third, at least under free-living conditions (heat shock in Mitsui *et al.* (196) and rich and minimal medium in this study), RpoH2 does not control any of the *groEL* genes and RpoH1 only controls expression of *groEL5*. However, *groESL5* is not required for nodulation or nitrogen fixation (196). Therefore, *groESL5* cannot be a single key target. My results suggest that the system in

S. meliloti is unlike *E. coli*, which is not surprising given the greater developmental complexity, and that there must be other crucial targets of RpoH1 and RpoH2.

What genes might be under the control of RpoH1 and RpoH2? There are at least two scenarios that could be true for either protein. First, the requirements for RpoH could solely be due to the need for properly folded proteins. The requirement during symbiosis may reflect the need to fold specific proteins induced during symbiosis and/or to respond to an increase in unfolded proteins due to stress within the nodule. The regulon would, therefore, be similar to that in *E. coli*. Second, although RpoH may direct expression of the classic HSPs, the requirement may reflect expression of other genes, perhaps specific to *Rhizobium*. For example, *rpoH2* in *Rhizobium* sp. strain TAL1145 regulates genes for exopolysaccharide synthesis, which is required for effective nodulation (162). Mitsui *et al.* (196) determined whether RpoH1 and RpoH2 control expression of nine Hsp homologs in *S. meliloti* (*groESL1-5, dnaK, clpA, clpB*, and *lon*) during heat shock. RpoH1-controlled expression of *groESL5* and partially controlled expression of *clpB* and *lon*. In contrast, RpoH2 did not control expression of any of these genes. Therefore, the regulon of RpoH1 at least partially overlaps with the regulon of RpoH in *E. coli*, but genes under the control of RpoH2 are currently unknown. Microarray experiments to determine the regulons of RpoH1 and RpoH2 in *S. meliloti* should be illuminating.

Three different phenotypes are associated with *rpoH1* and *rpoH2* in *S. meliloti*. The *rpoH1* mutant has a growth defect at high temperature and forms ineffective nodules on plants (226, 227), and the *rpoH1 rpoH2* double mutant is unable to nodulate (227). Suppressor mutant analysis suggests that the requirements for RpoH are not the same for all of the phenotypes. Spontaneous suppressor mutants have been isolated based on the ability to grow at high temperature (bypassing RpoH1) or to nodulate (bypassing RpoH1 and RpoH2). Interestingly,

regardless of how they were initially isolated, most of these mutants are able to suppress both the growth and nodulation defects. In contrast, none of the suppressor mutants are able to suppress the nitrogen fixation defect. This implies that the functions of RpoH1 during growth and RpoH1/RpoH2 during the early stages of symbiosis are similar but that there is a different or additional function of RpoH1 needed later during symbiosis. I do not know what has been altered in these suppressor mutants although I have shown that production of the GroEL and DnaK-DnaJ chaperones is not significantly altered. Analysis of the differences between the suppressor mutants and the parent strains should provide clues about the roles of RpoH1 and RpoH2 during free-living growth and symbiosis.

Although the requirements for RpoH1 and RpoH2 during symbiosis cannot be explained as a simple requirement for expression of *groESL*, the presence of multiple *groESL* genes and the connections to symbiosis make this gene family particularly interesting in the *Rhizobiaceae*. All of the nodule-forming *Rhizobium* that have been fully sequenced (*S. meliloti, B. japonicum,* and *M. loti*), as well as *R. leguminosarum,* contain multiple *groESL* genes. Although many single and double *groESL* mutants do not have symbiotic defects, some mutants do (84, 196, 222, 252). In *S. meliloti groEL1* mutants form nodules late and the nodules are Fix^- (223), and in *B. japonicum* a $groESL_3$ $groESL_4$ double mutant is unable to fix nitrogen (84). What roles do these genes play in symbiosis? In *S. meliloti* genetic and biochemical studies have demonstrated that two key regulatory proteins necessary for early gene expression during symbiosis, NodD1 and NodD3, are substrates of GroEL-GroES (223, 316). In addition, GroEL-GroES may help to form active nitrogenase later during symbiosis. In *B. japonicum*, the level of nitrogenase subunits in the $groESL_3$ mutant is greatly decreased although transcription of the genes is unaffected (84). In the free-living bacterium *Klebsiella pneumoniae,* GroEL regulates nitrogen

fixation possibly as a result of direct interactions with the regulatory protein NifA and nitrogenase subunits (116, 117). Finally, the GroEL-GroES chaperone complex may help to fold other proteins that are newly produced as the cells adapt and differentiate within the plant host.

Currently there is no clear reason why multiple *groESL* genes are present in these genomes. One possibility is that the genes are simply regulated differentially, providing GroES and GroEL under different conditions. Evidence for differential gene expression has been obtained in *S. meliloti* (196), *B. japonicum* (10, 83), and *R. leguminosarum* (252). Specifically in *S. meliloti,* only *groESL1* and *groESL5* are induced by heat shock (196), only *groESL5* is controlled by RpoH1 (196, this study), and only *groESL1* and *groESL2* are preceded by a CIRCE element that may indicate regulation by the HrcA repressor, which is used to regulate heat-inducible genes in some bacteria (210). An additional possibility is that the encoded chaperones have different ranges of substrates. Although the GroEL-GroES complex can assist in the folding of a wide variety of proteins, it cannot function universally. Directed evolution studies have demonstrated that small numbers of amino acid changes in GroES and GroEL can lead to shifts in the spectrum of substrates (303). Therefore, multiple *groESL* genes may allow the cell to fold a wider variety of proteins. As an extreme example, bacteriophage T4 encodes a protein of little sequence similarity to GroES that nevertheless substitutes for the host GroES, generating a new chaperone complex that can fold the major capsid protein (11, 295). In *R. leguminosarum*, the three GroEL proteins have different *in vitro* properties, including the ability to refold a specific denatured substrate (103). We found that *groEL3* is not interchangeable with *groEL1*, which would be consistent with different substrate specificities, whereas *groEL2* is interchangeable with *groEL1* (223), suggesting at least overlapping substrate specificities for that pair. As an added complexity, heteromeric complexes, as well as homomeric complexes, might

be made, which would dramatically increase the number of different types of GroEL-GroES chaperone species within the cells.

2.6 CONCLUSIONS

While many bacteria maintain one copy of the *rpoH* gene, some members of the α-proteobacteria, including several root-nodulating rhizobia, maintain multiple copies. *S. meliloti* maintains two copies of *rpoH*, and mutations in these genes are associated with symbiotic phenotypes (196, 226, 227). Because *groESL* is a key target in *E. coli* and a mutation in *groEL1* in *S. meliloti* results in a symbiotic phenotype, I tested whether the primary function of RpoH1 and RpoH2 during symbiosis is to provide sufficient transcription of *groESL*. The results of suppressor mutant analysis and overexpression of *groESL* indicate that this is not the case and that there must be other targets of *rpoH* that are necessary for symbiosis.

The next step toward determining the functions of RpoH1 and RpoH2 will be to determine their regulons. Sharon Long's laboratory at Stanford University has designed the *S. meliloti* Affymetrix GeneChip and optimized the protocols (16). In the following chapter, I used microarray analysis to determine the gene targets of RpoH1 and RpoH2 in *rpoH* mutant cells during stationary phase.

3.0 DETERMINING THE GENE TARGETS OF RPOH1 AND RPOH2 USING MICROARRAY ANALYSIS

3.1 INTRODUCTION

In bacteria, sigma factors are the subunits of RNA polymerase that recognize the promoters of genes. Typically, bacteria maintain a housekeeping sigma factor to transcribe genes essential for basic cellular processes and several alternative sigma factors to direct transcription of genes in response to changing environmental conditions. RpoH (σ^{32}) is an alternative sigma factor that was first identified in *E. coli* during response to heat stress (125, 169, 216, 314). When cells are challenged by heat or other stresses, RpoH levels rise transiently to direct transcription of heat shock genes, which encode proteins that function in protein folding and degradation, among other processes (321). In *E. coli*, the basal level of transcription of these genes directed by RpoH is also important under non-stress conditions (324).

Many members of the RpoH regulon have been identified through the analysis of individual or several genes, but these approaches were necessarily limited by scale. Recent studies have sought to determine the RpoH regulons in various bacterial species through global expression analyses. Two studies have determined the RpoH regulon in *E. coli* by inducing expression of *rpoH* under normal growth conditions and monitoring the resulting changes in gene expression over time (221, 323). As expected, overexpression of *rpoH* up-regulates the

classical heat shock genes, but many new targets were also identified. Interestingly, Nonaka *et al.* showed that 25% of the RpoH gene targets encode proteins that are associated with the cytoplasmic membrane, suggesting that the RpoH-mediated response to heat shock is involved in membrane homeostasis (221). Additionally, they found roles for RpoH targets in DNA and RNA protection, in transcriptional and translational processes, and in regulating transcription in response to environmental conditions.

RpoH regulons have also been determined in the pathogens *Neisseria gonorrhoeae* (128) and *Vibrio cholerae* (273). In *N. gonorrhoeae*, RpoH regulates stress response genes (128) but also appears to regulate genes in response to epithelial cell contact (61). In *V. cholerae*, the RpoH regulon is similar to the *E. coli* regulon, with many of the encoded products functioning in protein maintenance (273).

All of the above analyses were performed in bacteria that maintain a single *rpoH* gene, but certain members of the α-proteobacteria have multiple copies of *rpoH*. There is no single unifying characteristic for the 11 species that have multiple *rpoH* genes, although several are nitrogen-fixing bacteria (*Rhizobium* species, *Rhodobacter sphaeroides,* and *Rhodospirillum rubrum*) and others interact with eukaryotic hosts (*Rhizobium* species, two *Bartonella* species, and *Brucella melitensis*). Why might these bacteria maintain multiple copies of *rpoH*? One possibility is that the transcription factors are regulated differentially but control similar regulons. This would allow the bacteria to fine-tune their responses to different environmental conditions. Alternatively, the *rpoH* genes might encode transcription factors that control different regulons and therefore direct transcription of genes needed for different functions.

There is evidence that multiple *rpoH* genes are differentially regulated. In *Bradyrhizobium japonicum*, the $rpoH_1$ and $rpoH_2$ genes have σ^{70}-like promoter sequences,

although *rpoH$_1$* is up-regulated in response to heat shock (213, 214). The *B. japonicum rpoH$_3$* gene, in contrast, is organized within an operon with two other genes under control of a putative σ^{32} promoter, suggesting possible up-regulation in response to heat shock (213). The two *rpoH* homologs in the nitrogen-fixing symbiont *Sinorhizobium meliloti* are regulated such that *rpoH1* is expressed strongly during growth, stationary phase and symbiosis, while *rpoH2* is expressed only during late stationary phase in minimal media (226, 261). In *Rhodobacter sphaeroides*, the levels of $RpoH_I$ greatly exceed the levels of $RpoH_{II}$ under aerobic growth conditions (121), and only the gene for $RpoH_{II}$ is regulated by the extracytoplasmic stress response sigma factor RpoE (5).

There is also evidence that multiple RpoH proteins in the same organism could recognize different promoters. In *B. japonicum*, the three *rpoH* genes complement an *E. coli rpoH* mutant to different degrees (213), and $RpoH_1$ and $RpoH_2$ recognize the *dnaKJ* and *groESL* promoters *in vitro* with different efficiencies (212). Analysis of the *S. meliloti* single and double *rpoH* mutants suggests that the RpoH1 and RpoH2 regulons overlap (20, 227). However, Mitsui *et al.* (196) demonstrated that RpoH1 controls expression of three classical heat shock genes, *groESL5*, *lon*, and *clpB*, while RpoH2 does not. The $RpoH_I$ and $RpoH_{II}$ factors from *R. sphaeroides* recognize an overlapping, but not identical, set of promoters *in vitro*, although both are able to complement an *E. coli rpoH* mutant (121, 160). The *B. melitensis* 16 M *rpoH2* mutant is sensitive to heat and oxidative stress and displays attenuated virulence, and the *rpoH1* mutant behaves like the wild-type strain (55), which indicates that either the RpoHs control different targets or are regulated differentially. Currently, the evidence suggests that in most organisms studied the regulons at least partially overlap, while the *rpoH* genes are regulated differentially such that they might be specialized for responses to different environmental stresses.

Interestingly, several *rpoH* genes in the *Rhizobium* have been connected to symbiosis and nitrogen fixation. An *S. meliloti rpoH1* mutant is unable to fix nitrogen, and an *rpoH1 rpoH2* double mutant cannot form nodules on plant roots (226, 227). A mutation in the *Rhizobium* sp. strain TAL1145 *rpoH2*, named after its similarity to other *rpoH2* genes, results in a strain that nodulates less effectively (162). This RpoH also regulates expression of the genes necessary for exopolysaccharide synthesis, which has not been shown for any other RpoH. The regulon of RpoH1 in *S. meliloti* has been shown to overlap partially with the *E. coli* regulon because it directs transcription of *groESL5*, *lon*, and *clpB*, but not *dnaK* and *clpA* (196). However, given that the *Rhizobium* sp. strain TAL1145 RpoH2 protein directs exopolysaccharide synthesis, unlike in *E. coli*, the possibility exists that *S. meliloti*, too, has co-opted these stress response sigma factors for organism-specific functions, like symbiosis.

In Chapter 2 I showed that the requirements for *rpoH* during symbiosis cannot be explained solely by a loss of GroEL protein and that there must be other or additional crucial targets (20). To find these targets, I am using microarray experiments to define the RpoH1 and RpoH2 regulons in *S. meliloti*. In the first experiment, I have compared global gene expression among wild-type and *rpoH* mutant cells in late stationary phase, a condition under which there is evidence that both *rpoH* genes are expressed (20, 226). In this experiment, expression of RpoH targets should be decreased in *rpoH* mutants relative to the wild-type strain. The results from the microarray analysis are described below.

3.2 MATERIALS AND METHODS

3.2.1 Strains and growth conditions

Strains used in this study are Rm1021 (wild type), VO3128 (*rpoH1::aadA*), AB3 (*rpoH2::aacCI*), and AB9 (*rpoH1::aadA rpoH2::aacCI*), which have been previously published (20, 188, 226). Bacterial cultures were grown at 30°C in LB/MC medium (109) or M9 minimal medium containing 0.2% sucrose, 0.5 μg biotin ml^{-1}, 1 mM $MgSO_4$, and 0.25 mM $CaCl_2$. Streptomycin was added to the media at 500 μg ml^{-1}.

3.2.2 Sample preparation, cell lysis and RNA isolation

Three replicates for this experiment were started on the same day using the same medium. Overnight cultures were grown in LB/MC medium, diluted to OD_{595} = 0.05 the next day, and allowed to grow overnight to ensure even growth. 8 ml of cells were washed twice and diluted to an OD_{595} of 0.05 in 300 ml M9 minimal medium. Growth of the three replicates was staggered by 1.5 hours. Cultures were incubated with shaking for 48 hours, until the late stationary phase of growth. Cells were harvested by mixing 99 ml of culture with 11 ml of ice-cold stop solution (5% buffer-equilibrated phenol in ethanol) and centrifuging 5,000 rpm at 4°C. The supernatant was removed, and cells were frozen in liquid nitrogen and stored at -80°C.

Cell pellets were resuspended in 1 mg ml^{-1} lysozyme in TE and buffer RLT (Qiagen RNeasy Kit), and lysed by bead-beating with 0.09-0.135 mm glass beads (Thomas Scientific). The bead-beating procedure was performed at 4°C with three 30-second pulses and a one-minute incubation on ice between each pulse. Total RNA was isolated as described (16), except that an

RNeasy Midi Kit (Qiagen) was used. This protocol includes both on-column and off-column DNase digests to remove all contaminating chromosomal DNA. Absence of chromosomal DNA was confirmed by PCR amplification using primers to an intergenic region, and RNA integrity was validated on a 1.2% agarose formaldehyde gel. Over 100 ug of RNA was isolated for each sample. RNA for all three replicates was shipped on dry ice to Sharon Long's laboratory at Stanford University for microarray analysis.

3.2.3 cDNA synthesis, labeling, and Affymetrix GeneChip hybridization

Carol Toman (Stanford University) performed the cDNA synthesis, labeling and hybridization as described (16). Briefly, the GeneChip *Pseudomonas aeruginosa* Genome Array Expression Analysis protocol (Affymetrix, Santa Clara, CA) was used for first-strand cDNA synthesis. 12 μg of fragmented labeled cDNA was hybridized to GeneChips in a GeneChip Hybridization Oven 640, and the arrays were washed and stained according to the Affymetrix protocols.

3.2.4 Data analysis using Affymetrix software

Melanie Barnett (Stanford University) performed the data analysis using MICROARRAY SUITE Ver. 5, MICRODB Version 3, and DATA MINING TOOL Version 3 (Affymetrix) as described (16). Briefly, the GeneChips were globally scaled to a target signal intensity of 500. Experimental arrays were compared to baseline arrays in comparison expression analysis, such that each array for a mutant strain was compared to the wild-type arrays for each replicate. For example, the signal for the VO3128 replicate 1 chip was compared to the signals from all three wild-type chips. This yielded nine comparisons for each mutant strain. A decrease of average signal log

ratio ≥ 1 (corresponding to a 2-fold change) was considered significant if the software found the pairwise comparisons to be significantly different ($P \leq 0.05$). Of those genes designated significant, I have listed only the genes that showed a decrease in 8-9 of 9 comparisons.

3.2.5 Data analysis using Significance Analysis of Microarrays (SAM)

SAM is a Microsoft Excel-based program that identifies significantly changed gene expression (http://www-stat.stanford.edu/~tibs/SAM/) (293). I entered the signal log ratios for each replicate into the SAM program as a "one class" experiment comparing wild-type to *rpoH* mutant cells. The analysis was performed once for each comparison: wild type versus *rpoH1*, wild type versus *rpoH2*, and wild type versus *rpoH1 rpoH2*. I set the Delta value, which determines the number of significant genes and calculates the false positive rate for the set, at 0.75. A q-value is reported for each gene, which describes the chance that the target is a false positive.

3.3 RESULTS

3.3.1 Isolating sufficient RNA

Previous studies using transcriptional reporter fusions demonstrated that both *rpoH1* and *rpoH2* were expressed during stationary phase, with *rpoH1* induced early in stationary phase and *rpoH2* induced later in stationary phase (20, 226). I chose to compare global gene expression in wild-type and *rpoH* mutant cells after 48 hours of growth in minimal medium when it was likely that

both RpoH1 and RpoH2 would be active, which would allow me to determine the RpoH1 and RpoH2 regulons simultaneously.

An important step in microarray analysis is obtaining RNA. To lyse cells, Melanie Barnett and I first used an enzymatic procedure (cells resuspended in 1 mg ml^{-1} lysozyme) that had been used for microarray analysis of *S. meliloti* cells that were in the exponential or early stationary phases of growth (16), but we discovered that late stationary phase *S. meliloti* cells were resistant to the enzymatic lysis step. To overcome this difficulty, I tested two mechanical lysis methods—sonication and bead-beating—for the ability to efficiently lyse stationary phase cells. I determined that the bead-beating procedure yielded the largest quantity of RNA (data not shown), suggesting that it most effectively lysed stationary phase cells. To confirm that the RNA isolated after bead-beating was suitable for microarray analysis, I isolated 40 μg of RNA from log phase cells grown in LB/MC and from stationary phase cells grown in M9 minimal for test microarrays.

Carol Toman performed the hybridization to the Affymetrix GeneChips using 4 μg of cDNA, which was previously successful using RNA from cells in the exponential and early stationary phases of growth. However, when Melanie Barnett compared the signals from the log phase and stationary phase samples, she found that the signal from stationary phase hybridization was too low for a meaningful comparison (data not shown). Ribosomal RNA was overrepresented on the chips, suggesting that mRNA levels in stationary phase cells are quite low. Therefore, Melanie Barnett and Carol Toman decided that we needed 12 μg of cDNA from the late stationary phase cells for chip hybridization. Synthesis of such a large quantity of cDNA requires at least 100 μg of RNA. Therefore, I harvested nearly 100 ml of cells (instead of 16 ml) and increased the scale of the experiment more than 10-fold using the RNeasy Midi Kit. When

12 μg of cDNA from the late stationary phase cells was used for hybridization, the total signal from the chips was adequate for analysis, and the background signal was within the range from experiments using less cDNA and was therefore acceptable.

3.3.2 RpoH1 and RpoH2 targets identified during late stationary phase

I obtained RNA from wild-type and *rpoH* mutant cells in late stationary phase, and Carol Toman performed the cDNA synthesis and microarray hybridization. Melanie Barnett performed the initial analysis of the microarray data using the Affymetrix software. In this analysis, experimental arrays were compared to baseline arrays in comparison expression analysis, such that each array for a mutant strain was compared to the wild-type arrays from all replicates. The significance of the change in gene expression was determined for each gene in nine total comparisons, and the amount of the change in gene expression was expressed as a signal log ratio (SLR). I compiled lists of genes that exhibit significant differences in gene expression and that show a two-fold decrease or greater (SLR ≥ 1) in wild-type versus *rpoH1* cells (Table 4), wild-type versus *rpoH2* cells (Table 5), and wild-type versus *rpoH1 rpoH2* cells (Table 6). Strikingly, expression of far fewer genes was identified as significantly decreased at a two-fold cutoff in the *rpoH1* mutant (9 genes) than in the *rpoH2* (46 genes) or *rpoH1 rpoH2* (85 genes) mutants. Of the 85 targets whose expression was decreased in the *rpoH1 rpoH2* mutant, 55 show a greater fold-decrease in the double mutant relative to either single mutant or appear in the double mutant only, suggesting that these targets might be regulated by both RpoH1 and RpoH2. Although I am primarily concerned with genes that have decreased expression in *rpoH* mutant cells compared to wild-type cells because I expect direct targets to be transcribed less frequently in mutant cells, I have also listed genes that have increased expression in *rpoH* mutant cells

(Table 7). Interestingly, these results indicate that RpoH2 might negatively regulate the *rpoH2* gene.

To obtain the most conservative number of likely RpoH targets, I performed additional analysis of the microarray data using SAM (Significance Analysis of Microarrays), which is a program that identifies significantly changed gene expression by performing gene-specific *t* tests (293). The program can estimate the false discovery rate (FDR) for the data set through permutations of the *t* scores, and the FDR cutoff is set by the user. This cutoff, or Delta value, defines the number of false positives that are acceptable for a given data set. I performed data analysis with SAM using a Delta of 0.75 so that the program would identify genes whose q-value was no greater than 51%, meaning that there was no more than a 51% chance the target was a false positive. The q-values for significant genes identified by SAM are listed in Tables 4-6. Of the 88 unique putative targets identified by the Affymetrix software, only 34 were considered significant by SAM, and the majority of these were only significant in the wild type versus *rpoH1 rpoH2* comparison.

Table 4. Genes with significantly decreased expression in *rpoH1* cells versus wild-type cells during stationary phase

Name[a]	Description	Signal Log Ratio[b]	Standard Deviation	Fold Decrease[c]	SAM q-value (%)[d]
SMb20303	Hypothetical	-1.17	0.27	2.2	
SMb20551	Hypothetical	-1.47	0.30	2.8	
SMb21379	Conserved hypothetical	-1.10	0.37	2.1	
SMc00949	Conserved hypothetical	-1.55	0.61	2.9	
SMc01030	Pyruvate dehydrogenase α_2 subunit	-1.04	0.36	2.1	
SMc02769	Conserved hypothetical transmembrane	-3.20	0.23	9.1	0
SMc02900*	Conserved hypothetical	-1.05	0.22	2.1	
SMc03794	Hypothetical transmembrane	-5.57	0.75	47	22
SMc04040	Probable heat shock (IbpA)	-2.02	0.30	4.0	22

[a] Targets possibly regulated by both RpoH1 and RpoH2 are underlined. These genes demonstrate an increased signal log ratio in the *rpoH1 rpoH2* mutant compared to the *rpoH1* mutant, suggesting they may also be regulated by RpoH2.

[b] Average from three biological replicates.

[c] Fold decrease was calculated from the average signal log ratio.

[d] Only genes found to be significant using SAM at Delta = 0.75 have a q-value.

* Significantly decreased in *rpoH1*, *rpoH2*, and *rpoH1 rpoH2* mutant cells.

Table 5. Genes with significantly decreased expression in *rpoH2* cells versus wild-type cells during stationary phase

Name[a]	Description	Signal Log Ratio[b]	Standard Deviation	Fold Decrease[c]	SAM q-value (%)[d]
SMa0136	Hypothetical	-2.38	0.80	5.2	
SMa1158	Conserved hypothetical	-1.61	0.60	3.0	
SMa1364	Putative ABC transporter, periplasmic solute-binding	-2.42	0.82	5.3	
SMa2061	Conserved hypothetical	-3.04	0.76	8.2	
SMa2301	Putative response regulator	-4.08	0.60	17	46
SMa2349	Probable oxidoreductase	-3.78	1.10	14	
SMa2351	Possible oxidoreductase, molybdopterin-binding subunit	-2.66	1.03	6.3	
SMa2353	Probable oxidoreductase	-2.46	1.00	5.5	
SMb20116	Conserved hypothetical	-2.24	0.54	4.7	
SMb20117	Hypothetical sugar transferase	-2.86	0.46	7.2	51
SMb20302	Conserved hypothetical	-5.41	1.76	42	
SMb20361	Putative ionic voltage-gated channel	-1.96	0.78	3.9	
SMb20575	Putative 3-carboxy-cis,cis-muconate cycloisomerase	-1.65	0.52	3.1	
SMb20590	Hypothetical	-1.28	0.60	2.4	
SMb21683	Hypothetical	-4.29	0.51	19	0
SMc00030	Hypothetical signal peptide	-2.05	0.57	4.1	
SMc00031	Hypothetical transmembrane	-1.34	0.32	2.5	
SMc00048	Conserved hypothetical	-1.14	0.37	2.2	
SMc00049	Conserved hypothetical	-1.18	0.37	2.3	
SMc00106	Conserved hypothetical	-3.88	0.56	15	46
SMc00110	Probable glucose dehydrogenase (pyrroloquinoline-quinone)	-1.62	0.54	3.1	
SMc00367	Conserved hypothetical	-2.96	0.38	7.7	46
SMc00814	Hypothetical signal peptide	-2.94	0.45	7.6	0
SMc00952	Conserved hypothetical	-1.43	0.49	2.7	
SMc00969	Conserved hypothetical	-2.09	0.83	4.2	
SMc00970	Putative exodeoxyribonuclease	-1.40	0.52	2.6	

SMc01723	Hypothetical transmembrane	-3.89	0.99	15	51
SMc01757iSMc01758f1	Intergenic spacer	-1.42	0.36	2.7	
SMc01758	60 KD chaperonin B (GroEL4)	-1.76	0.55	3.4	
SMc01759	Conserved hypothetical	-1.93	0.54	3.8	
SMc01960	Putative oxidoreductase	-1.81	0.42	3.5	
SMc01961iSMc01962f1	Intergenic spacer	-1.31	0.50	2.5	
SMc02832	Putative periplasmic binding	-1.05	0.41	2.1	
SMc02900*	Conserved hypothetical	-1.43	0.54	2.7	
SMc03176	Hypothetical	-1.65	0.69	3.1	
SMc03246	Putative integrase DNA	-1.11	0.53	2.2	
SMc03246iSMc03247f1	Intergenic spacer	-1.04	0.39	2.1	
SMc03802	Conserved hypothetical	-2.62	0.60	6.1	51
SMc03803	Conserved hypothetical	-2.56	0.69	5.9	51
SMc03836	Putative acyl-CoA thioesterase I	-1.94	0.66	3.8	
SMc03873iSMc03874f2	Intergenic spacer	-6.65	0.82	99	0
SMc04146	Conserved hypothetical	-3.23	0.46	9.3	46
SMc04181	Putative transmembrane	-1.42	0.43	2.7	
SMc04202	Putative transmembrane	-1.25	0.49	2.4	
SMc04202iSMc04203f1	Intergenic spacer	-1.18	0.45	2.3	
SMc04334iSMc04335f1	Intergenic spacer	-2.19	0.56	4.5	

[a] Targets possibly regulated by both RpoH1 and RpoH2 are underlined. These genes demonstrate an increased signal log ratio in the *rpoH1 rpoH2* mutant compared to the *rpoH2* mutant, suggesting they may also be regulated by RpoH1.

[b] Average from three biological replicates.

[c] Fold decrease was calculated from the average signal log ratio.

[d] Only genes found to be significant using SAM at Delta = 0.75 have a q-value.

* Significantly decreased in *rpoH1*, *rpoH2*, and *rpoH1 rpoH2* mutant cells.

Table 6. Genes with significantly decreased expression in *rpoH1 rpoH2* cells versus wild-type cells during stationary phase

Name[a]	Description**	Signal Log Ratio[b]	Standard Deviation	Fold Decrease[c]	SAM q-value (%)[d]
SMa0136	Hypothetical	-2.97	0.98	7.8	
SMa0436iSMa0439f3	Intergenic spacer	-2.41	0.43	5.3	
SMa1158	Conserved hypothetical	-1.64	0.54	3.1	
SMa1364	Putative ABC transporter, periplasmic solute-binding	-3.23	0.84	9.3	24
SMa2061	Conserved hypothetical	-2.80	0.76	6.9	22
SMa2063	Hypothetical	-2.82	1.53	7.0	35
SMa2301	Putative response regulator	-4.60	0.66	24	22
SMa2349	Probable oxidoreductase**	-2.87	0.34	7.3	22
SMa2351	Possible oxidoreductase, molybdopterin-binding subunit**	-2.61	1.36	6.1	
SMa2353	Probable oxidoreductase**	-1.62	0.55	3.1	29
SMb20116	Conserved hypothetical	-1.84	0.56	3.6	
SMb20117	Hypothetical sugar transferase	-3.96	0.81	15	10
SMb20302	Conserved hypothetical	-5.58	1.55	47	
SMb20303	Hypothetical	-1.67	0.35	3.2	22
SMb20361	Putative ionic voltage-gated channel	-3.53	1.46	11	
SMb20551	Hypothetical	-2.17	0.41	4.5	29
SMb20575	Putative 3-carboxy-cis,cis-muconate cycloisomerase	-1.61	0.46	3.0	
SMb20589iSMb20590f1	Intergenic spacer	-1.28	0.65	2.4	
SMb20590	Hypothetical	-1.41	0.65	2.6	
SMb21028	Conserved hypothetical	-1.45	0.37	2.7	
SMb21295	Putative small heat shock, hsp20**	-2.72	0.78	6.5	
SMb21296	Hypothetical	-2.19	0.45	4.5	22
SMb21379	Conserved hypothetical	-2.53	0.34	5.7	22
SMb21683	Hypothetical	-4.55	0.48	23	22
SMc00030	Hypothetical signal peptide	-1.89	0.49	3.7	

SMc00043	Probable superoxide dismutase Fe**	-1.04	0.33	2.1	
SMc00048	Conserved hypothetical	-1.13	0.25	2.2	
SMc00049	Conserved hypothetical	-1.08	0.23	2.1	
SMc00106	Conserved hypothetical	-3.33	0.53	10	22
SMc00110	Probable glucose dehydrogenase (pyrroloquinoline-quinone)	-1.68	0.51	3.2	
SMc00366iSMc00367f1	Intergenic spacer	-2.18	0.36	4.5	24
SMc00367	Conserved hypothetical	-4.60	0.60	24	10
SMc00469iSMc00468f1	Intergenic spacer	-1.55	0.33	2.9	
SMc00814	Hypothetical signal peptide	-2.95	0.29	7.7	22
SMc00876	Putative MRP ATP-binding	-1.15	0.30	2.2	
SMc00949	Conserved hypothetical	-3.29	0.45	9.7	22
SMc00952	Conserved hypothetical	-2.34	0.55	5.0	
SMc00969	Conserved hypothetical	-2.41	0.70	5.3	
SMc00970	Putative exodeoxyribonuclease	-1.59	0.60	3.0	
SMc01030	Pyruvate dehydrogenase α_2 subunit	-1.21	0.39	2.3	
SMc01031	Pyruvate dehydrogenase β_2 subunit	-1.11	0.41	2.2	
SMc01033	Probable arylesterase	-1.15	0.56	2.2	
SMc01035	Probable dihydrolipoamide dehydrogenase	-1.21	0.26	2.3	37
SMc01180	Conserved hypothetical transmembrane	-1.27	0.84	2.4	
SMc01280	Probable protease**	-1.48	0.53	2.8	
SMc01723	Hypothetical transmembrane	-4.64	1.11	25	22
SMc01757iSMc01758f1	Intergenic spacer	-1.07	0.24	2.1	
SMc01758	60 KD chaperonin B (GroEL4)**	-1.64	0.32	3.1	
SMc01759	Conserved hypothetical	-1.80	0.34	3.5	
SMc01959iSMc01960f1	Intergenic spacer	-1.59	0.50	3.0	
SMc01960	Putative oxidoreductase**	-1.76	0.37	3.4	29
SMc01961iSMc01962f1	Intergenic spacer	-1.99	0.5	4.0	22
SMc02051	Conserved hypothetical	-1.46	0.27	2.7	
SMc02052	Conserved hypothetical	-1.48	0.38	2.8	
SMc02382	Conserved hypothetical	-2.50	1.16	5.6	
SMc02390	Putative glutathione s-transferase**	-2.49	1.12	5.6	

SMc02433	Probable ATP-dependent protease (heat shock) (ClpB)**	-1.90	0.42	3.7	22
SMc02443	Probable glutaredoxin 3	-2.64	1.06	6.2	
SMc02558	Conserved hypothetical	-1.24	0.73	2.4	
SMc02562	Phosphoenolpyruvate carboxykinase	-1.10	0.66	2.1	
SMc02575	Probable heat shock (HslV)**	-1.28	0.53	2.4	
SMc02656	Hypothetical	-1.74	0.53	3.3	37
SMc02703	Conserved hypothetical	-1.26	0.37	2.4	
SMc02769	Conserved hypothetical transmembrane	-3.93	0.27	15	10
SMc02832	Putative periplasmic binding	-1.08	0.46	2.1	
SMc02900	Conserved hypothetical	-3.72	0.39	13	22
SMc03246	Putative integrase DNA	-1.52	0.45	2.9	
SMc03246iSMc03247f1	Intergenic spacer	-1.27	0.50	2.4	
SMc03789	Hypothetical	-1.00	0.38	2.0	
SMc03794	Hypothetical transmembrane	-4.55	0.35	23	10
SMc03802	Conserved hypothetical	-3.22	0.69	9.2	22
SMc03803	Conserved hypothetical	-3.61	1.19	12	
SMc03836	Putative acyl-CoA thioesterase I	-2.47	0.90	5.5	35
SMc03857	Probable signal recognition particle	-1.98	0.55	3.9	
SMc03858	Putative chorismate mutase	-1.05	0.50	2.1	
SMc03859	Probable 30S ribosomal S16	-1.40	0.45	2.6	22
SMc03873iSMc03874f2	Intergenic spacer	-6.06	0.97	66	22
SMc04040	Probable heat shock (IbpA)**	-2.61	0.38	6.1	29
SMc04092	Hypothetical	-1.17	0.43	2.2	
SMc04146	Conserved hypothetical	-2.87	0.52	7.3	22
SMc04202	Putative transmembrane	-2.82	0.83	7.0	
SMc04202iSMc04203f1	Intergenic spacer	-2.62	0.54	6.1	22
SMc04305iSMc04306f2	Intergenic spacer	-1.12	0.44	2.2	
SMc04334iSMc04335f1	Intergenic spacer	-2.00	0.56	4.0	
SMc04406	Hypothetical transmembrane	-3.09	1.60	8.5	

[a] Targets possibly regulated by both RpoH1 and RpoH2 are underlined. These genes do not demonstrate a significant decrease in either single mutant, suggesting that they may be regulated by both RpoH1 and RpoH2.

[b] Average from three biological replicates.

[c] Fold decrease was calculated from the average signal log ratio.

[d] Only genes found to be significant using SAM at Delta = 0.75 have a q-value.

**Possible stress response proteins.

Table 7. Genes with significantly increased expression in *rpoH* mutant cells versus wild-type cells in stationary phase

***rpoH* mutant**	**Name**	**Description**	**Signal Log Ratio[a]**	**Standard Deviation**	**Fold Increase[b]**
rpoH1	SMb21094	Probable argininosuccinate lyase	1.47	1.09	2.8
	SMb21097	Putative amino acid uptake ABC transporter periplasmic solute-binding precursor	1.37	1.04	2.6
	SMc00091	Putative sulfate adenylate transferase subunit 2 cysteine biosynthesis	1.21	0.98	2.3
rpoH2	SMa0933	Probable TraC conjugal transfer	1.89	0.59	3.7
	SMc01430	Probable acetolactate synthase isozyme III small subunit	1.07	0.23	2.1
	SMc01431	Probable acetolactate synthase isozyme III large subunit	1.04	0.29	2.1
	SMc03873	Putative RNA polymerase sigma factor (RpoH2)	2.55	0.61	5.8
rpoH1 rpoH2	SMa0320	Putative	1.31	0.80	2.5
	SMa0930	Probable TraD conjugal transfer	3.16	0.57	8.9
	SMa0933	Probable TraC conjugal transfer	2.12	0.45	4.3
	SMa2297	Hypothetical	1.21	0.32	2.3
	SMc01609	Putative 6,7-dimethyl-8-ribityllumazine synthase	1.00	0.50	2.0
	SMc03873	Putative RNA polymerase sigma factor (RpoH2)	2.71	0.62	6.5
	SMc04153	Putative aminomethyltransferase	1.08	0.38	2.1

[a] Average from three biological replicates.

[b] Fold decrease was calculated from the average signal log ratio.

Many of these genes encode proteins involved in cell processes and metabolism, according to the *S. meliloti* genome annotation. Of these, twelve are putative stress response proteins, encoding proteases, chaperones and proteins involved in resistance to oxidative stress (Table 6). The majority of targets identified encode hypothetical or conserved hypothetical proteins (43 of 88). 12 intergenic regions were also significantly down-regulated. Although these targets might represent long leader sequences, it is possible that some might be small regulatory RNAs or unannotated open reading frames.

When I compared my potential RpoH targets in *S. meliloti* to the *E. coli* RpoH regulon (221), I identified four genes in common: *clpB*, *groEL* (*groEL4*), *hslV*, and *ibpA*. *clpB*, *hslV*, and *ibpA* could be regulated by both RpoH1 and RpoH2, while *groEL4* was only identified as a putative RpoH2 target. *clpB* and *ibpA* were significant in both the Affymetrix and SAM analyses. Previous studies had identified *groESL5* (20, 227), *clpB* (196), and *lon* as targets of RpoH1. From this stationary phase microarray experiment, only *clpB* was identified, and it could be co-regulated by RpoH1 and RpoH2.

A previous study defined a consensus sequence for RpoH1 in *S. meliloti* (cnCTTgAA-N_{17}-CCAnaT) using the promoters of *groESL5*, *clpB*, and *lon* (196). I searched for this sequence in the 400 bp upstream of all genes that were considered significantly decreased in both computer analyses using the program DNA-PATTERN (296), and the results appear in Table 8. A putative RpoH1 binding site was identified upstream of seven of 34 genes, which included *clpB* and *ibpA*. All seven genes were identified as either regulated by RpoH1 or by RpoH1 and RpoH2. I also searched for the RpoH consensus sequence defined for α-purple proteobacteria (CTTG-N_{17}-CyTATnT) (267), although this consensus sequence was determined from promoters that are not confirmed RpoH-controlled genes. In addition to the genes identified with the

RpoH1 sequence, 12 more genes have a putative upstream RpoH consensus sequence (Table 9). These genes include both RpoH2-specific genes and genes that could be regulated by both RpoH1 and RpoH2.

Table 8. Potential RpoH1 promoter sequences upstream of putative RpoH targets

Gene	Regulated by RpoH1 or RpoH2	Sequence[a]
SMb20117	1, 2	CTCTTGAAGAGCACGGCGCATCGACTAGAT*
SMb21296	1, 2	TTCTTGAATTTACCTTTTCCGCTCCAATAT**
SMc00949	1, 2	CTCTTGATGTTCCATATGCGACAACCCAGAT*
SMc02433 (*clpB*)	1, 2	CTCTTTAATTCAGAAGTGCGCTGCCCCATAT*
SMc03794	1	CTCTTGCAAGCGGCGTGGCCAGCACCCACAT*
SMc03802	1, 2	CCCTTGAGATTTCGAAGGGCCGCACCATTT**
SMc04040 (*ibpA*)	1, 2	CTCTTGAACTCGTGCGCGGGCATTCCCATGT*

[a] Consensus RpoH1 sequence: CNCTTGAA (N_{16-17}) CCANAT (196)

*One mismatch to consensus.

**Two mismatches to consensus.

Table 9. Potential α-proteobacteria RpoH promoter sequences upstream of putative RpoH targets

Gene	Regulated by RpoH1 or RpoH2	Sequence[a]
SMa1364	1, 2	CTCGACGTTTCCGCAGACAGAGCTATCT**
SMa2061	2	CTCGAAATCATACCTGGACGCCCTAACT**
SMa2349	2	CTTGTGAACGCGCCTGCCGGATCATTTCT**
SMb20303	1, 2	CTCGTTCGTTGAAAAGTGCTGCCTTCTAT** CTGGATTCGAGCCGTATCCGCACCGATAT**
SMb21379	1, 2	TTGGGGAGCGGGCGGCGTGACCCTATCT**
SMb21683	2	CTTGCCAATCATCACCCAGCGCTTCCTT** CTTGCCGGCGGCCGGTCGGCTCCTATTC*
SMc00106	2	CTCGACGCGCCGCCTTGCGGTCCTATCT*
SMc00814	2	CTTCAGCGTGATCGGCCCGACACCCATGT**
SMc02656	1, 2	CTGGTGGCGGTTCGGCTGGTGAACTATCT**
SMc02769	1, 2	CTTGCGTCGGGGCTGTGGGTTTCCCACAT**
SMc02900	1, 2	CTACTTAAATCTTAGGCAAATGCCTATTT**
SMc03836	1, 2	GTTTTTGCCACGCGAAAGAGCCCTATAT**

[a] Consensus RpoH sequence for α-purple proteobacteria: CTTG (N_{17-18}) CYTATNT (267)

*One mismatch to consensus.

**Two mismatches to consensus.

3.4 DISCUSSION

My previous work had demonstrated that the requirement for RpoH1 and RpoH2 during symbiosis in *S. meliloti* cannot be explained solely by a loss of *groEL* expression and that there must be other crucial targets (20). To identify these targets by global gene expression analysis, I isolated RNA from wild-type, *rpoH1*, *rpoH2*, and *rpoH1 rpoH2* cells grown for 48 hours in minimal medium. Microarray analysis was performed by Carol Toman and Melanie Barnett (Stanford University), and I completed additional computer analysis of the data. In total, we identified nine down-regulated genes in the *rpoH1* mutant, 46 genes in the *rpoH2* mutant and 85 genes in the *rpoH1 rpoH2* mutant for a total of 88 unique genes, 34 of which were significantly decreased in both the Affymetrix and SAM statistical analyses.

Of the nine targets that were identified as significantly decreased in the *rpoH1* mutant, only two of these appeared to be RpoH1-specific. Additionally, only one of three known RpoH1 targets (*clpB*) was identified as significantly decreased in cells containing an *rpoH1* mutation (196). Because known RpoH1 targets (*groESL5* and *lon*) were missing from the analysis, our data suggest that *rpoH1* is not active under the conditions tested. When I designed the experiment, I looked for conditions where both *rpoH* genes appeared to be expressed so that I could monitor global transcription in the mutant cells simultaneously. Previous studies using transcriptional fusions had shown that both *rpoH1* and *rpoH2* are expressed during late stationary phase in minimal medium, suggesting that both transcription factors were active during stationary phase (20, 226). Because expression of *rpoH2* is induced later in stationary phase than *rpoH1*, I chose the earliest time point where *rpoH2* was significantly expressed.

There was a risk that the transcriptional fusion data did not accurately reflect the activity of the RpoH1 protein, because the activity from the reporter can remain after the gene is no longer transcribed. Therefore, it was possible that the regulons would have to be studied under separate conditions. To determine the RpoH1 regulon, I could monitor cells earlier in stationary phase or under heat shock conditions. Because more is known about RpoH1 activity during heat shock (196, 227) and because of technical issues involved with stationary phase cells, I will be performing microarray analysis on heat-shocked wild-type and *rpoH* mutant cells to determine the complete set of RpoH-regulated genes (Appendix A).

Previous studies have shown that *rpoH2* is transcribed during stationary phase in minimal media (226, 261), and the relatively large number of potential RpoH2 targets identified here confirms that this sigma factor is active under the experimental conditions assayed. A closely related sigma factor, RpoH2 in *Sinorhizobium* sp. BL3, is also produced primarily in late stationary phase (286). When combined with data from transcriptional fusions indicating that *rpoH2* is not expressed during the exponential phase of growth or during symbiosis (226), these results indicate that *rpoH2* is specialized for cellular responses during stationary phase, particularly in nutrient-limiting media. Although the targets of RpoH2 identified here by microarray analysis have not been confirmed, several putative targets may be involved in stress response, including *groEL4* (261) and several genes that encode putative oxidoreductases. Interestingly, although it was tested using S1 nuclease protection assays and transcriptional reporter fusions (20, 196), *groEL4* has not been previously reported to be regulated by RpoH2. However, transcriptional reporter fusions did suggest that RpoH1 and RpoH2 might control expression of *groEL4*, but the difference was not believed to be significant (Chapter 2). Unlike

the similar RpoH2 in *Rhizobium* sp. strain TAL1145 (162), *S. meliloti* RpoH2 does not appear to regulate the *exo* genes for exopolysaccharide synthesis under the conditions tested.

Genetic analysis of the *rpoH1*, *rpoH2*, and *rpoH1 rpoH2* mutants during symbiosis suggested that the regulons of RpoH1 and RpoH2 must at least partially overlap because both *rpoH1* and rpoH2 single mutants can form nodules on the roots of host plants, but an *rpoH1 rpoH2* double mutant cannot (20, 227). The microarray data are consistent with partially overlapping regulons since expression of 55 of the 85 genes identified in *rpoH1 rpoH2* mutant cells could be co-regulated by RpoH1 and RpoH2; expression of these genes either shows a greater fold-decrease in the double mutant than in either single mutant or is down-regulated in the double mutant only. In addition, expression of only two genes was down-regulated solely in *rpoH1* mutant cells compared to the wild type and expression of 31 genes was down-regulated solely in rpoH2 mutant cells. I also identified putative RpoH1 binding sites (196) or generic RpoH binding sites (for α-proteobacteria) (267) upstream of some of these genes. Based on this preliminary analysis of the data, it is possible that RpoH1 and RpoH2 control transcription of the same genes under different conditions. For example, RpoH2 is functional during late stationary phase in minimal media (226, 261, this study), whereas RpoH1 directs transcription of genes during early stationary phase (226) and in response to heat shock in rich medium (196, Appendix A). Because I have not yet identified the full regulons, it is possible that the two regulons fully overlap rather than partially overlap, which would still be consistent with the genetic data.

Previous studies demonstrated that the RpoH1 regulon partially overlaps with the *E. coli* RpoH regulon such that RpoH1 controls expression of *groESL5* (20, 196), *clpB* (196), and *lon*, but not *dnaK* or *clpA*. Here we identified the *E.coli* regulon homologs *clpB*, *ibpA*, *hslV*, and *groEL4* as potential RpoH targets. In *E. coli* RpoH is predicted to play a role in membrane

homeostasis during heat stress, as ~25% of its targets encode membrane-associated proteins (221). At least ten of the potential RpoH targets for *S. meliloti* are annotated as encoding possible membrane-associated proteins. This number could grow, since some of the proteins annotated as hypothetical could have membrane-related functions and we are missing some known RpoH1 targets from our analysis. The possible connection to membrane integrity is intriguing, because the *rpoH1* mutant is hypersensitive to detergents and hydrophobic dye (196), phenotypes that are associated with membrane integrity defects (129, 168).

3.5 CONCLUSIONS

To determine the regulons of RpoH1 and RpoH2, I performed a microarray experiment to monitor gene expression in wild-type and *rpoH* mutant cells during late stationary phase in M9 minimal medium. In this chapter I describe the statistical analysis of the microarray data. Of the nine down-regulated genes in the *rpoH1* mutant, 46 genes in the *rpoH2* mutant, and 85 genes in the *rpoH1 rpoH2* mutant, 34 were also identified as significantly changed by SAM. My results suggest that there is significant overlap between the RpoH1 and RpoH2 regulons and that there is some overlap with the *E. coli* RpoH regulon.

The low number of genes whose expression was significantly decreased in *rpoH1* mutant cells, in addition to known targets that were not identified, suggests that the time point analyzed was too late during stationary phase for RpoH1 activity. Therefore, to determine the complete set of RpoH-regulated gene targets, I intend to perform additional microarray experiments. First, I am comparing gene expression in wild-type and *rpoH* mutant cells during heat shock to obtain RpoH1 targets that were missing from the stationary phase analysis (Appendix A). Second, I

will induce expression of each *rpoH* gene in *rpoH1 rpoH2* cells and compare transcription to uninduced cells (Appendix B). Unlike the mutant analyses described here and in Appendix A, in the induction experiment, the genes for which RpoH1 or RpoH2 are sufficient for transcription will be up-regulated. Successful completion of both types of microarray experiments is ideal because it maximizes the likelihood of identifying true RpoH targets and obtaining complete regulons.

From the microarray results, I will choose a set of potential target genes for experimental confirmation. The criteria for these genes will be that expression is identified as significantly changed by both Affymetrix and SAM analyses and/or that they appear in the results of more than one microarray experiment (stationary phase, heat shock, or *rpoH* induction). I will experimentally confirm these targets by 5' RACE (5' rapid amplification of cDNA ends) to compare RNA levels in wild-type and *rpoH* mutant cells. 5' RACE has the advantage over Northern analysis or RT-PCR, which are often used to confirm microarray results, in that it identifies the start site of transcription, from which the promoter can be inferred. This will allow me to determine the consensus sequence for the RpoH1 and RpoH2 binding sites, and I will then search for binding sites upstream of all genes identified from the array experiments. Because of indirect effects that alter transcription, not all genes with decreased expression in a microarray experiment will be real targets. Therefore, genes with significantly altered expression in a microarray experiment and possessing a binding site will be considered true RpoH target genes.

4.0 ONLY ONE OF FIVE *GROEL* GENES IS REQUIRED FOR VIABILITY AND SUCCESSFUL SYMBIOSIS IN *SINORHIZOBIUM MELILOTI*

This chapter is adapted from previously published material in *Journal of Bacteriology*, Volume 189, Number 5, pages 1884-1889. It is reproduced here with permission from the American Society for Microbiology Journals Division.

4.1 ABSTRACT

Many bacterial species contain multiple copies of the genes that encode the chaperone GroEL and its co-chaperone GroES, including all of the fully sequenced root-nodulating bacteria that interact symbiotically with legumes to generate fixed nitrogen. In particular, in *Sinorhizobium meliloti* there are four *groESL* operons and one *groEL* gene. To uncover functional redundancies of these genes during growth and symbiosis, I attempted to construct strains containing all combinations of *groEL* mutations. Although a double *groEL1 groEL2* mutant cannot be constructed, I demonstrate that the quadruple *groEL1 groESL3 groEL4 groESL5* and *groEL2 groESL3 groEL4 groESL5* mutants are viable. Therefore, like *E. coli* and other species, *S.*

meliloti only requires one *groEL* for viability and either *groEL1* or *groEL2* will suffice. The *groEL1 groESL5* double mutant is more severely affected for growth at both 30° and 40°C than the single mutants, suggesting overlapping functions in stress response. During symbiosis the quadruple *groEL2 groESL3 groEL4 groESL5* mutant acts like the wild-type strain, but the quadruple *groEL1 groESL3 groEL4 groESL5* mutant acts like the *groEL1* single mutant, which forms ineffective nodules. Therefore, the only *groEL* gene required for symbiosis is *groEL1*. However, I show that the other *groEL* genes are expressed in the nodule, suggesting minor roles during symbiosis. Combining my data with other data, I conclude that *groESL1* encodes the housekeeping GroEL-GroES chaperone and that *groESL5* is specialized for stress response.

4.2 INTRODUCTION

The *groESL* operon encodes the chaperone GroEL and its co-chaperone GroES, which function as a multimeric complex that binds protein substrates and enables them to fold properly. Many bacterial species have only one *groESL* operon, and in *E. coli* the single copy is required for viability at temperatures as low as 17°C (75). Other bacterial species, however, have more than one *groESL* operon and additional *groEL* genes. The reason for maintaining multiple copies has not been fully determined for any species. One possibility is that the genes may be differentially regulated to provide GroEL-GroES at different times or at different levels. Such regulation has been observed in a number of species (20, 83, 108, 161, 170, 196, 252). A second possibility is that the genes may encode proteins with different substrate specificities. Although the substrates of the GroEL-GroES complexes in species with multiple *groESL* operons have not been determined, there is some evidence consistent with this hypothesis. For example, in *Rhizobium*

leguminosarum the three GroEL proteins have different *in vitro* properties for folding one substrate (103), and in *Sinorhizobium meliloti* GroEL3 is not able to functionally replace GroEL1 (20). A third possibility is that the proteins may be specialized for functions that do not include protein folding. For example, in *E. coli* GroEL is able to bind folded σ^{32}, decreasing σ^{32}–dependent transcription (127).

I am particularly interested in the role of multiple *groEL* genes in the root-nodulating bacteria of the Rhizobiales. These bacteria interact symbiotically with partner legume species by inducing the formation of nodules, colonizing the nodules, and then fixing nitrogen for the host plant. Multiple *groEL* copies have been found in all of the fully sequenced genomes of root nodulators: *Bradyrhizobium japonicum* (157), *Mesorhizobium loti* (156), *Rhizobium etli* (112), *R. leguminosarum* (319), and *S. meliloti* (97). In two cases *groEL* has been connected to symbiosis. In *B. japonicum* the *groESL3* operon is regulated along with nitrogen fixation genes, and a *groEL3 groEL4* double mutant is unable to fix nitrogen (Fix^-) (84). In *S. meliloti groEL1* is required for full induction of nodulation genes and nitrogen fixation (223).

S. meliloti has five *groEL* loci in the genome: *groESL1* and *groEL4* are located on the chromosome, *groESL2* and *groESL3* are located on the pSyma megaplasmid, and *groESL5* is located on the pSymb megaplasmid (97). Only the *groEL1* locus has been identified in mutant screens. Originally *groEL1* was discovered in a screen for reduced *nod* gene expression (223). The *nod* genes encode enzymes that produce Nod factor, which elicits nodule formation. The genes are controlled by several related transcription factors (NodD1, NodD2, and NodD3), some of which require plant inducers for activity. Biochemical studies have demonstrated that GroEL copurifies with NodD1 and NodD3, and GroEL-GroES modulates NodD activity (223, 316). However, the studies did not address which GroEL-GroES complexes are involved. The *groEL1*

locus was also identified in a screen for genes required for the production of *N*-acyl homoserine lactones used in quorum sensing (183). The deficiency in *N*-acyl homoserine lactone production may be due to a direct interaction of GroEL-GroES with the TraR regulator.

All of the single *S. meliloti groEL* mutants are viable (20, 196, 222, 223), but *groEL1* and *groEL2* cannot be disrupted at the same time (222). The effect of *groEL* mutations on growth rate has only been determined for the *groEL1* mutant, which has a longer doubling time at 30°C than the wild-type strain (183, 223). The effect of *groEL* mutations on symbiosis has been determined for all five *groEL* mutants and only *groEL1* is associated with symbiotic defects (20, 196, 222, 223). *groEL1* mutants are delayed in nodulation and are unable to fix nitrogen (223). Interestingly, *groEL2,* but not *groEL3,* can substitute for *groEL1* during symbiosis if expressed at high levels (20, 223).

Previous work has demonstrated that all of the *S. meliloti groEL* genes are expressed during free-living growth in rich and minimal media with *groESL1* expressed at high levels and the others expressed at low levels (20, 196). Transcription of only *groESL1* and *groESL5* increases upon heat shock (196). Two regulatory systems that bacteria use for controlling genes in response to heat stress are the RpoH sigma factor (124, 125), which directs transcription from specific promoters, and the HrcA repressor (210, 248, 264), which binds to a *cis*-acting element called CIRCE (326). *S. meliloti* has two genes that are known to encode RpoH sigma factors (226, 227). RpoH2 does not control any of the *groEL* genes, and RpoH1 only controls *groESL5* (20, 196). The *S. meliloti* genome contains one gene that is predicted to encode HrcA (97), and putative CIRCE elements are located upstream of *groESL1* and *groESL2*. However, the functionality of this CIRCE/HrcA system has not been demonstrated.

My goal was to uncover functional redundancies of the *groEL* genes in *S. meliloti* during growth and symbiosis by attempting to construct strains containing all combinations of *groE* mutations. In this chapter I demonstrate that *S. meliloti* cells only require one *groEL* for viability and either *groEL1* or *groEL2* will suffice. However, only *groEL1* is necessary and sufficient for symbiosis. Although the roles of *groEL2*, *groEL3*, and *groEL4* are still unclear, we present evidence that *groEL1* and *groESL5* have overlapping functions and suggest that *groESL5* is specialized for stress response.

4.3 MATERIALS AND METHODS

4.3.1 Strains and growth conditions

The bacterial strains used in this study were isogenic to the wild-type strain Rm1021 (188). The *groEL* mutations in strains JO138 (*groEL1Δ::gus-aph*) and JO60 (*groEL2Δ::gus-aph*) (222) were transduced into Rm1021 to remove the plasmid pPH1JI, which had been used for homogenotization, generating AB249 and AB247. AF14 (*groESL3Δ::tet*) (20), VO3193 (*groEL4Δ*) (20), and NI001 (*groESL5Δ::aacC1*) (196) have been previously published. Multiple *groE* mutants were constructed by generalized transduction using N3 phage. Transcriptional *groEL-gfp-gus* fusions were located in the chromosome or megaplasmids of AB140 (*groEL1::*pAB11), AB129 (*groEL2::*pAB10), AB145 (*groEL3::*pAB12), AB147 (*groEL4::*pAB13), and AB150 (*groEL5::*pAB14) in a manner such that the *groEL* gene is not disrupted (20). Cells were grown in LB medium supplemented with $MgSO_4$ and $CaCl_2$ (LB/MC medium) (109) at 30°C, unless otherwise indicated.

4.3.2 Plant assays

Medicago sativa plants were grown on nitrogen free medium at pH 6.0 and inoculated with *S. meliloti* cells as previously described (225). Nodulation was determined at three weeks post inoculation, and nitrogen fixation was determined at six weeks post inoculation. Fix^+ bacteria result in tall, green plants with pink nodules, whereas Fix^- bacteria result in stunted, chlorotic plants with white nodules.

4.3.3 Western blot analysis

To obtain samples for Western blot analysis, cells were grown overnight at 30°C in LB/MC medium with streptomycin, diluted to an OD_{595} of 0.1, and grown to mid log phase ($0.6 \leq OD_{595} \leq 0.8$). Cultures were then split and grown for an additional hour at 30°C or heat-shocked for an hour at 42°C, after which cells were harvested and frozen at -80°C. Cells were resuspended in 1× phosphate-buffered saline at 0.1 ml per OD_{595} unit and disrupted by sonication. The resulting extracts were combined with 2× Laemmli sample buffer, and equal volumes were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to nitrocellulose. Blots were probed with a 1:5,000 dilution of rabbit polyclonal antibodies to *E. coli* GroEL (Stressgen) followed by a 1:15,000 dilution of anti-rabbit horseradish peroxidase-conjugated secondary antibody (Amersham), developed with enhanced chemiluminescence reagents (Pierce), and imaged using a Fujifilm LAS-3000 imaging system.

4.3.4 Isolation of protein aggregates from *groE* mutants

This protocol is adapted from Tomoyasu *et al.* (289). To obtain samples of *groE* single mutants for aggregate analysis, cells were grown overnight at 30°C in LB/MC medium with streptomycin, diluted to an OD_{595} of 0.05, and grown to $OD_{595} \geq 0.4$. Cultures were split and grown for 2.5 hours at 30°C or heat-shocked for 2.5 hours at 42°C. Cultures were placed on ice while the optical density was measured. 5.5 ml of each culture was harvested (two tubes per strain, 11 ml total) and stored at -80°C.

Cells were resuspended in 40 µl buffer A (10 mM potassium phosphate buffer at pH 6.5, 1 mM EDTA, 20% (w/v) sucrose, 1 mg mL^{-1} lysozyme). After incubation on ice for 1 hour, 360 µl buffer B (10 mM potassium phosphate buffer at pH 6.5, 1 mM EDTA) was added. Cells were lysed by sonication (8 cycles of 15-second pulses). Extracts were centrifuged 2000 × *g* for 15 minutes at 4°C to remove unbroken cells. Supernatants were placed in new tubes and centrifuged 15,000 × *g* for 20 minutes at 4°C. Supernatants were discarded and pellets were stored at -80°C.

Pellets were resuspended in 400 µl buffer B by brief sonication. After centrifugation (15,000 × *g* for 20 minutes at 4°C), the supernatant was removed, and the pellet was resuspended in 320 µl buffer B by sonication. 80 µl of 10% NP40 was added to extract membrane proteins. After centrifugation (15,000 × *g* for 30 minutes at 4°C), the supernatant was removed, and the pellet was again resuspended as above, followed by an additional NP40 extraction. After centrifugation (15,000 × *g* for 30 minutes at 4°C), the pellet was washed once in buffer B and centrifuged again (15,000 × *g* for 10 minutes at 4°C). The supernatant was removed, and the

pellet was resuspended in 180 μl Buffer B. After pooling the aggregates from both tubes, the samples were concentrated by centrifugation (Microcon YM-3).

Samples were combined with 2× Laemmli sample buffer and heated for 10 minutes at 75°C. Equal volumes of sample were resolved by SDS-PAGE on a 12.5% acrylamide gel. The gel was silver stained using a protocol from Jeffrey Brodsky. Briefly, the gel was fixed for 30 minutes in fixative (50% ethanol, 12% acetic acid, 0.5 ml L^{-1} formaldehyde) and washed in 50% ethanol for 40 minutes. After brief washes first in 0.01% sodium thiosulfate and then water, the gel was stained in 0.1% silver nitrate for 1 hour. The gel was incubated in developer (0.28 M sodium carbonate, 0.0002% sodium thiosulfate, 480 μl L^{-1} formaldehyde) until bands were clearly visible, and the reaction was stopped in 10 mM EDTA.

4.3.5 Assays of β-glucuronidase activity

To quantitate β-glucuronidase (GUS) activity in free-living bacteria, cells were grown in LB/MC medium, harvested at the indicated times, and frozen at –80°C. The cells were then permeabilized using lysozyme (200 μg ml^{-1}, 37°C for 10 min), and β-glucuronidase activity was assayed using *p*-nitrophenyl-β-D-glucuronide as described previously (151). GUS activity is expressed in nmol per min per OD_{595} unit × 1,000.

To visualize β-glucuronidase activity *in planta*, nodules were sectioned and stained as described previously (279). Briefly, after sectioning with surgical razor blades, nodules were incubated with stain (1 mM X-gluc, 50 mM sodium phosphate buffer at pH 7, 0.02% SDS) overnight. Nodules were then viewed through a dissecting microscope and photographed on a dark background.

4.4 RESULTS

4.4.1 *groEL* requirements during free-living growth

Previous work has shown that all five *groE* operons can be disrupted (20, 196, 222, 223), but a *groEL1 groEL2* double mutant is not viable (222). To uncover functional redundancies among other *groEL* genes besides *groEL1* and *groEL2*, I attempted to construct all of the possible double, triple, and quadruple mutants. I confirmed that the *groEL1 groEL2* double mutant cannot be constructed but was able to construct strains containing all other combinations of mutations. Since the two quadruple mutants (abbreviated *1⁻ 3⁻ 4⁻ 5⁻* and *2⁻ 3⁻ 4⁻ 5⁻*) are viable, one of either *groEL1* or *groEL2* is necessary and sufficient for growth. Therefore, *S. meliloti* is like all other bacterial species with multiple *groEL* genes tested so far in requiring only one *groEL* gene for growth under non-stress conditions (170, 224, 252, 268).

To determine if the mutations affected growth under free-living conditions, I compared the growth of the single, double, triple, and quadruple *groE* mutants to the wild type in LB/MC medium at 30° and 40°C. At 30°C (Fig. 14A) I found that most of the *groE* mutants grew like the wild-type strain. The exceptions were that among the single mutants, the *groEL1* mutant displayed a slight but reproducible growth defect, as shown previously (183, 223), and among the multiple mutants, strains containing mutations in both *groEL1* and *groESL5* displayed a slightly more pronounced growth defect. In particular, these strains exhibited a longer lag phase and doubling time, although they reached the same maximum cell density. At 40°C (Fig. 14B) I found that all the *groE* single mutants had a growth defect, reaching lower cell densities than the wild type, with the *groEL1* mutant always being the most affected. The *groEL1 groESL3* double mutant had a slightly larger growth defect than either single mutant. Strains containing

mutations in both *groEL1* and *groESL5* exhibited a severe growth defect. All other double, triple and quadruple mutants displayed growth phenotypes similar to the single mutants (data not shown). In summary, although most of the *groE* mutants exhibit only minor growth defects if any, the double *groEL1 groESL5* mutant is temperature sensitive for growth.

In *E. coli groESL* is required for growth at low temperature (75), and the activity of GroEL-GroES in part determines the lower temperature limit at which the bacteria can grow (80). In addition, a mutation in another chaperone encoding gene, *dnaK,* renders the cells both cold sensitive and temperature sensitive for growth (29). Therefore, to determine whether *groE* mutants in *S. meliloti* are cold sensitive, I grew the quadruple mutants at 20°, 15°, and 10°C. Neither quadruple mutant was cold sensitive (Fig. 14C).

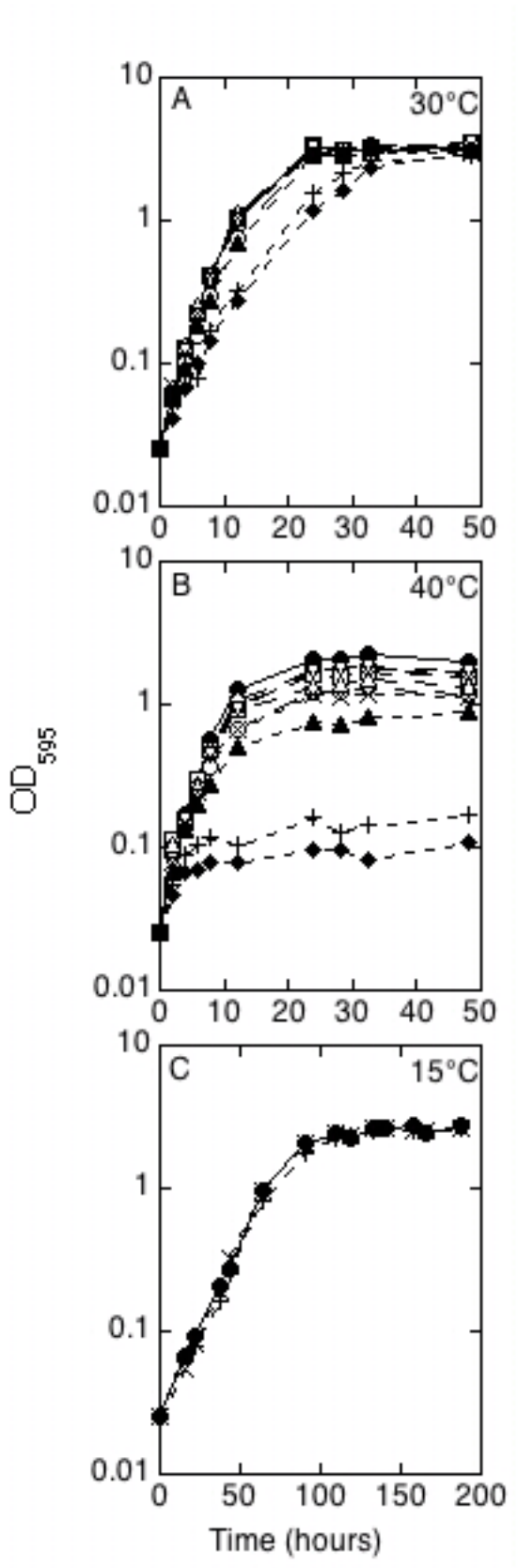
A
30°C
B
40°C
C
15°C
10
1
0.1
0.01
0
10
20
30
40
50
0
50
100
150
200
OD595
Time (hours)

Figure 14. Comparison of the growth of single and multiple *groE* mutants with the wild-type strain as measured using OD_{595}.

Cells were grown in LB/MC medium with streptomycin at 30°C (A), 40°C (B) and 15°C (C). Strains are Rm1021 (wild type; filled circles), AB249 (*groEL1*; open circles), AB247 (*groEL2*; open squares), AF14 (*groESL3*; open diamonds), VO3193 (*groEL4*; open triangles), NI001 (*groESL5*; open inverted triangles), AB221 (*groEL1 groESL3*; filled triangles), AB219 (*groEL1 groESL5*; filled diamonds), AB257 (1^- 3^- 4^- 5^-; plus signs), and AB238 (2^- 3^- 4^- 5^-; crosses). The experiment was repeated three times with essentially identical results, and the panels show data from one experiment.

I performed Western analysis on *groE* mutant strains using polyclonal antibodies to the *E. coli* GroEL protein (Fig. 15). At 30°C a single band for GroEL was obtained in wild-type cells. This band is predominantly due to GroEL1, as shown by the *2⁻ 3⁻ 4⁻ 5⁻* quadruple mutant, but also includes GroEL2, as shown by the *1⁻ 3⁻ 4⁻ 5⁻* quadruple mutant. After subjecting cells to heat shock at 42°C, the levels of GroEL1, but not GroEL2, increased, which is consistent with data on transcription (196). In wild-type cells, a second band of lower molecular weight appeared following heat shock. Production of the second band was dependent on the RpoH1 sigma factor (data not shown). Because *groESL5* is the only *groE* locus controlled by RpoH1 (20, 196) and transcription of *groESL5* increases upon heat shock (196), I hypothesized that the second band corresponded to GroEL5 protein. Consistent with this hypothesis, the second band was not produced in cells containing the *groESL5* deletion. Interestingly, the GroEL5 band was observed in the *groEL1* mutant even at 30°C, indicating that GroEL5 production increases when GroEL1 is absent. The effect is specific to the *groEL1* mutation because the band was not present at 30°C in the triple *groEL2 groESL3 groEL4* mutant strain.

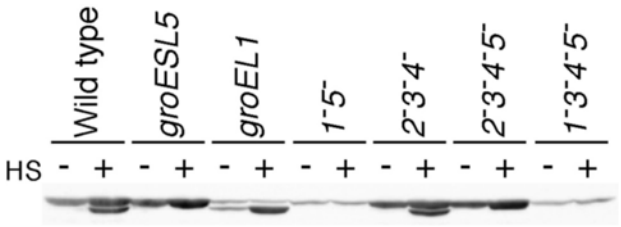

Figure 15. Western analysis of GroEL in *groE* mutant strains.

Cells were grown to mid-log phase in LB/MC medium with streptomycin at 30°C. Cultures were split and grown for an additional hour at 30°C (HS -) or at the heat-shock temperature of 42°C (HS +) before processing for Western analysis using a polyclonal antibody to *E. coli* GroEL. The strains were Rm1021 (wild type), NI001 (*groESL5*), AB249 (*groEL1*), AB219 (*groEL1 groESL5*), AB243 (*groEL2 groESL3 groEL4*), AB238 (*2⁻ 3⁻ 4⁻ 5⁻*), and AB257 (*1⁻ 3⁻ 4⁻ 5⁻*). The figure shows one representative blot out of three experiments.

The Western analysis did not allow me to determine whether production of GroEL2, GroEL3 or GroEL4 was increased in the *groEL1* mutant because the signals from these proteins are most likely masked by the high levels of GroEL1. To resolve this issue and confirm my findings for GroEL5, I transduced *groEL-gus* transcriptional fusions (20) into the *groEL1* mutant (Table 10). During the exponential and stationary phases of growth (6 and 24 hours, respectively), expression of *groEL2* and *groEL5* was significantly increased in the *groEL1* mutant compared to the wild-type strain. Expression of *groEL4* was significantly increased in the *groEL1* mutant only at 24 hours, suggesting that the effect on *groEL4* expression depends on growth phase. Expression of *groEL3* was not affected by the *groEL1* mutation. Therefore, loss of GroEL1 results in up-regulation of all of the other *groEL* genes except *groEL3*.

Table 10. Effect of a *groEL1* mutation on *groEL-gus* gene expression

Genotype	6 hr		24 hr	
	β-glucuronidase activity[a]	Fold difference	β-glucuronidase activity[a]	Fold difference
Wild type *groEL2-gus*	2,099		1,344	
groEL1::Tn*5 groEL2-gus*	6,804[b]	3.2	12,379[b]	9.2
Wild type *groEL3-gus*	75		189	
groEL1::Tn*5 groEL3-gus*	100	1.3	309	1.6
Wild type *groEL4-gus*	379		1,069	
groEL1::Tn*5 groEL4-gus*	359	0.95	2,047[c]	1.9
Wild type *groEL5-gus*	1,086		1,569	
groEL1::Tn*5 groEL5-gus*	7,238[b]	6.7	11,990[b]	7.6

[a] Average from three experiments.

[b] Significantly different from wild-type levels ($P < 0.01$) using the Student's *t* test.

[c] Significantly different from wild-type levels ($P < 0.05$) using the Student's *t* test.

Other studies have used isolation of protein aggregates to identify possible chaperone substrates (140, 289). In the absence of a particular chaperone, its protein substrates will misfold and can be detected in aggregate form. To explore substrate specificities, I isolated aggregates from the single *groE* mutants under heat-shock and nonheat-shock conditions in a preliminary experiment. The protocol was designed to reduce background by extracting membrane proteins from the aggregate pellet by solubilizing them in NP40 detergent (289). As shown in Fig. 16, very few aggregates are isolated from wild-type and single *groE* mutant cells grown at 30°C, although several bands are present in the *groEL1* mutant. At 42°C, many aggregates were isolated from each strain. Although a few bands seem to be specific to *groEL1*/*groEL2* or *groESL5*, there are too many bands present in the wild-type strain to identify aggregate bands as unique. I could not identify any bands specific to *groESL3* or *groEL4*, the *groE* genes about which the least is known.

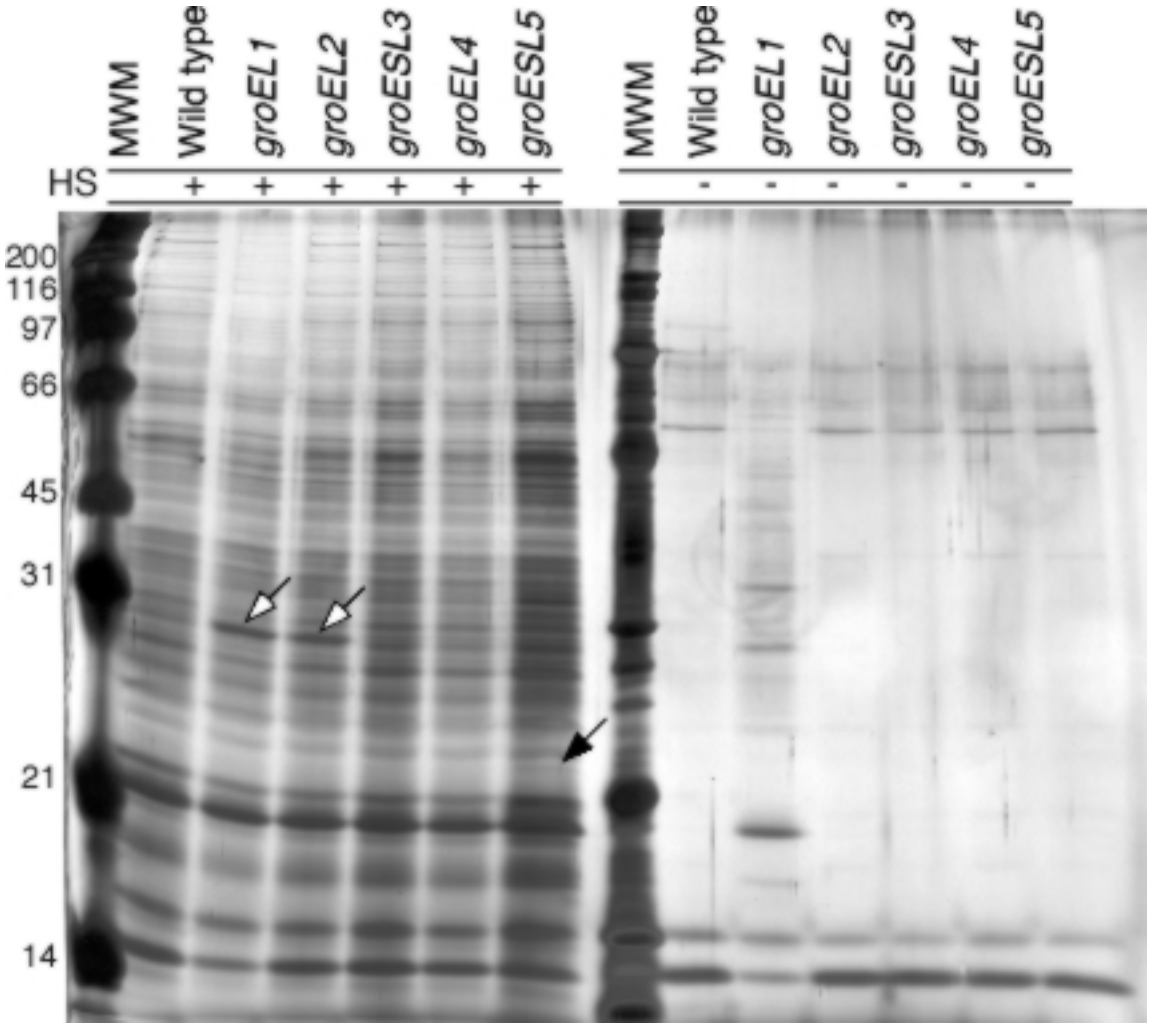

Figure 16. Isolation of protein aggregates from single *groE* mutant cells.

Cells were grown through log phase in LB/MC medium with streptomycin at 30°C. Cultures were split and grown for 2.5 hours at 30°C (HS -) or at the heat-shock temperature of 42°C (HS +) before processing for aggregate analysis. Arrows indicate bands possibly specific to a particular *groEL* mutant at 42°C (*groEL1*/*groEL2*, open arrows; *groEL5*, filled arrow). The strains were Rm1021 (wild type), AB249 (*groEL1*), AB247 (*groEL2*), AF14 (*groESL3*), VO3193 (*groEL4*), and NI001 (*groESL5*). Molecular weigh markers (MWM) are in kDa.

4.4.2 *groEL* requirements during symbiosis

To determine which combinations of *groEL* genes are important for symbiosis, I tested the effects on the formation of effective nodules. Among the single *groE* mutants only *groEL1* mutants have a symbiotic defect, resulting in Fix$^-$ nodules (20, 196, 222, 223). To uncover redundant functions, I inoculated alfalfa (*Medicago sativa*) plants with the triple and quadruple mutants and observed the plants for nodule formation and nitrogen fixation. Any mutant that contained the *groEL1* mutation formed Fix$^-$ nodules. All other mutants were similar to the wild-type strain in ability to nodulate and fix nitrogen.

Additionally, an undergraduate in the lab, Amanda Foltz, determined the effect of multiple *groE* mutations on the expression of *nod* genes (19). *groEL1* was previously shown to be required for full induction of *nod* gene expression in response to the plant inducer luteolin (223). To determine whether other *groE* genes play a role in *nod* gene expression, a multicopy plasmid containing *nodD1* and *nodC-lacZ* was introduced into wild-type cells and the single and quadruple mutants (19). The *groEL1* mutant and the *1$^-$ 3$^-$ 4$^-$ 5$^-$* quadruple mutant displayed lower expression of *nodC-lacZ* than the wild-type strain both in the absence and the presence of the plant inducer luteolin. In addition, the amount of induction caused by the addition of luteolin was reduced. In contrast, single *groEL2*, *groESL3*, *groEL4*, and *groESL5* mutants, as well as the *2$^-$ 3$^-$ 4$^-$ 5$^-$* quadruple mutant, displayed full *nodC-lacZ* expression. Therefore, only *groEL1* is necessary and sufficient for full induction of the *nod* genes, and the only *groEL* gene required for symbiosis is *groEL1*.

Expression of each of the *groE* loci within nodules on a variety of hosts has been detected in one or more of the global transcript or protein analyses (4, 16, 18, 58, 59, 215). To directly

compare levels of gene expression within alfalfa nodules, I inoculated plants with bacteria containing a matched set of *groEL-gus* transcriptional fusions (20). As shown in Fig. 17, all five *groEL* genes are expressed within the nodule, although at different levels. To quantitate expression, Valerie Oke harvested bacteria from nodules and determined β-glucuronidase activity (19). Similar to results obtained under free-living conditions (20), *groEL1-gus* was expressed at high levels; *groEL2-gus*, *groEL4-gus*, and *groEL5-gus* were expressed at low levels; and *groEL3-gus* was expressed at very low levels. Therefore, although only *groEL1* is required for effective nodules, all of the *groE* genes are expressed during symbiosis to some degree.

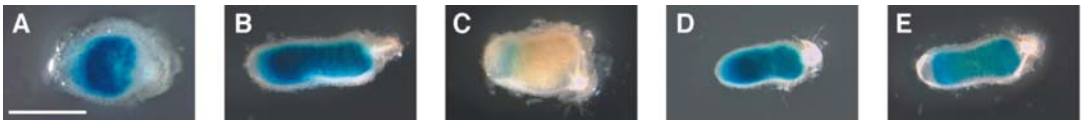

Figure 17. Expression of *groE* genes within the nodule.

Plants were inoculated with bacteria containing the following *groEL-gus* reporter fusions: *groEL1* (A), *groEL2* (B), *groEL3* (C), *groEL4* (D), and *groEL5* (E). Nodules were harvested 3 weeks post infection, hand-sectioned, and stained using X-gluc for β-glucuronidase activity. Ten nodules were examined per strain in each of three experiments, and the panels show one representative nodule. The tip of the nodule is on the left, and the root-proximal portion is on the right. The line in (A) represents 1 mm. Strains from left to right are AB140, AB129, AB145, AB147, and AB150.

4.5 DISCUSSION

I conclude that *groESL1* encodes the major housekeeping GroEL-GroES chaperone in *S. meliloti* for the following reasons. First, *groESL1* is located on the chromosome near many of the same neighboring genes as the single *groESL* operon in the closely related bacterium *Agrobacterium tumefaciens* (97, 310). Second, *groESL1* is expressed at much higher levels than the other *groE* genes during growth in culture and during symbiosis (20 and this study). Third, *groEL1* is sufficient for both growth in culture and successful symbiosis (this study). In addition to the housekeeping role, *groESL1* is also partially controlled by heat shock (196), indicating a role in stress response. This control is independent of RpoH1 and RpoH2 (20, 196) but may depend on a CIRCE/HrcA regulatory system (210) since a putative CIRCE element is located upstream of *groESL1*.

groESL5 is probably specialized for stress response since gene expression (196) and protein production (this study) is induced by heat shock, and *groESL5* is the only *groE* locus controlled by RpoH1 (20, 196). In addition, I have shown that *groESL5* is up-regulated in the absence of *groEL1*. I postulate that the loss of the major housekeeping chaperone results in unfolded proteins that trigger *groESL5* expression. The partially overlapping function of *groEL1* and *groEL5* can be observed by the synergistic effect of the two mutations on growth at both 30° and 40°C.

The roles of the other *groE* genes are still unclear. *groESL1* and *groESL2* encode very similar proteins (two amino acid differences for GroES and one amino acid difference for GroEL). Either *groEL1* or *groEL2* is sufficient during growth (this study), and *groEL2* can

substitute for *groEL1* during symbiosis if present on a multicopy plasmid (223). This suggests that the proteins are interchangeable but that *groESL2* is normally not expressed at high enough levels to be sufficient for symbiosis. As with *groESL1*, *groESL2* is preceded by a putative CIRCE element although heat shock control has not been observed (196). I have shown that the *groEL2* gene is up-regulated in the absence of *groEL1*. Given that transcriptional repression by the CIRCE/HrcA system in other bacteria depends upon levels of GroEL (10, 197), the increase in *groEL2* transcription could be mediated through its putative CIRCE element. Why does *groESL2* exist? Outside of the ORFs and CIRCE elements, the *groESL1* and *groESL2* DNA sequences are quite different, which would be consistent with differential regulation. We speculate that *groESL2* is expressed at high levels under some unknown condition when *groESL1* is not expressed well or in addition to *groESL1* when larger amounts of GroEL-GroES are needed.

The roles of *groESL3* and *groEL4* remain unknown. Presumably the genes produce GroES and GroEL under different conditions, encode chaperones that fold different ranges of substrates, and/or encode proteins specialized for non-folding functions. Previously we have shown that *groESL3*, which encodes the most divergent of the GroEL-GroES homologs, is unable to functionally replace *groEL1* (20). In addition, *groESL3* is the only *groE* locus that is not up-regulated in response to the loss of *groEL1*. These results would be consistent with different substrate specificities or different functions.

4.6 CONCLUSIONS

All of the root-nodulating rhizobia whose genomes have been sequenced maintain multiple copies of *groE* genes. My goal was to determine why *S. meliloti* maintains its five copies. The results of my genetic analysis of the multiple *groE* mutants, when combined with previous studies (20, 196, 222, 223), demonstrates that *groESL1* encodes the major housekeeping chaperonin, while *groESL5* is specialized for the stress response. Because it is so similar to *groESL1*, *groESL2* likely folds the same substrates, though perhaps under different conditions. The functions of *groESL3* and *groEL4* are unclear. These results, however, indicate that *S. meliloti* does not need multiple *groEL* genes for symbiosis, because *groEL1* is both necessary and sufficient.

In a preliminary experiment, I have attempted to identify differential substrates by isolating protein aggregates from single *groEL* mutant cells. Unfortunately, there was too much background in the wild-type strain to allow easy identification of unique proteins in the mutant strains. Possibly the cells were heat-shocked for too long, such that proteins began to aggregate in wild-type cells. Additionally, two-dimensional gel electrophoresis might be better suited to resolving these samples. An important caveat of this experiment is that *groESL2*, *groESL3*, and *groEL4* are not up-regulated in response to heat shock (196), and the proteins encoded by these genes might not be active under this condition. If these GroELs target specialized substrates, their substrates might not be present at high levels during heat shock and therefore might not be isolated in this assay. Further analysis of *groEL* function largely depends on determining the conditions under which the protein products play an active role.

APPENDIX A

HEAT SHOCK MICROARRAY EXPERIMENT TO DETERMINE GENE TARGETS OF RPOH1 AND RPOH2

A.1 INTRODUCTION

In Chapter 3 I discussed the results of a microarray experiment comparing gene expression in wild-type and *rpoH* mutant cells during late stationary phase in minimal medium. A study using transcriptional gene fusions indicated that during stationary phase expression of *rpoH1* is induced before *rpoH2* (226). I chose to harvest cells late in stationary phase in an attempt to define the RpoH1 and RpoH2 regulons in one experiment. However, the time point was too late for strong RpoH1 activity because only a few RpoH1 targets were identified and two of three known targets were missed.

RpoH was discovered in *E. coli* as an alternative sigma factor that directs the cellular response to heat stress (321), and the *V. cholerae* RpoH regulon was determined by comparing global gene expression in wild-type and *rpoH* mutant cells during heat shock (273). Previous studies in *S. meliloti* have shown that RpoH1 is active during heat shock conditions. Production of several heat shock proteins in response to heat stress is reduced in an *rpoH1* mutant (227), and

the heat-shock induction of *clpB*, *groESL5*, and *lon* is RpoH1-dependent (196). The best-studied target of RpoH1, *groESL5*, appears to encode the primary heat-shock response GroEL (19).

Therefore, to identify the RpoH1 regulon, I heat-shocked wild-type and *rpoH* mutant cells and isolated RNA for microarray analysis, after determining the appropriate heat shock conditions. I included the *rpoH2* and the *rpoH1 rpoH2* mutants in the analysis because one study suggested that RpoH2 might play a minor overlapping role during heat shock (227). The RNA samples have been sent to our collaborators at Stanford University for microarray analysis.

A.2 MATERIALS AND METHODS

Strains and growth conditions

Strains used in this study are Rm1021 (wild type), VO3128 (*rpoH1::aadA*), AB3 (*rpoH2::aacCI*), and AB9 (*rpoH1::aadA rpoH2::aacCI*), which have been previously published (20, 188, 226). Bacterial cultures were grown in LB/MC medium (109). Streptomycin was added to the medium at 500 μg ml^{-1}. *S. meliloti* cells were grown at 30°C unless otherwise indicated.

Heat shock and sample preparation

Three replicates for this experiment were started on the same day using the same medium. Cells were grown overnight in LB/MC medium, diluted to an OD_{595} of 0.05 the next day and allowed

to grow overnight again to ensure even growth. When the experiment was started, the growth of the three replicates was staggered by one hour. Cultures were diluted to an OD_{595} of 0.05 in 65 ml LB/MC and allowed to grow to the mid log phase of growth (0.5-0.7 OD_{595}). The wild-type culture was split so that 30 ml remained at 30°C for 15 minutes as a control, and 30 ml of each culture was heat-shocked for 15 minutes at 42°C. Cells were harvested by mixing with ice-cold stop solution (5% buffer-equilibrated phenol in ethanol) and centrifuging 5,000 rpm at 4°C. The supernatant was removed, and cells were frozen in liquid nitrogen and stored at -80°C.

Cell lysis and RNA isolation

Cell pellets were resuspended in 1 mg ml^{-1} lysozyme in TE and incubated for 5 minutes at room temperature to lyse the cells. Total RNA was isolated as described (16) using an RNeasy Mini Kit (Qiagen). This protocol includes both on-column and off-column DNase digests. Absence of chromosomal DNA was confirmed by PCR, and RNA integrity was validated on a 1.2% agarose formaldehyde gel. At least 80 ug of RNA was isolated for each sample and shipped to Sharon Long's laboratory at Stanford University for microarray analysis.

Real-time PCR analysis

Heat shock conditions were determined through reverse transcription of the mRNA for select gene targets followed by real-time PCR analysis for relative quantitation. Cultures were split and either kept at 30°C as a control or heat shocked at 42°C, and cells were harvested at the indicated times. 10 μg of RNA was reverse transcribed using SuperScript II reverse transcriptase (Invitrogen) with primers specific for *rpoH1*, *rpoH2*, *groESL5*, and SMc04228 (an endogenous

control gene that is constitutively expressed) at 20 ng per primer. The reverse transcription reaction was diluted to 100 µl in TE and 1 µl was subsequently used in each real-time reaction. Real-time PCR primers were designed using Primer Express Software version 3.0 (Applied Biosystems). The real-time PCR reactions were performed according to the SYBR® Green PCR Master Mix protocol (Applied Biosystems). Briefly, reactions contained 25 µl Master Mix, 0.25 µl forward and reverse primers from 100 µM stock, and 1 µl cDNA template in 50 µl total volume. Reactions for each sample were performed in triplicate using an Applied Biosystems 7300 system under the following conditions: 94°C for 10 minutes, then 40 cycles of 94°C for 15 seconds, 60°C for 30 seconds, and 72°C for 45 seconds. Real-time data were collected during the 72°C stage. During analysis, all target values were normalized to the endogenous control (SMc04228).

A.3 RESULTS AND DISCUSSION

Because RpoH1 did not appear to be active at 48 hours in minimal medium in the stationary phase microarray experiment (Chapter 3), I attempted to find conditions under which RpoH1 was functional. Previous work had shown that the *rpoH1* mutant is sensitive to high temperature and that the RpoH1 target *groESL5* is induced by heat shock (196, 227). The *rpoH1* gene is also expressed early in the stationary phase of growth (226), but obtaining RNA during stationary phase is technically difficult (Chapter 3). Therefore, I decided to monitor global gene expression in wild-type and *rpoH* mutant cells in response to heat shock.

To determine conditions for heat shock, I grew cultures of wild-type and *rpoH1* cells in M9 minimal medium, to match the medium that was used in the stationary phase microarray experiment (Chapter 3). Cultures were split once they had reached the mid log phase of growth and either kept at 30°C or heat-shocked at 42°C. Samples were taken every 15 minutes from 0 to 1 hour. After reverse transcription of select targets, I performed real-time PCR analysis of the cDNA. Significant induction of *groESL5* expression occurred within 15 minutes of heat shock (Fig. 18). Although *groESL5* was reported to be controlled solely by RpoH1 (20, 196), there was induction of *groESL5* expression even in the absence of RpoH1. One study performed in M9 minimal medium suggests that RpoH2 might control GroEL5 production in response to heat shock (the protein band was not specifically labeled) (227), but the rest of the work on *groESL5* regulation in response to heat shock was performed in LB/MC medium (19, 196). Therefore, I repeated the above experiment in LB/MC medium using wild-type, *rpoH1*, *rpoH2*, and *rpoH1 rpoH2* cells with a maximum heat shock of 30 minutes. The results demonstrate that transcription of *rpoH1* (Fig. 19A), *rpoH2* (Fig. 19B), and *groESL5* (Fig. 19C) is induced after a 15-minute heat shock. Importantly, induction of *groESL5* was greatly reduced in *rpoH1* and *rpoH1 rpoH2* mutant cells, suggesting that RpoH1 regulates transcription of *groESL5* in response to heat shock in LB/MC medium. Because cells in M9 minimal medium also grow more slowly and yield less RNA (data not shown), I chose to perform the heat shock experiment using LB/MC medium.

To obtain samples for microarray analysis, I isolated total RNA from wild-type, *rpoH1*, *rpoH2*, and *rpoH1 rpoH2* cells that were heat-shocked at 42°C for 15 minutes and from wild-type cells that were kept at 30°C as a non-heat shock control. I chose a short 15-minute heat shock to minimize the appearance of downstream targets that are not directly regulated by

RpoH1 and RpoH2. RNA integrity was confirmed by formaldehyde gel, and absence of chromosomal DNA contamination was verified by PCR amplification of an intergenic region. I am now waiting for the microarray results from our collaborators.

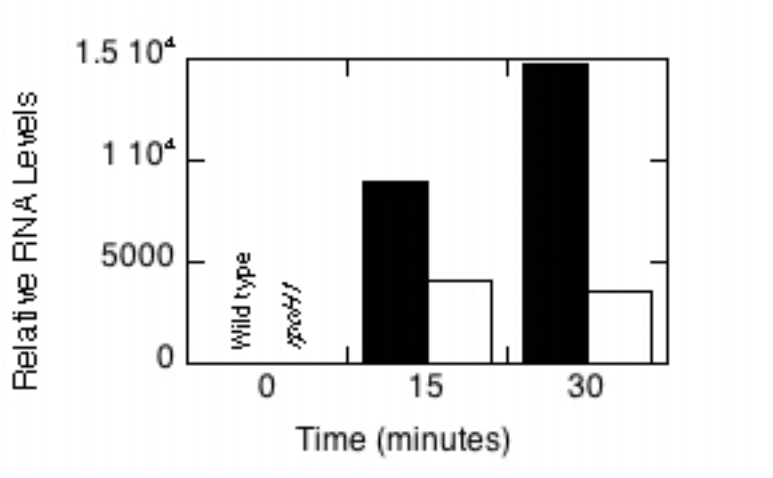

Figure 18. Relative expression level of *groESL5* in wild-type and *rpoH1* cells in response to heat shock in M9 minimal medium.

Cells were grown to mid log phase in M9 minimal medium, heat-shocked at 42°C, and harvested at the indicated times. Total RNA was isolated, and samples were processed for real-time PCR. Strains are Rm1021 (wild type, black bars) and VO3128 (*rpoH1*, white bars). All values are relative to the wild type signal (= 1) at time point 0.

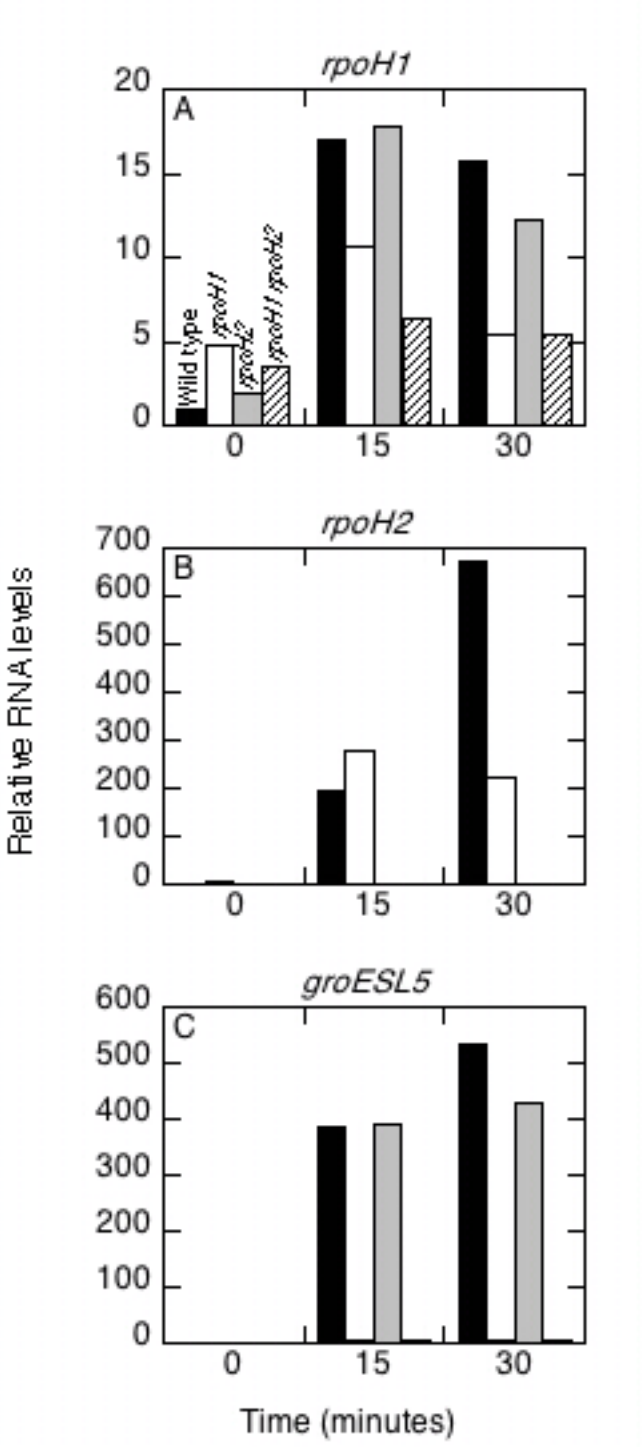
rpoH1
A
20
15
10
5
0
Wild type
rpoH1
rpoH2
rpoH1 rpoH2
0
15
30
rpoH2
B
700
600
500
400
300
200
100
0
0
15
30
groESL5
C
600
500
400
300
200
100
0
0
15
30
Relative RNA levels
Time (minutes)

Figure 19. Relative expression levels of *rpoH1, rpoH2,* and *groESL5* in wild-type and *rpoH* mutant cells in response to heat shock in LB/MC medium.

Cells were grown to mid log phase in LB/MC medium, heat-shocked at 42°C, and harvested at the indicated times. Total RNA was isolated, and samples were processed for real-time PCR. The transcripts analyzed are *rpoH1* (A), *rpoH2* (B), and *groESL5* (C). Strains are Rm1021 (wild type, black bars), VO3128 (*rpoH1*, white bars), AB3 (*rpoH2*, gray bars), and AB9 (*rpoH1 rpoH2*, diagonal lines). All values are relative to the wild type signal (= 1) at time point 0. Relative expression levels between Fig. 18 and Fig. 19 cannot be compared, because more cDNA was used for real-time PCR in Fig. 18.

A.4 CONCLUSIONS

Results from the stationary phase microarray experiment (Chapter 3) indicate that RpoH1 was not active under those growth conditions. Therefore, to determine the RpoH1 regulon, I designed a heat shock experiment and confirmed RpoH1 activity before isolating RNA for microarray analysis. Both *rpoH* single mutants and the double mutant were included in the analysis because transcription of both *rpoH* genes was induced upon heat shock.

When I receive the microarray data from Carol Toman and Melanie Barnett, I will perform data analysis with SAM. I will then experimentally confirm the targets as described in Chapter 3 (section 3.5). This analysis will identify the transcriptional start sites for genes whose expression is significantly decreased in *rpoH* mutant cells, which will allow me to infer the promoter sequence and determine consensus binding sequences for RpoH1 and RpoH2.

APPENDIX B

INDUCTION OF *RPOH1* AND *RPOH2* IN AN *RPOH1 RPOH2* DOUBLE MUTANT STRAIN TO DETERMINE THE GENE TARGETS OF RPOH1 AND RPOH2

B.1 INTRODUCTION

In Chapter 3 and Appendix A, I described experiments to define the RpoH1 and RpoH2 regulons by comparing global gene expression in wild-type and *rpoH* mutant cells. Although this approach is expected to identify RpoH targets whose expression is decreased, it will not distinguish between direct targets and secondary effects that alter transcription in mutant cells. A powerful and complementary approach to the mutant comparison is to induce expression of *rpoH1* and *rpoH2* under conditions where they are not typically expressed and to monitor the resulting changes in global gene expression over a time course. Genes that are induced early in the time course are likely to be direct RpoH targets, while genes induced later could be indirect targets or could have promoters that are weakly bound by RpoH.

The use of complementary mutant and induction analyses has been used previously to determine the regulons of sigma factors in *Bacillus subtilis* (28, 69). The *E. coli* (221, 323) and the *N. gonorrhoeae* (128) RpoH regulons have been determined by inducing *rpoH* expression during normal growth conditions in part because *rpoH* mutants grow slowly. For the *E. coli*

experiments, changes in gene expression were monitored within minutes of induction to minimize downstream effects, and samples were taken at several time points to determine how the response to *rpoH* gene-induction changed over time (221, 323).

I wanted to induce expression of *S. meliloti rpoH1* and *rpoH2* in *rpoH1 rpoH2* double mutant cells and harvest cells soon after induction to minimize the influence of indirect effects that alter transcription. In this experiment, targets whose expression is increased shortly after induction would be considered likely RpoH targets. I chose to work in *rpoH1 rpoH2* double mutant cells to eliminate transcription of gene targets by the endogenous proteins. Although *rpoH1* and *rpoH2* are primarily expressed during stationary phase, expression of *rpoH1* is induced between the mid and late log phases of growth (226, Bittner and Oke, unpublished results).

The primary difficulty with this experiment is that inducible promoters have not been well-developed in *S. meliloti*. The promoters commonly used *E. coli* are either strongly transcribed but not well repressed (*lac* and *tac*) (182, Margolin, personal communication to V. Oke) or are tightly repressed but not strongly transcribed (*araBAD*) (Bittner and Oke, unpublished results). The xylose-inducible system used in the α-proteobacterium *Caulobacter crescentus* (189) is also ineffective (Peck and Long, personal communication to V. Oke). However, the recently-identified *S. meliloti melA* promoter is promising because it appears to be well repressed by succinate and strongly induced by α-galactosides (27, 95). In this section, I describe the construction and testing of inducible *rpoH* constructs that have been placed under control of the *melA* promoter.

B.2 MATERIALS AND METHODS

Strains, plasmids and growth conditions

Bacterial strains are listed in Table 11. Bacterial cultures were grown in LB medium, LB/MC medium (109), M9 minimal medium containing 0.2% succinate, 0.5 μg biotin ml^{-1}, 1 mM $MgSO_4$, and 0.25 mM $CaCl_2$, Vincent minimal medium (VMM) containing 0.4% succinate, 14.7 mM K_2HPO_4, 11.5 mM KH_2PO_4, 15.7 mM NH_4Cl, 1 mM $MgSO_4$, 460 μM $CaCl_2$, 37 μM $FeCl_3$, 4.1 μM biotin, 48.5 μM H_3BO_3, 10 μM $MnSO_4$, 1 μM $ZnSO_4$, 0.5 μM $CuSO_4$, 0.27 μM $CoCl_2$, and 0.5 μM $NaMoO_4$ (261), or standard minimal medium (SMM) containing 0.3% succinate, 13.8 mM K_2HPO_4, 11.7 mM KH_2PO_4, 6.8 mM NaCl, 15.0 mM NH_4Cl, 2.0 mM $MgSO_4$, 0.84 μM $CoCl_2$, 0.82 μM biotin, and 1.8 μM thiamine (123). Streptomycin was added to the media at 500 μg ml^{-1}, and spectinomycin was added at 50 μg ml^{-1}. *S. meliloti* cells were grown at 30°C unless otherwise indicated. Plasmids were introduced into *S. meliloti* cells by triparental conjugation (109). Chromosomally located constructs were moved between *S. meliloti* strains by generalized transduction using N3 phage (184). The *rpoH1::aphII* mutation from strain HY658N (227) was transferred by transduction into strain AB3 (20), creating AB71, to ensure isogenicity with our lab strains.

Table 11. Strains used in Appendix B

Strain	Relevant characteristics	Reference
AB3	*rpoH2::aacCI*	(20)
AB71	*rpoH1::aphII rpoH2::aacCI*	This study
AB306	*rpoH1::aphII rpoH2::aacCI*/pCAP11 (vector)	This study
AB308	*rpoH1::aphII rpoH2::aacCI*/pAB15 (P_{melA}-*rpoH1*)	This study
AB310	*rpoH1::aphII rpoH2::aacCI*/pAB16 (P_{melA}-*rpoH2*)	This study
HY658N	*rpoH1::aphII*	(227)

Construction of plasmids to induce expression of *rpoH1* and *rpoH2*

To place *rpoH1* under the control of the *S. meliloti melA* (melibiose-inducible) promoter, a 0.91 kb DNA fragment that extends from the *rpoH1* start codon to 3 bp downstream of the stop codon was amplified with primers that generate *Avr* II restriction sites (Table 12). The upstream primer also includes a consensus Shine-Dalgarno sequence for translation. The fragment was inserted into *Avr* II-digested and phosphatased pCAP11 (Gift of Catalina Arango Pinedo), a variant of the broad-host-range vector pMB393 (14), that contains the *melA* promoter, creating pAB15 (P_{melA}-*rpoH1*). To place *rpoH2* under the control of the *melA* promoter, a 0.85 kb fragment that extends from the *rpoH2* start codon to 3 bp after the stop codon was amplified with primers that generate *Avr* II restriction sites. The upstream primer was identical to the *rpoH1* upstream primer up to the start codon so that both constructs would have identical Shine-Dalgarno sequences and spacing between the start codon and the ribosome-binding site. The fragment was inserted into *Avr* II-digested and phosphatased pCAP11, creating pAB16 (P_{melA}-*rpoH2*). Both constructs were verified by multiple restriction digests and sequencing.

Table 12. Primers used in Appendix B

Primers	**Sequence**[a]
rpoH1 upstream	cggccctaggacaggaggatacgcgATGGCCCGCAATACCTTG
rpoH1 downstream	aatgcctaggcctTTAAGCGCCTTCAAC
rpoH2 upstream	gaagcctaggacaggaggatacgcgATGATCAAGATTGCCATG
rpoH2 downstream	gaagcctagggctTCAATGCATCGACGC

[a] Sequences read 5' to 3'. *Avr* II sites, Shine-Dalgarno sequences, and start/stop codons are underlined. Sequences corresponding to the *rpoH1* or *rpoH2* genes are capitalized.

Melibiose induction

Prior to induction, cells were grown in LB/MC medium overnight, diluted to an OD_{595} of 0.05 in LB/MC the next day and allowed to grow overnight again to ensure even growth. Cells were then washed twice and diluted to an OD_{595} of 0.05 in minimal medium containing succinate or glycerol. When cultures reached an $OD_{595} \geq 0.5$, cells were centrifuged, washed twice in minimal medium lacking a carbon source, and resuspended in minimal medium containing 1% melibiose as the sole carbon source. Time points were taken every 30 minutes from 0 to 4 hours. Cells were harvested as described (16, Appendix A.2.2).

Cell lysis and RNA isolation

Cells were lysed and total RNA was isolated as described (16, Appendix A.2.3).

Real-time PCR analysis

Melibiose-induction conditions were determined through reverse transcription of the mRNA for select gene targets followed by real-time PCR analysis for relative quantitation. Real-time PCR analysis was performed as described above (Appendix A.2.4) with the exception that 2.5 μg of RNA and 12 ng per primer were used in the reverse transcription reaction, which was diluted to a final volume of 65 μl in TE.

B.3 RESULTS AND DISCUSSION

An important tool for the analysis of sigma factor regulons has been the ability to induce production of the transcription factor and to assay changes in gene expression shortly after induction (28, 69). It is important that transcription of the gene that encodes the sigma factor is tightly repressed, so that little of the protein is active before induction. When time points are taken shortly after induction, the likelihood of identifying direct targets is increased, and the contribution of downstream effects is minimized. However, an inducible promoter that is both tightly repressed and strongly induced has not been developed for *S. meliloti*. I and others have been testing the *S. meliloti melA* promoter for use as an inducible system in this organism. The *melA* promoter controls expression of the *melA-agp* operon, which encodes proteins necessary for the transport and utilization of α-galactosides in *S. meliloti* (95). Expression from this promoter is tightly repressed by succinate and strongly induced by α-galactosides (27, 95).

pCAP11 (gift from Catalina Arango Pinedo in Daniel Gage's laboratory at the University of Connecticut) is a multicopy plasmid containing the minimal *melA* promoter fragment downstream of a transcriptional terminator to prevent read-through from vector promoters. To initially characterize the requirements for *melA* induction, I placed a *gus* transcriptional reporter fusion downstream of the *melA* promoter. Cells were grown in M9 succinate minimal medium, washed twice with medium lacking a carbon source, and then induced with M9 minimal medium containing different concentrations of the α-galactoside melibiose. Cells induced by 1-2% melibiose yielded the highest induction levels, which reached approximately 25-fold above background after 6 hours (data not shown). A 6-hour induction period is much longer than desired; sufficient levels of induction from inducible systems in *E. coli* can be obtained within

minutes. To attempt to shorten the induction period, I grew the cells in medium containing glycerol, which is a neutral rather than a repressive carbon source, and added melibiose to the medium during mid log phase to induce the promoter. However, this procedure yielded similar induction times to the previous experiment, and background levels of expression were much higher (data not shown). Therefore, I decided to use the original growth conditions where all cells were initially grown in medium containing succinate to efficiently repress the promoter prior to induction with 1% melibiose.

To make inducible *rpoH* constructs, I placed *rpoH1* (pAB15) and *rpoH2* (pAB16) downstream of the *melA* promoter and moved both plasmids and the control vector (pCAP11) into *rpoH1 rpoH2* mutant cells. To test induction conditions in these new strains, I grew cells in M9 succinate medium containing streptomycin and 50 μg ml^{-1} spectinomycin. It was apparent that the strain containing the *rpoH1* construct grew more quickly than the strains containing the *rpoH2* construct or the control vector. To find conditions under which all three strains grew evenly, I altered spectinomycin concentration (50 μg ml^{-1} and 35 μg ml^{-1}), temperature (30°C and 25°C), and growth medium (VMM and SMM). The strains grew most evenly in SMM containing 50 μg ml^{-1} spectinomycin at 30°C (data not shown), and therefore these conditions are used in the following preliminary experiment.

To test induction requirements for the *rpoH* constructs, I grew cells to mid log phase, washed twice in SMM lacking a carbon source, and resuspended cells in SMM containing 1% melibiose as a carbon source. Cells were harvested every 30 minutes from 0 to 4 hours. Real-time PCR analysis of cDNA generated from *rpoH1* (Fig. 20A) and *rpoH2* (Fig. 20B) indicated that *rpoH1* was maximally induced between 3 and 4 hours and that *rpoH2* was maximally induced between 2.5 and 3.5 hours. The relative induction levels of *rpoH1* are misleading

because the reverse transcription primer will anneal to both the induced *rpoH1* transcript and the transcript from the *rpoH1* mutant allele. Therefore, the observed induction levels are most likely lower than the actual induction levels because of background transcription from the mutant allele in the absence of inducer.

The level of induction of *groESL5* is most puzzling. At the time point at zero hours, there are already high levels of *groESL5* transcript in cells expressing *rpoH1* (Fig. 20C). This suggests that the *melA* promoter is not well repressed, allowing sufficient production of RpoH1 to increase *groESL5* transcript levels in the absence of inducer. After induction with melibiose, *groESL5* levels immediately decrease, followed by an increase up to 4 hours that never reaches pre-induction levels. The drop in *groESL5* transcript levels suggests that the mRNA could be degraded when cells are switched to a medium containing melibiose. However, this experiment should be repeated to determine whether this pattern continues in subsequent experiments. Interestingly, induction of *rpoH2* also increased *groESL5* levels, although to a lesser extent than induction of *rpoH1*. Two previous studies had concluded that RpoH2 does not control expression of *groESL5* (20, 196). However, my work on expression of *groESL5* in response to heat shock (Appendix A) has suggested that expression of *groESL5* is regulated by another transcription factor in addition to RpoH1 in minimal medium. This is the first experiment to show that an increase in *groESL5* transcription correlates to an increase *rpoH2* expression.

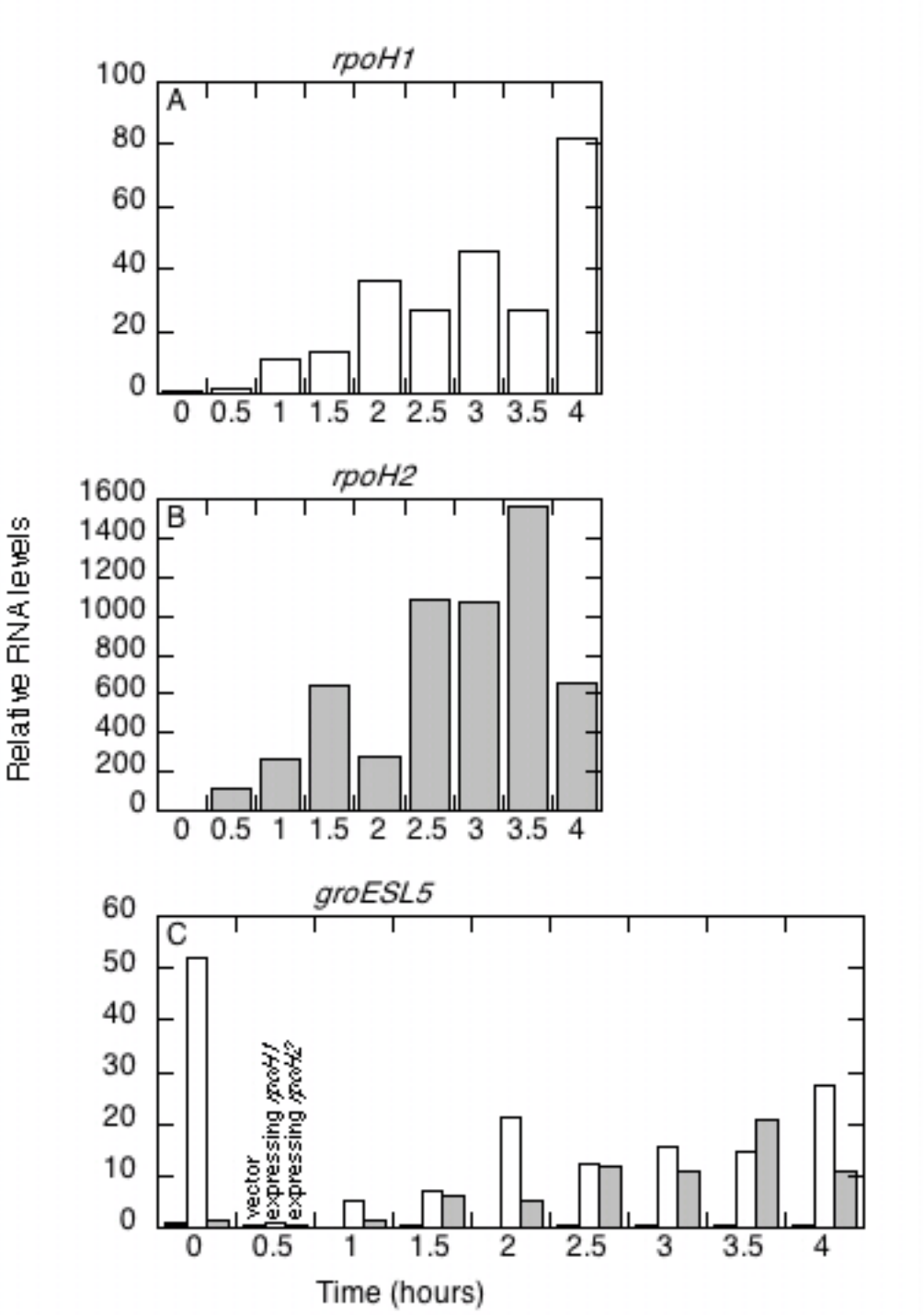
rpoH1
A
100
80
60
40
20
0
0 0.5 1 1.5 2 2.5 3 3.5 4
rpoH2
B
1600
1400
1200
1000
800
600
400
200
0
0 0.5 1 1.5 2 2.5 3 3.5 4
Relative RNA levels
groESL5
C
60
50
40
30
20
10
0
vector
expressing rpoH1
expressing rpoH2
0 0.5 1 1.5 2 2.5 3 3.5 4
Time (hours)

Figure 20. Relative expression levels of *rpoH1*, *rpoH2*, and *groESL5* in *rpoH1 rpoH2* mutant cells expressing *rpoH1* or *rpoH2*.

Cells were grown to mid log phase in SMM medium containing succinate, washed twice with medium lacking a carbon source, induced with SMM medium containing 1% melibiose, and harvested at the indicated times. Total RNA was isolated, and samples were processed for real-time PCR. (A) Relative *rpoH1* expression levels in AB308 (*rpoH1 rpoH2*/P_{melA}-*rpoH1*); values are relative to time point 0 (= 1). (B) Relative *rpoH2* expression levels in AB310 (*rpoH1 rpoH2*/P_{melA}-*rpoH2*); values are relative to time point 0 (= 1). (C) Relative *groESL5* expression levels in AB306 (*rpoH1 rpoH2*/vector; black bars), AB308 (*rpoH1 rpoH2*/P_{melA}-*rpoH1*; white bars), and AB310 (*rpoH1 rpoH2*/P_{melA}-*rpoH2*; gray bars); values are relative to the vector alone control (AB306) at time point 0 (= 1).

B.4 CONCLUSIONS

To complement the stationary phase and heat shock microarray experiments to determine complete RpoH1 and RpoH2 regulons, I wanted to perform microarray experiments with *rpoH1 rpoH2* mutant cells that have been engineered to induce expression of *rpoH1* and *rpoH2*. By monitoring changes in transcription shortly after induction, I hoped to reduce the chance of identifying downstream targets that are not directly regulated by RpoH1 and RpoH2. However, the most commonly used inducible promoters are not well repressed or highly induced in *S. meliloti*.

Here I have attempted to induce *rpoH1* and *rpoH2* in *rpoH1 rpoH2* mutant cells using the melibiose-inducible *S. meliloti melA* promoter. While I was able to obtain high levels of induction for *rpoH1* and *rpoH2*, the levels of *groESL5* indicate that the *melA* promoter is not well-repressed. Unexpectedly, *groESL5* transcript levels decrease immediately after induction, which suggests that the *groESL5* mRNA could be degraded. To determine whether this effect is genuine, the experiment should be repeated and include additional time points both before and after the switch to growth in melibiose-containing medium. In addition, results from this preliminary experiment suggest that RpoH2 could contribute to regulation of *groESL5* when RpoH1 is not present, which has not been shown in any previous study (20, 196). My work on *groESL5* expression in response to heat shock (Appendix A) suggested that a sigma factor other than RpoH1 could regulate *groESL5* in a minimal medium, and these results indicate that the other sigma factor could be RpoH2.

Finally, induction of the *melA* promoter was comparatively slow, on the order of hours instead of minutes, under all conditions tested. The combination of a promoter that is not tightly repressed and a long induction period increases the likelihood of identifying downstream targets that are not directly regulated by RpoH1 or RpoH2. However, this experiment would still allow me to determine RpoH targets by identifying genes whose expression is increased in cells expressing *rpoH1* or *rpoH2*. This could confirm results obtained in the previous experiments that monitored changes in gene expression in *rpoH* mutant cells (Chapter 3 and Appendix A), and it might also identify additional targets that were missed in the other experiments.

BIBLIOGRAPHY

1. **Alesandrini, F., R. Mathis, G. Van de Sype, D. Herouart, and A. Puppo.** 2003. Possible roles of a cysteine protease and hydrogen peroxide in soybean nodule development and senescence. New Phytologist **158:**131-8.
2. **Alsmark, C. M., A. C. Frank, E. O. Karlberg, B.-A. Legault, D. H. Ardell, B. Canbäck, A.-S. Eriksson, A. K. Naslund, S. A. Handley, M. Huvet, B. La Scola, M. Holmberg, and S. G. E. Andersson.** 2004. The louse-borne human pathogen *Bartonella quintana* is a genomic derivative of the zoonotic agent *Bartonella henselae*. Proceedings of the National Academy of Sciences of the United States of America **101:**9716-9721.
3. **Amor, B. B., S. L. Shaw, G. E. Oldroyd, F. Maillet, R. V. Penmetsa, D. Cook, S. R. Long, J. Dénarié, and C. Gough.** 2003. The *NFP* locus of *Medicago truncatula* controls an early step of Nod factor signal transduction upstream of a rapid calcium flux and root hair deformation. The Plant Journal **34:**495-506.
4. **Ampe, F., E. Kiss, F. Sabourdy, and J. Batut.** 2003. Transcriptome analysis of *Sinorhizobium meliloti* during symbiosis. Genome Biology **4:**R15.
5. **Anthony, J. R., K. L. Warczak, and T. J. Donohue.** 2005. A transcriptional response to singlet oxygen, a toxic byproduct of photosynthesis. Proceedings of the National Academy of Sciences of the United States of America **102:**6502-7.
6. **Apel, K., and H. Hirt.** 2004. Reactive oxygen species: metabolism, oxidative stress, and signal transduction. Annual Review of Plant Biology **55:**373-99.
7. **Ardissone, S., P. Frendo, E. Laurenti, W. Jantschko, C. Obinger, A. Puppo, and R. P. Ferrari.** 2004. Purification and physical-chemical characterization of the three hydroperoxidases from the symbiotic bacterium *Sinorhizobium meliloti*. Biochemistry **43:**12692-9.
8. **Ardourel, M., N. Demont, F. Debellé, F. Miallet, F. de Billy, J. C. Promé, J. Dénarié, and G. Truchet.** 1994. *Rhizobium meliloti* lipooligosaccharide nodulation factors: different structural requirements for bacterial entry into target root hair cells and induction of plant symbiotic developmental responses. The Plant Cell **6:**1357-1374.
9. **Arnold, W., A. Rump, W. Klipp, U. B. Priefer, and A. Pühler.** 1988. Nucleotide sequence of a 24,206-base-pair DNA fragment carrying the entire nitrogen fixation gene cluster of *Klebsiella pneumoniae*. Journal of Molecular Biology **203:**715-38.

10. **Babst, M., H. Hennecke, and H.-M. Fischer.** 1996. Two different mechanisms are involved in the heat-shock regulation of chaperonin gene expression in *Bradyrhizobium japonicum*. Molecular Microbiology **19:**827-839.
11. **Bakkes, P. J., B. W. Faber, H. van Heerikhuizen, and S. M. van der Vies.** 2005. The T4-encoded cochaperonin, gp31, has unique properties that explain its requirement for the folding of the T4 major capsid protein. Proceedings of the National Academy of Sciences of the United States of America **102:**8144-8149.
12. **Banfalvi, Z., and A. Kondorosi.** 1989. Production of root hair deformation factors by *Rhizobium meliloti* nodulation genes in *Escherichia coli*: HsnD (NodH) is involved in the plant host-specific modification of the NodABC factor. Plant Molecular Biology **13:**1-12.
13. **Bardwell, J. C., and E. A. Craig.** 1987. Eukaryotic Mr 83,000 heat shock protein has a homologue in *Escherichia coli*. Proceedings of the National Academy of Sciences of the United States of America **84:**5177-81.
14. **Barnett, M. J., V. Oke, and S. R. Long.** 2000. New genetic tools for use in the *Rhizobiaceae* and other bacteria. BioTechniques **29:**240-245.
15. **Barnett, M. J., J. A. Swanson, and S. R. Long.** 1998. Multiple genetic controls on *Rhizobium meliloti syrA*, a regulator of exopolysaccharide abundance. Genetics **148:**19-32.
16. **Barnett, M. J., C. J. Toman, R. F. Fisher, and S. R. Long.** 2004. A dual-genome Symbiosis Chip for coordinate study of signal exchange and development in a prokaryote-host interaction. Proceedings of the National Academy of Sciences of the United States of America **101:**16636-16641.
17. **Batut, J., M.-L. Daveran-Mingot, M. David, J. Jacobs, A. M. Garnerone, and D. Kahn.** 1989. *fixK*, a gene homologous with *fnr* and *crp* from *Escherichia coli*, regulates nitrogen fixation genes positively and negatively in *Rhizobium meliloti*. The EMBO Journal **8:**1279-1286.
18. **Becker, A., H. Bergès, E. Krol, C. Bruand, S. Rüberg, D. Capela, E. Lauber, E. Meilhoc, F. Ampe, F. J. de Bruijn, J. Fourment, A. Francez-Charlot, D. Kahn, H. Küster, C. Liebe, A. Pühler, S. Weidner, and J. Batut.** 2004. Global changes in gene expression in *Sinorhizobium meliloti* 1021 under microoxic and symbiotic conditions. Molecular Plant-Microbe Interactions **17:**292-303.
19. **Bittner, A. N., A. Foltz, and V. Oke.** 2007. Only one of five *groEL* genes is required for viability and successful symbiosis in *Sinorhizobium meliloti*. Journal of Bacteriology **189:**1884-9.
20. **Bittner, A. N., and V. Oke.** 2006. Multiple *groESL* operons are not key targets of RpoH1 and RpoH2 in *Sinorhizobium meliloti*. Journal of Bacteriology **188:**3507-3515.
21. **Boada, J., T. Roig, X. Perez, A. Gamez, R. Bartrons, M. Cascante, and J. Bermúdez.** 2000. Cells overexpressing fructose-2,6-bisphosphatase showed enhanced pentose phosphate pathway flux and resistance to oxidative stress. FEBS Letters **480:**261-4.
22. **Bochkareva, E. S., N. M. Lissin, G. C. Flynn, J. E. Rothman, and A. S. Girshovich.** 1992. Positive cooperativity in the functioning of molecular chaperone GroEL. Journal of Biological Chemistry **267:**6796-800.
23. **Bochkareva, E. S., N. M. Lissin, and A. S. Girshovich.** 1988. Transient association of newly synthesized unfolded proteins with the heat-shock GroEL protein. Nature **336:**254-7.

24. **Braig, K., Z. Otwinowski, R. Hegde, D. C. Boisvert, A. Joachimiak, A. L. Horwich, and P. B. Sigler.** 1994. The crystal structure of the bacterial chaperonin GroEL at 2.8 Å. Nature **371:**578-86.
25. **Braig, K., M. Simon, F. Furuya, J. F. Hainfeld, and A. L. Horwich.** 1993. A polypeptide bound by the chaperonin GroEL is localized within a central cavity. Proceedings of the National Academy of Sciences of the United States of America **90:**3978-82.
26. **Brewin, N. J.** 2004. Plant cell wall remodelling in the *Rhizobium*-legume symbiosis. Critical Reviews in Plant Science **23:**293-316.
27. **Bringhurst, R. M., and D. J. Gage.** 2000. An AraC-like transcriptional activator is required for induction of genes needed for α-galactoside utilization in *Sinorizobium meliloti*. FEMS Microbiology Letters **188:**23-27.
28. **Britton, R. A., P. Eichenberger, J.-E. González-Pastor, P. Fawcett, R. Monson, R. Losick, and A. D. Grossman.** 2002. Genome-wide analysis of the stationary-phase sigma factor (Sigma-H) regulon of *Bacillus subtilis*. Journal of Bacteriology **184:**4881-4890.
29. **Bukau, B., and G. C. Walker.** 1989. Cellular defects caused by deletion of the *Escherichia coli dnaK* gene indicate roles for heat shock protein in normal metabolism. Journal of Bacteriology **171:**2337-2346.
30. **Burgess, R. R., and A. A. Travers.** 1970. *Escherichia coli* RNA polymerase: purification, subunit structure, and factor requirements. Federation Proceedings **29:**1164-9.
31. **Burgess, R. R., A. A. Travers, J. J. Dunn, and E. K. Bautz.** 1969. Factor stimulating transcription by RNA polymerase. Nature **221:**43-6.
32. **Burston, S. G., N. A. Ranson, and A. R. Clarke.** 1995. The origins and consequences of asymmetry in the chaperonin reaction cycle. Journal of Molecular Biology **249:**138-52.
33. **Campbell, G. R., B. L. Reuhs, and G. C. Walker.** 2002. Chronic intracellular infection of alfalfa nodules by *Sinorhizobium meliloti* requires correct lipopolysaccharide core. Proceedings of the National Academy of Sciences of the United States of America **99:**3938-43.
34. **Cao, P., M. S. McClain, M. H. Forsyth, and T. L. Cover.** 1998. Extracellular release of antigenic proteins by *Helicobacter pylori*. Infection and Immunity **66:**2984-6.
35. **Capela, D., F. Barloy-Hubler, J. Gouzy, G. Bothe, F. Ampe, J. Batut, P. Boistard, A. Becker, M. Boutry, E. Cadieu, S. Dréano, 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. Thébault, M. Vandenbol, S. Weidner, and F. Galibert.** 2001. Analysis of the chromosome sequence of the legume symbiont *Sinorhizobium meliloti*. Proceedings of the National Academy of Sciences of the United States of America **98:**9877-9882.
36. **Cardenas, L., J. E. Thomas-Oates, N. Nava, I. M. Lopez-Lara, P. K. Hepler, and C. Quinto.** 2003. The role of nod factor substituents in actin cytoskeleton rearrangements in *Phaseolus vulgaris*. Molecular Plant-Microbe Interactions **16:**326-34.
37. **Cardenas, L., J. Vidali, J. Dominguez, H. Perez, F. Sanchez, P. K. Hepler, and C. Quinto.** 1998. Rearrangement of actin microfilaments in plant root hairs responding to *Rhizobium etli* nodulation signals. Plant Physiology **116:**871-7.

38. **Cebolla, A., J. M. Vinardell, E. Kiss, B. Oláh, F. Roudier, A. Kondorosi, and É. Kondorosi.** 1999. The mitotic inhibitor *ccs52* is required for endoreduplication and ploidy-dependent cell enlargement in plants. The EMBO Journal **18:**4476-84.
39. **Chen, L., and P. B. Sigler.** 1999. The crystal structure of a GroEL/peptide complex: plasticity as a basis for substrate diversity. Cell **99:**757-68.
40. **Chen, S., A. M. Roseman, A. S. Hunter, S. P. Wood, S. G. Burston, N. A. Ranson, A. R. Clarke, and H. R. Saibil.** 1994. Location of a folding protein and shape changes in GroEL-GroES complexes imaged by cryo-electron microscopy. Nature **371:**261-4.
41. **Chen, S., and D. F. Smith.** 1998. Hop as an adaptor in the heat shock protein 70 (Hsp70) and hsp90 chaperone machinery. Journal of Biological Chemistry **273:**35194-200.
42. **Choi, E. Y., G. S. Ahn, and K. W. Jeon.** 1991. Elevated levels of stress proteins associated with bacterial symbiosis in *Amoeba proteus* and soybean root nodule cells. Biosystems **25:**205-12.
43. **Chuang, S. E., and F. R. Blattner.** 1993. Characterization of twenty-six new heat shock genes of *Escherichia coli*. Journal of Bacteriology **175:**5242-52.
44. **Chuang, S. E., D. L. Daniels, and F. R. Blattner.** 1993. Global regulation of gene expression in *Escherichia coli*. Journal of Bacteriology **175:**2026-36.
45. **Corbin, D., L. Barran, and G. Ditta.** 1983. Organization and expression of *Rhizobium meliloti* nitrogen fixation genes. Proceedings of the National Academy of Sciences of the United States of America **80:**3005-3009.
46. **D'Haeze, W., J. Glushka, R. De Rycke, M. Holsters, and R. W. Carlson.** 2004. Structural characterization of extracellular polysaccharides of *Azorhizobium caulinodans* and importance for nodule initiation on *Sesbania rostrata*. Molecular Microbiology **52:**485-500.
47. **da Silva, A. C., R. C. Simao, M. F. Susin, R. L. Baldini, M. Avedissian, and S. L. Gomes.** 2003. Downregulation of the heat shock response is independent of DnaK and sigma32 levels in *Caulobacter crescentus*. Molecular Microbiology **49:**541-53.
48. **David, M., M.-L. Daveran, J. Batut, A. Dedieu, O. Domergue, J. Ghai, C. Hertig, P. Boistard, and D. Kahn.** 1988. Cascade regulation of *nif* gene expression in *Rhizobium meliloti*. Cell **54:**671-683.
49. **Davies, B. W., and G. C. Walker.** 2008. A highly conserved protein of unknown function is required by *Sinorhizobium meliloti* for symbiosis and environmental stress protection. Journal of Bacteriology **190:**1118-23.
50. **Davies, B. W., and G. C. Walker.** 2007. Disruption of *sitA* compromises *Sinorhizobium meliloti* for manganese uptake required for protection against oxidative stress. Journal of Bacteriology **189:**2101-9.
51. **Davies, B. W., and G. C. Walker.** 2007. Identification of novel *Sinorhizobium meliloti* mutants compromised for oxidative stress protection and symbiosis. Journal of Bacteriology **189:**2110-13.
52. **Deaton, J., C. G. Savva, J. Sun, A. Holzenburg, J. Berry, and R. Young.** 2004. Solubilization and delivery by GroEL of megadalton complexes of the λ holin. Protein Science **13:**1778-86.
53. **Deaton, J., J. Sun, A. Holzenburg, D. K. Struck, J. Berry, and R. Young.** 2004. Functional bacteriorhodopsin is efficiently solubilized and delivered to membranes by the chaperonin GroEL. Proceedings of the National Academy of Sciences of the United States of America **101:**2281-6.

54. **Debellé, F., and S. B. Sharma.** 1986. Nucleotide sequence of *Rhizobium meliloti* RCR2011 genes involved in host specificity of nodulation. Nucleic Acids Research **14:**7453-7472.
55. **Delory, M., R. Hallez, J. J. Letesson, and X. De Bolle.** 2006. An RpoH-like heat shock sigma factor is involved in stress response and virulence in *Brucella melitensis* 16M. Journal of Bacteriology **188:**7707-10.
56. **DelVecchio, V. G., V. Kapatral, R. J. Redkar, G. Patra, C. Mujer, T. Los, N. Ivanova, I. Anderson, A. Bhattacharyya, A. Lykidis, G. Reznik, L. Jablonski, N. Larsen, M. D'Souza, A. Bernal, M. Mazur, E. Goltsman, E. Selkov, P. H. Elzer, S. Hagius, D. O'Callaghan, J.-J. Letesson, R. Haselkorn, N. Kyrpides, and R. Overbeek.** 2001. The genome sequence of the facultative intracellular pathogen *Brucella melitensis*. Proceedings of the National Academy of Sciences of the United States of America **99:**443-448.
57. **Denefle, P., A. Kush, F. Norel, A. Paquelin, and C. Elmerich.** 1987. Biochemical and genetic analysis of the *nifHDKE* region of *Rhizobium* ORS571. Molecular & General Genetics **207:**280-7.
58. **Djordjevic, M. A.** 2004. *Sinorhizobium meliloti* metabolism in the root nodule: a proteomic perspective. Proteomics **4:**1859-1872.
59. **Djordjevic, M. A., H. C. Chen, S. Natera, G. Van Noorden, C. Menzel, S. Taylor, C. Renard, O. Geiger, t. *S.* D. S. Consortium, and G. F. Weiller.** 2003. A global analysis of protein expression profiles in *Sinorhizobium meliloti*: Discovery of new genes for nodule occupancy and stress adaptation. Molecular Plant-Microbe Interactions **16:**508-524.
60. **Djordjevic, M. A., J. W. Redmond, M. Batley, and B. G. Rolfe.** 1987. Clovers secrete specific phenolic compounds which either stimulate or repress nod gene expression in *Rhizobium trifolii*. The EMBO Journal **6:**1173-1179.
61. **Du, Y., and C. G. Arvidson.** 2006. RpoH mediates the expression of some, but not all, genes induced in *Neisseria gonorrhoeae* adherent to epithelial cells. Infection and Immunity **74:**2767-76.
62. **Du, Y., J. Lenz, and C. G. Arvidson.** 2005. Global gene expression and the role of sigma factors in *Neisseria gonorrhoeae* in interactions with epithelial cells. Infection and Immunity **73:**4834-45.
63. **Dylan, T., D. R. Helinski, and G. S. Ditta.** 1990. Hypoosmotic adaptation in *Rhizobium meliloti* requires β-(1□2)-glucan. Journal of Bacteriology **172:**1400-8.
64. **Dylan, T., P. Nagpal, D. R. Helinski, and G. S. Ditta.** 1990. Symbiotic pseudorevertants of *Rhizobium meliloti ndv* mutants. Journal of Bacteriology **172:**1409-17.
65. **Earl, C. D., C. W. Ronson, and F. M. Ausubel.** 1987. Genetic and structural analysis of the *Rhizobium meliloti fixA, fixB, fixC,* and *fixX* genes. Journal of Bacteriology **169:**1127-1136.
66. **Egelhoff, T. T., and S. R. Long.** 1985. *Rhizobium meliloti* nodulation genes: identification of *nodDABC* gene products, purification of *nodA* protein, and expression of *nodA* in *Rhizobium meliloti*. Journal of Bacteriology **164:**591-9.
67. **Ehrhardt, D. W., E. M. Atkinson, and S. R. Long.** 1992. Depolarization of alfalfa root hair membrane potential by *Rhizobium meliloti* Nod factors. Science **256:**998-1000.

68. **Ehrhardt, D. W., R. Wais, and S. R. Long.** 1996. Calcium spiking in plant root hairs responding to Rhizobium nodulation signals. Cell **85:**673-681.
69. **Eichenberger, P., S. T. Jensen, E. M. Conlon, C. van Ooij, J. Silvaggi, J.-E. González-Pastor, M. Fujita, S. Ben-Yehuda, P. Stragier, J. S. Liu, and R. Losick.** 2003. The σ^E regulon and the identification of additional sporulation genes in *Bacillus subtilis*. Journal of Molecular Biology **327:**945-972.
70. **Engelke, T., D. Jording, D. Kapp, and A. Pühler.** 1989. Identification and sequence analysis of the *Rhizobium meliloti dctA* gene encoding the C_4-dicarboxylate carrier. Journal of Bacteriology **171:**5551-60.
71. **Erickson, J. W., and C. A. Gross.** 1989. Identification of the sigma E subunit of *Escherichia coli* RNA polymerase: a second alternate sigma factor involved in high-temperature gene expression. Genes & Development **3:**1462-71.
72. **Erickson, J. W., V. Vaughn, W. A. Walter, F. C. Neidhardt, and C. A. Gross.** 1987. Regulation of the promoters and transcripts of *rpoH*, the *Escherichia coli* heat shock regulatory gene. Genes & Development **1:**419-432.
73. **Ewalt, K. L., J. P. Hendrick, W. A. Houry, and F. U. Hartl.** 1997. *In vivo* observation of polypeptide flux through the bacterial chaperonin system. Cell **90:**491-500.
74. **Farr, G. W., K. Furtak, M. B. Rowland, N. A. Ranson, H. R. Saibil, T. Kirchhausen, and A. L. Horwich.** 2000. Multivalent binding of nonnative substrate proteins by the chaperonin GroEL. Cell **100:**561-73.
75. **Fayet, O., T. Ziegelhoffer, and C. Georgopoulos.** 1989. The *groES* and *groEL* heat shock gene products of *Escherichia coli* are essential for bacterial growth at all temperatures. Journal of Bacteriology **171:**1379-1385.
76. **Fellay, R., X. Perret, V. Viprey, W. J. Broughton, and S. Brenner.** 1995. Organization of host-inducible transcripts on the symbiotic plasmid of *Rhizobium* sp. NGR234. Molecular Microbiology **16:**657-67.
77. **Felle, H., É. Kondorosi, A. Kondorosi, and M. Schultze.** 1996. Rapid alkalinization in alfalfa root hairs in response to rhizobial lipochitooligosaccharide signals. The Plant Journal **10:**295-301.
78. **Felle, H., É. Kondorosi, A. Kondorosi, and M. Schultze.** 1998. The role of ion fluxes in Nod factor signalling in *Medicago sativa*. The Plant Journal **13:**455-63.
79. **Fenton, W. A., Y. Kashi, K. Furtak, and A. L. Horwich.** 1994. Residues in chaperonin GroEL required for polypeptide binding and release. Nature **371:**614-9.
80. **Ferrer, M., T. N. Chernikova, M. M. Yakimov, P. N. Golyshin, and K. N. Timmis.** 2003. Chaperonins govern growth of *Escherichia coli* at low temperatures. Nature Biotechnology **21:**1266-1267.
81. **Figurski, D. H., and D. R. Helinski.** 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in *trans*. Proceedings of the National Academy of Sciences of the United States of America **76:**1648-52.
82. **Finan, T. M., A. M. Hirsch, J. A. Leigh, E. Johansen, G. A. Kuldau, S. Deegan, G. C. Walker, and E. R. Signer.** 1985. Symbiotic mutants of *Rhizobium meliloti* that uncouple plant from bacterial differentiation. Cell **40:**869-77.
83. **Fischer, H.-M., M. Babst, T. Kaspar, G. Acuña, F. Arigoni, and H. Hennecke.** 1993. One member of a *groESL*-like chaperonin multigene family in *Bradyrhizobium*

japonicum is co-regulated with symbiotic nitrogen fixation genes. The EMBO Journal **12:**2901-12.

84. **Fischer, H.-M., K. Schneider, M. Babst, and H. Hennecke.** 1999. GroEL chaperonins are required for the formation of a functional nitrogenase in *Bradyrhizobium japonicum*. Archives of Microbiology **171:**279-289.
85. **Fisher, R. F., H. L. Brierley, J. T. Mulligan, and S. R. Long.** 1987. Transcription of *Rhizobium meliloti* nodulation genes. Identification of a *nodD* transcription initiation site *in vitro* and *in vivo*. The Journal of Biological Chemistry **262:**6849-55.
86. **Fisher, R. F., T. T. Egelhoff, J. T. Mulligan, and S. R. Long.** 1988. Specific binding of proteins from *Rhizobium meliloti* cell-free extracts containing NodD to DNA sequences upstream of inducible nodulation genes. Genes & Development **2:**282-93.
87. **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. Journal of Molecular Biology **233:**336-48.
88. **Foucher, F., and E. Kondorosi.** 2000. Cell cycle regulation in the course of nodule organogenesis in *Medicago*. Plant Molecular Biology **43:**773-86.
89. **Friedland, J. S., R. Shattock, D. G. Remick, and G. E. Griffin.** 1993. Mycobacterial 65-kD heat shock protein induces release of proinflammatory cytokines from human monocytic cells. Clinical and Experimental Immunology **91:**58-62.
90. **Friedman, D. I., E. R. Olson, C. Georgopoulos, K. Tilly, I. Herskowitz, and F. Banuett.** 1984. Interactions of bacteriophage and host macromolecules in the growth of bacteriophage lambda. Microbiological Reviews **48:**299-325.
91. **Frisk, A., C. A. Ison, and T. Lagergard.** 1998. GroEL heat shock protein of *Haemophilus ducreyi*: association with cell surface and capacity to bind to eukaryotic cells. Infection and Immunity **66:**1252-7.
92. **Gage, D. J.** 2002. Analysis of infection thread development using Gfp- and DsRed-expressing *Sinorhizobium meliloti*. Journal of Bacteriology **184:**7042-6.
93. **Gage, D. J.** 2004. Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. Microbiology and Molecular Biology Reviews **68:**280-300.
94. **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*). Journal of Bacteriology **178:**7159-7166.
95. **Gage, D. J., and S. R. Long.** 1998. α-galactoside uptake in *Rhizobium meliloti*: Isolation and characterization of *agpA*, a gene encoding a periplasmic binding protein required for melibiose and raffinose utilization. Journal of Bacteriology **180:**5739-5748.
96. **Gagnon, H., and R. K. Ibrahim.** 1998. Aldonic acids: a novel family of *nod* gene inducers of *Mesorhizobium loti*, *Rhizobium lupini*, and *Sinorhizobium meliloti*. Molecular Plant-Microbe Interactions **11**.
97. **Galibert, F., T. M. Finan, S. R. Long, A. Pühler, A. P. Abola, F. Ampe, F. Barloy-Hubler, M. J. Barnett, A. Becker, P. Boistard, G. Bothe, M. Boutry, L. Bowser, J. Buhrmester, E. Cadieu, D. Capela, P. Chain, A. Cowie, R. W. Davis, S. Dréano, N. A. Federspiel, R. F. Fisher, S. Gloux, T. Godrie, A. Goffeau, B. Golding, J. Gouzy, M. Gurjal, I. Hernandez-Lucas, A. Hong, L. Huizar, R. W. Hyman, T. Jones, D. Kahn, M. Kahn, S. Kalman, D. H. Keating, E. Kiss, C. Komp, V. Lelaure, D. Masuy, C. Palm, M. C. Peck, T. M. Pohl, D. Portetelle, B. Purnelle, U. Ramsperger, R.**

Surzycki, P. Thébault, M. Vandenbol, F.-J. Vorhölter, S. Weidner, D. H. Wells, K. Wong, K.-C. Yeh, and J. Batut. 2001. The composite genome of the legume symbiont *Sinorhizobium meliloti*. Science **293:**668-672.

98. **Galitski, T., A. J. Saldanha, C. A. Styles, E. S. Lander, and G. R. Fink.** 1999. Ploidy regulation of gene expression. Science **285:**251-4.
99. **Gamer, J., H. Bujard, and B. Bukau.** 1992. Physical interaction between heat shock proteins DnaK, DnaJ, and GrpE and the bacterial heat shock transcription factor sigma 32. Cell **69:**833-42.
100. **Gamer, J., G. Multhaup, T. Tomoyasu, J. S. McCarty, S. Rüdiger, H.-J. Schönfeld, C. Schirra, H. Bujard, and B. Bukau.** 1996. A cycle of binding and release of the DnaK, DnaJ and GrpE chaperones regulates the activity of the *Escherichia coli* heat shock transcription factor σ^{32}. The EMBO Journal **15:**607-617.
101. **Garduno, R. A., G. Faulkner, M. A. Trevors, N. Vats, and P. S. Hoffman.** 1998. Immunolocalisation of Hsp60 in *Legionella pneumophila*. Journal of Bacteriology **180:**505-13.
102. **Garduno, R. A., E. Garduno, and P. S. Hoffman.** 1998. Surface-associated hsp60 chaperonin of *Legionella pneumophila* mediates invasion in a HeLa cell model. Infection and Immunity **66:**4602-10.
103. **George, R., S. M. Kelly, N. C. Price, A. Erbse, M. Fisher, and P. A. Lund.** 2004. Three GroEL homologues from *Rhizobium leguminosarum* have distinct in vitro properties. Biochemical and Biophysical Research Communications **324:**822-828.
104. **Georgopoulos, C. P., and H. Eisen.** 1974. Bacterial mutants which block phage assembly. Journal of Supramolecular Structure **2:**349-59.
105. **Georgopoulos, C. P., R. W. Hendrix, S. R. Casjens, and A. D. Kaiser.** 1973. Host participation in bacteriophage lambda head assembly. Journal of Molecular Biology **76:**45-60.
106. **Georgopoulos, C. P., R. W. Hendrix, A. D. Kaiser, and W. B. Wood.** 1972. Role of the host cell in bacteriophage morphogenesis: effects of a bacterial mutation on T4 head assembly. Nature: New Biology **239:**38-41.
107. **Gérard, H. C., J. A. Whittum-Hudson, H. R. Schumacher, and A. P. Hudson.** 2004. Differential expression of three *Chlamydia trachomatis* hsp60-encoding genes in active vs. persistent infections. Microbial Pathogenesis **36:**35-9.
108. **Glatz, A., I. Horváth, V. Varvasovszki, E. Kovács, Z. Török, and L. Vigh.** 1997. Chaperonin genes of the *Synechocystis* PCC 6803 are differentially regulated under light-dark transition during heat stress. Biochemical and Biophysical Research Communications **239:**291-297.
109. **Glazebrook, J., and G. C. Walker.** 1991. Genetic techniques in *Rhizobium meliloti*. Methods in Enzymology **204:**398-418.
110. **Glenn, A. R., W. G. Reeve, R. P. Tiwari, and M. J. Dilworth.** 1999. Acid tolerance in root nodule bacteria. Novartis Foundation Symposium **221:**112-26; discussion 126-30.
111. **Goloubinoff, P., J. T. Christeller, A. A. Gatenby, and G. H. Lorimer.** 1989. Reconstitution of active dimeric ribulose bisphosphate carboxylase from an unfolded state depends on two chaperonin proteins and Mg-ATP. Nature **342:**884-9.
112. **González, V., R. I. Santamaría, P. Bustos, I. Hernández-González, A. Medrano-Soto, G. Moreno-Hagelsieb, S. C. Janga, M. A. Ramírez, V. Jiménez-Jacinto, J. Collado-Vides, and G. Dávila.** 2006. The partitioned *Rhizobium etli* genome: genetic and

metabolic redundancy in seven interacting replicons. Proceedings of the National Academy of Sciences of the United States of America **103:**3834-3839.

113. **Göttfert, M., D. Holzhauser, D. Bani, and H. Hennecke.** 1992. Structural and functional analysis of two different *nodD* genes in *Bradyrhizobium japonicum* USDA110. Molecular Plant-Microbe Interactions **5:**257-65.
114. **Gould, P. S., H. R. Burgar, and P. A. Lund.** 2007. Homologous *cpn60* genes in *Rhizobium leguminosarum* are not functionally equivalent. Cell Stress & Chaperones **12:**123-31.
115. **Gould, P. S., M. Maguire, and P. A. Lund.** 2007. Distinct mechanisms regulate expression of the two major *groEL* homologues in *Rhizobium leguminosarum*. Archives of Microbiology **187:**1-14.
116. **Govezensky, D., E. S. Bochkareva, A. Zamir, and A. S. Girshovich.** 1994. Chaperonins as potential gene regulatory factors: *in vitro* interaction and solubilization of NifA, the *nif* transcriptional activator, with GroEL. The Journal of Biological Chemistry **269:**14003-14006.
117. **Govezensky, D., T. Greener, G. Segal, and A. Zamir.** 1991. Involvement of GroEL in *nif* gene regulation and nitrogenase assembly. Journal of Bacteriology **173:**6339-46.
118. **Goyal, K., R. Qamra, and S. C. Mande.** 2006. Multiple gene duplication and rapid evolution in the *groEL* gene: functional implications. Journal of Molecular Evolution **63:**781-7.
119. **Gragerov, A., E. Nudler, N. Komissarova, G. A. Gaitanaris, M. E. Gottesman, and V. Nikiforov.** 1992. Cooperation of GroEL/GroES and DnaK/DnaJ heat shock proteins in preventing protein misfolding in *Escherichia coli*. Proceedings of the National Academy of Sciences of the United States of America **89:**10341-4.
120. **Gray, T. E., and A. R. Fersht.** 1991. Cooperativity in ATP hydrolysis by GroEL is increased by GroES [published erratum appears in FEBS Lett 1992 Sep21;310(1):99]. FEBS Letters **292:**254-8.
121. **Green, H. A., and T. J. Donohue.** 2006. Activity of *Rhodobacter sphaeroides* RpoHII, a second member of the heat shock sigma factor family. Journal of Bacteriology **188:**5712-21.
122. **Grenert, J. P., B. D. Johnson, and D. O. Toft.** 1999. The importance of ATP binding and hydrolysis by hsp90 in formation and function of protein heterocomplexes. Journal of Biological Chemistry **274:**17525-33.
123. **Griffitts, J. S., and S. R. Long.** 2008. A symbiotic mutant of *Sinorhizobium meliloti* reveals a novel genetic pathway involving succinoglycan biosynthetic functions. Molecular Microbiology **67:**1292-306.
124. **Gross, C. A.** 1996. Function and regulation of the heat shock proteins, p. 1382-1399. *In* F. C. Neidhardt, R. Curtiss, III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed, vol. 1. ASM Press, Washington, D.C.
125. **Grossman, A. D., J. W. Erickson, and C. A. Gross.** 1984. The *htpR* gene product of *E. coli* is a sigma factor for heat-shock promoters. Cell **38:**383-90.
126. **Gruber, T. M., and D. A. Bryant.** 1997. Molecular systematic studies of eubacteria, using sigma70-type sigma factors of group 1 and group 2. Journal of Bacteriology **179:**1734-47.

127. **Guisbert, E., C. Herman, C. Z. Lu, and C. A. Gross.** 2004. A chaperone network controls the heat shock response in *E. coli*. Genes & Development **18:**2812-2821.
128. **Gunesekere, I. C., C. M. Kahler, D. R. Powell, L. A. S. Snyder, N. J. Saunders, J. I. Rood, and J. K. Davies.** 2006. Comparison of the RpoH-Dependent Regulon and General Stress Response in *Neisseria gonorrhoeae*. Journal of Bacteriology **188:**4769-4776.
129. **Gustafsson, P., K. Nordström, and S. Normark.** 1973. Outer penetration barrier of *Escherichia coli* K-12: kinetics of the uptake of gentian violet by wild type and envelope mutants. Journal of Bacteriology **116:**893-900.
130. **Györgypal, Z., A. Kondorosi, and E. Kondorosi.** 1991. Diverse Signal Sensitivity of NodD Protein Homologs from Narrow and Broad Host Range Rhizobia. Molecular Plant-Microbe Interactions **4:**356-64.
131. **Harley, C. B., and R. P. Reynolds.** 1987. Analysis of *E. coli* promoter sequences. Nucleic Acids Res **15:**2343-61.
132. **Hartwig, U. A., C. A. Maxwell, C. M. Joseph, and D. A. Phillips.** 1990. Effects of alfalfa nod gene-inducing flavonoids on *nodABC* transcription in *Rhizobium meliloti* strains containing different *nodD* genes. Journal of Bacteriology **172:**2769-73.
133. **Heidstra, R., W. C. Yang, Y. Yalcin, S. Peck, A. M. Emons, A. van Kammen, and T. Bisseling.** 1997. Ethylene provides positional information on cortical cell division but is not involved in Nod factor-induced root hair tip growth in *Rhizobium*-legume interaction. Development **124:**1781-7.
134. **Herman, C., D. Thévenet, R. d'Ari, and P. Bouloc.** 1995. Degradation of σ^{32}, the heat shock regulator in *Escherichia coli*, is governed by HflB. Proceedings of the National Academy of Sciences of the United States of America **92:**3516-3520.
135. **Herouart, D., S. Sigaud, S. Moreau, P. Frendo, D. Touati, and A. Puppo.** 1996. Cloning and characterization of the *katA* gene of *Rhizobium meliloti* encoding a hydrogen peroxide-inducible catalase. Journal of Bacteriology **178:**6802-9.
136. **Higgins, D. G., and P. M. Sharp.** 1989. Fast and sensitive multiple sequence alignments on a microcomputer. Computer Applications in the Biosciences: CABIOS **5:**151-3.
137. **Honma, M. A., and F. M. Ausubel.** 1987. *Rhizobium meliloti* has three functional copies of the *nodD* symbiotic regulatory gene. Proceedings of the National Academy of Sciences of the United States of America **84:**8558-62.
138. **Horváth, I., A. Glatz, V. Varvasovszki, Z. Török, T. Páli, G. Balogh, E. Kovács, L. Nádasdi, S. Benkö, F. Joó, and L. Vígh.** 1998. Membrane physical state controls the signaling mechanism of the heat shock response in *Synechocystis* PCC 6803: identification of hsp17 as a "fluidity gene". Proceedings of the National Academy of Sciences of the United States of America **95:**3513-8.
139. **Horváth, I., G. Multhoff, A. Sonnleitner, and L. Vígh.** 2008. Membrane-associated stress proteins: More than simply chaperones. Biochimica et Biophysica Acta **[Epub ahead of print]**.
140. **Horwich, A. L., K. B. Low, W. A. Fenton, I. N. Hirshfield, and K. Furtak.** 1993. Folding *in vivo* of bacterial cytoplasmic proteins: role of GroEL. Cell **74:**909-917.
141. **Houry, W. A.** 2001. Mechanism of substrate recognition by the chaperonin GroEL. Biochemistry and Cell Biology **79:**569-77.
142. **Houry, W. A., D. Frishman, C. Eckerskorn, F. Lottspeich, and F. U. Hartl.** 1999. Identification of *in vivo* substrates of the chaperonin GroEL. Nature **402:**147-54.

143. **House, B. L., M. W. Mortimer, and M. L. Kahn.** 2004. New recombination methods for *Sinorhizobium meliloti* genetics. Applied and Environmental Microbiology **70:**2806-2815.
144. **Hunt, J. F., A. J. Weaver, S. J. Landry, L. Gierasch, and J. Deisenhofer.** 1996. The crystal structure of the GroES co-chaperonin at 2.8 Å resolution. Nature **379:**37-45.
145. **Hwang, K., D. W. Jeong, J. W. Lee, I. H. Kim, H. I. Chang, H. J. Kim, and I. Y. Kim.** 1999. Alteration of the NAD+/NADH ratio in CHO cells by stable transfection with human cytosolic glycerol-3-phosphate dehydrogenase: resistance to oxidative stress. Molecules and Cells **9:**429-35.
146. **Jackson, G. S., R. A. Staniforth, D. J. Halsall, T. Atkinson, J. J. Holbrook, A. R. Clarke, and S. G. Burston.** 1993. Binding and hydrolysis of nucleotides in the chaperonin catalytic cycle: implications for the mechanism of assisted protein folding. Biochemistry **32:**2554-63.
147. **Jain, V., M. Kumar, and D. Chatterji.** 2006. ppGpp: stringent response and survival. Journal of Microbiology **44:**1-10.
148. **Jamet, A., S. Sigaud, G. Van de Sype, A. Puppo, and D. Herouart.** 2005. The *katA* catalase gene is regulated by OxyR in both free-living and symbiotic *Sinorhizobium meliloti*. Journal of Bacteriology **187:**376-81.
149. **Jamet, A., S. Sigaud, G. Van de Sype, A. Puppo, and D. Hérouart.** 2003. Expression of the bacterial catalase genes during *Sinorhizobium meliloti-Medicago sativa* symbiosis and their crucial role during the infection process. Molecular Plant-Microbe Interactions **16:**217-225.
150. **Janaszak, A., W. Majczak, B. Nadratowska, A. Szalewska-Palasz, G. Konopa, and A. Taylor.** 2007. A sigma54-dependent promoter in the regulatory region of the *Escherichia coli rpoH* gene. Microbiology **153:**111-23.
151. **Jefferson, R. A., S. M. Burgess, and D. Hirsh.** 1986. β-Glucuronidase from *Escherichia coli* as a gene-fusion marker. Proceedings of the National Academy of Sciences of the United States of America **83:**8447-51.
152. **JGI** 2005, posting date. *Rhodospirillum rubrum* genome. [Online.]
153. **Joo, D. M., A. Nolte, R. Calendar, Y. N. Zhou, and D. J. Jin.** 1998. Multiple regions on the *Escherichia coli* heat shock transcription factor σ^{32} determine core RNA polymerase binding specificity. Journal of Bacteriology **180:**1095-1102.
154. **Juhnke, H., B. Krems, P. Kotter, and K. D. Entian.** 1996. Mutants that show increased sensitivity to hydrogen peroxide reveal an important role for the pentose phosphate pathway in protection of yeast against oxidative stress. Molecular & General Genetics **252:**456-64.
155. **Kallipolitis, B. H., and P. Valentin-Hansen.** 1998. Transcription of *rpoH*, encoding the *Escherichia coli* heat-shock regulator sigma32, is negatively controlled by the cAMP-CRP/CytR nucleoprotein complex. Molecular Microbiology **29:**1091-9.
156. **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. S., M. Sugimoto, C. Takeuchi, M. Yamada, and S. Tabata.** 2000. Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. DNA Research **7:**331-338.

157. **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 Research **9:**189-197.

158. **Kanemori, M., K. Nishihara, H. Yanagi, and T. Yura.** 1997. Synergistic roles of HslVU and other ATP-dependent proteases in controlling *in vivo* turnover of sigma32 and abnormal proteins in *Escherichia coli*. Journal of Bacteriology **179:**7219-25.

159. **Kanemori, M., H. Yanagi, and T. Yura.** 1999. Marked instability of the sigma(32) heat shock transcription factor at high temperature. Implications for heat shock regulation. Journal of Biological Chemistry **274:**22002-7.

160. **Karls, R. K., J. Brooks, P. Rossmeissl, J. Luedke, and T. J. Donohue.** 1998. Metabolic roles of a *Rhodobacter sphaeroides* member of the σ^{32} family. Journal of Bacteriology **180:**10-19.

161. **Karunakaran, K. P., Y. Noguchi, T. D. Read, A. Cherkasov, J. Kwee, C. Shen, C. C. Nelson, and R. C. Brunham.** 2003. Molecular analysis of the multiple GroEL proteins of Chlamydiae. Journal of Bacteriology **185:**1958-1966.

162. **Kaufusi, P. H., L. S. Forsberg, P. Tittabutr, and D. Borthakur.** 2004. Regulation of exopolysaccharide synthesis in *Rhizobium* sp. strain TAL1145 involves an alternative sigma factor gene, *rpoH2*. Microbiology **150:**3473-3482.

163. **Kawe, M., and A. Plückthun.** 2006. GroEL walks the fine line: the subtle balance of substrate and co-chaperonin binding by GroEL. A combinatorial investigation by design, selection and screening. Journal of Molecular Biology **357:**411-26.

164. **Kerner, M. J., D. J. Naylor, Y. Ishihama, T. Maier, H. C. Chang, A. P. Stines, C. Georgopoulos, D. Frishman, M. Hayer-Hartl, M. Mann, and F. U. Hartl.** 2005. Proteome-wide analysis of chaperonin-dependent protein folding in *Escherichia coli*. Cell **122:**209-20.

165. **Kirby, A. C., S. Meghji, S. P. Nair, P. White, K. Reddi, T. Nishihara, K. Nakashima, A. C. Willis, R. Sim, M. Wilson, and et al.** 1995. The potent bone-resorbing mediator of *Actinobacillus actinomycetemcomitans* is homologous to the molecular chaperone GroEL. Journal of Clinical Investigation **96:**1185-94.

166. **Krummenacher, P., and F. Narberhaus.** 2000. Two genes encoding a putative multidrug efflux pump of the RND/MFP family are cotranscribed with an *rpoH* gene in *Bradyrhizobium japonicum*. Gene **241:**247-54.

167. **Kusukawa, N., and T. Yura.** 1988. Heat shock protein GroE of *Escherichia coli*: key protective roles against thermal stress. Genes & Development **2:**874-882.

168. **Lagares, A., G. Caetano-Anollés, K. Niehaus, J. Lorenzen, H. D. Ljunggren, A. Pühler, and G. Favelukes.** 1992. A *Rhizobium meliloti* lipopolysaccharide mutant altered in competitiveness for nodulation of alfalfa. Journal of Bacteriology **174:**5941-5952.

169. **Landick, R., V. Vaughn, E. T. Lau, R. A. VanBogelen, J. W. Erickson, and F. C. Neidhardt.** 1984. Nucleotide sequence of the heat shock regulatory gene of *E. coli* suggests its protein product may be a transcription factor. Cell **38:**175-82.

170. **Lee, W. T., K. C. Terlesky, and F. R. Tabita.** 1997. Cloning and characterization of two *groESL* operons of *Rhodobacter sphaeroides*: transcriptional regulation of the heat-induced *groESL* operon. Journal of Bacteriology **179:**487-495.

171. **Leigh, J. A., E. R. Signer, and G. C. Walker.** 1985. Exopolysaccharide-deficient mutants of *Rhizobium meliloti* that form ineffective nodules. Proceedings of the National Academy of Sciences of the United States of America **82:**6231-5.
172. **Lemaux, P. G., S. L. Herendeen, P. L. Bloch, and F. C. Neidhardt.** 1978. Transient rates of synthesis of individual polypeptides in *E. coli* following temperatures shifts. Cell **13:**427-34.
173. **Lerouge, P., P. Roche, C. Faucher, F. Maillet, G. Truchet, J. C. Promé, and J. Dénarié.** 1990. Symbiotic host-specificity of *Rhizobium meliloti* is determined by a sulphated and acylated glucosamine oligosaccharide signal. Nature **344:**781-4.
174. **Liberek, K., T. P. Galitski, M. Zylicz, and C. Georgopoulos.** 1992. The DnaK chaperone modulates the heat shock response of *Escherichia coli* by binding to the sigma 32 transcription factor. Proceedings of the National Academy of Sciences of the United States of America **89:**3516-20.
175. **Liberek, K., and C. Georgopoulos.** 1993. Autoregulation of the *Escherichia coli* heat shock response by the DnaK and DnaJ heat shock proteins. Proceedings of the National Academy of Sciences of the United States of America **90:**11019-23.
176. **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.
177. **Lin, Z., D. Madan, and H. S. Rye.** 2008. GroEL stimulates protein folding through forced unfolding. Nature Structural & Molecular Biology **15:**303-11.
178. **Lin, Z., and H. S. Rye.** 2006. GroEL-mediated protein folding: making the impossible, possible. Critical Reviews in Biochemistry and Molecular Biology **41:**211-39.
179. **Losick, R., and P. Stragier.** 1992. Crisscross regulation of cell-type-specific gene expression during development in *B. subtilis*. Nature **355:**601-4.
180. **Macario, A. J., C. B. Dugan, and E. Conway de Macario.** 1991. A *dnaK* homolog in the archaebacterium *Methanosarcina mazei* S6. Gene **108:**133-7.
181. **MacGregor, B. J., R. K. Karls, and T. J. Donohue.** 1998. Transcription of the *Rhodobacter sphaeroides cycA* P1 promoter by alternate RNA polymerase holoenzymes. Journal of Bacteriology **180:**1-9.
182. **Margolin, W., and S. R. Long.** 1994. *Rhizobium meliloti* contains a novel second homolog of the cell division gene *ftsZ*. Journal of Bacteriology **176:**2033-43.
183. **Marketon, M. M., and J. E. González.** 2002. Identification of two quorum-sensing systems in *Sinorhizobium meliloti*. Journal of Bacteriology **184:**3466-3475.
184. **Martin, M. O., and S. R. Long.** 1984. Generalized transduction in *Rhizobium meliloti*. Journal of Bacteriology **159:**125-9.
185. **Matamoros, M. A., D. A. Dalton, J. Ramos, M. R. Clemente, M. C. Rubio, and M. Becana.** 2003. Biochemistry and Molecular Biology of Antioxidants in the Rhizobia-Legume Symbiosis. Plant Physiology **133:**499-509.
186. **McCarty, J. S., S. Rüdiger, H.-J. Schönfeld, J. Schneider-Mergener, K. Nakahigashi, T. Yura, and B. Bukau.** 1996. Regulatory region C of the *E. coli* heat shock transcription factor, σ^{32}, constitutes a DnaK binding site and is conserved among eubacteria. Journal of Molecular Biology **256:**829-837.
187. **McNally, D., and M. A. Fares.** 2007. *In silico* identification of functional divergence between the multiple *groEL* gene paralogs in *Chlamydiae*. BMC Evolutionary Biology **7:**81-93.

188. **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. Journal of Bacteriology **149:**114-22.
189. **Meisenzahl, A. C., L. Shapiro, and U. Jenal.** 1997. Isolation and characterization of a xylose-dependent promoter from *Caulobacter crescentus*. Journal of Bacteriology **179:**592-600.
190. **Mellor, R. B.** 1989. Bacteroids in the *Rhizobium*-legume symbiosis inhabit a plant internal lytic compartment: implications for other microbial endosymbioses. Journal of Experimental Botany **40:**831-9.
191. **Mergaert, P., K. Nikovics, Z. Kelemen, N. Maunoury, D. Vaubert, A. Kondorosi, and E. Kondorosi.** 2003. A novel family in *Medicago truncatula* consisting of more than 300 nodule-specific genes coding for small, secreted polypeptides with conserved cysteine motifs. Plant Physiology **132:**161-73.
192. **Mergaert, P., T. Uchiumi, B. Alunni, G. Evanno, A. Cheron, O. Catrice, A. E. Mausset, F. Barloy-Hubler, F. Galibert, A. Kondorosi, and E. Kondorosi.** 2006. Eukaryotic control on bacterial cell cycle and differentiation in the *Rhizobium*-legume symbiosis. Proceedings of the National Academy of Sciences of the United States of America **103:**5230-5.
193. **Merrick, M., J. Gibbins, and A. Toukdarian.** 1987. The nucleotide sequence of the sigma factor gene *ntrA* (*rpoN*) of *Azotobacter vinelandii*: analysis of conserved sequences in NtrA proteins. Molecular & General Genetics **210:**323-30.
194. **Merrick, M. J.** 1993. In a class of its own--the RNA polymerase sigma factor sigma 54 (sigma N). Molecular Microbiology **10:**903-9.
195. **Messer, W., and C. Weigel.** 1997. DnaA initiator--also a transcription factor. Molecular Microbiology **24:**1-6.
196. **Mitsui, H., T. Sato, Y. Sato, N. Ito, and K. Minamisawa.** 2004. *Sinorhizobium meliloti* $RpoH_1$ is required for effective nitrogen-fixing symbiosis with alfalfa. Molecular Genetics and Genomics **271:**416-425.
197. **Mogk, A., G. Homuth, C. Scholz, L. Kim, F. X. Schmid, and W. Schumann.** 1997. The GroE chaperonin machine is a major modulator of the CIRCE heat shock regulon of *Bacillus subtilis*. The EMBO Journal **16:**4579-4590.
198. **Moran, M. A., A. Buchan, J. M. González, J. F. Heidelberg, W. B. Whitman, R. P. Kiene, J. R. Henriksen, G. M. King, R. Belas, C. Fuqua, L. Brinkac, M. Lewis, S. Johri, B. Weaver, G. Pai, J. A. Eisen, E. Rahe, W. M. Sheldon, W. Ye, T. R. Miller, J. Carlton, D. A. Rasko, I. T. Paulsen, Q. Ren, S. C. Daugherty, R. T. Deboy, R. J. Dodson, A. S. Durkin, R. Madupu, W. C. Nelson, S. A. Sullivan, M. J. Rosovitz, D. H. Haft, J. Selengut, and N. Ward.** 2004. Genome sequence of *Silicibacter pomeroyi* reveals adaptations to the marine environment. Nature **432:**910-913.
199. **Morita, M., M. Kanemori, H. Yanagi, and T. Yura.** 1999. Heat-induced synthesis of σ^{32} in *Escherichia coli*: structural and functional dissection of *rpoH* mRNA secondary structure. Journal of Bacteriology **181:**401-410.
200. **Morita, M. T., M. Kanemori, H. Yanagi, and T. Yura.** 2000. Dynamic interplay between antagonistic pathways controlling the sigma 32 level in *Escherichia coli*. Proceedings of the National Academy of Sciences of the United States of America **97:**5860-5.

201. **Morita, M. T., Y. Tanaka, T. S. Kodama, Y. Kyogoku, H. Yanagi, and T. Yura.** 1999. Translational induction of heat shock transcription factor σ^{32}: evidence for a built-in RNA thermosensor. Genes & Development **13:**655-665.
202. **Mulligan, J. T., and S. R. Long.** 1989. A family of activator genes regulates expression of *Rhizobium meliloti* nodulation genes. Genetics **122:**7-18.
203. **Mur, L. A., P. Kenton, A. J. Lloyd, H. Ougham, and E. Prats.** 2008. The hypersensitive response; the centenary is upon us but how much do we know? Journal of Experimental Botany **59:**501-20.
204. **Nagai, H., R. Yano, J. W. Erickson, and T. Yura.** 1990. Transcriptional regulation of the heat shock regulatory gene *rpoH* in *Escherichia coli*: involvement of a novel catabolite-sensitive promoter. Journal of Bacteriology **172:**2710-5.
205. **Nagai, H., H. Yuzawa, M. Kanemori, and T. Yura.** 1994. A distinct segment of the σ^{32} polypeptide is involved in DnaK-mediated negative control of the heat shock response in *Escherichia coli*. Proceedings of the National Academy of Sciences of the United States of America **91:**10280-10284.
206. **Nagai, H., H. Yuzawa, and T. Yura.** 1991. Interplay of two cis-acting mRNA regions in translational control of σ^{32} synthesis during the heat shock response of *Escherichia coli*. Proceedings of the National Academy of Sciences of the United States of America **88:**10515-10519.
207. **Nakahigashi, K., E. Z. Ron, H. Yanagi, and T. Yura.** 1999. Differential and independent roles of a sigma(32) homolog (RpoH) and an HrcA repressor in the heat shock response of *Agrobacterium tumefaciens*. Journal of Bacteriology **181:**7509-15.
208. **Nakahigashi, K., H. Yanagi, and T. Yura.** 2001. DnaK chaperone-mediated control of activity of a sigma(32) homolog (RpoH) plays a major role in the heat shock response of *Agrobacterium tumefaciens*. Journal of Bacteriology **183:**5302-10.
209. **Nakahigashi, K., H. Yanagi, and T. Yura.** 1995. Isolation and sequence analysis of *rpoH* genes encoding σ^{32} homologs from gram negative bacteria: conserved mRNA and protein segments for heat shock regulation. Nucleic Acids Research **23:**4383-4390.
210. **Narberhaus, F.** 1999. Negative regulation of bacterial heat shock genes. Molecular Microbiology **31:**1-8.
211. **Narberhaus, F., R. Kaser, A. Nocker, and H. Hennecke.** 1998. A novel DNA element that controls bacterial heat shock gene expression. Molecular Microbiology **28:**315-23.
212. **Narberhaus, F., M. Kowarik, C. Beck, and H. Hennecke.** 1998. Promoter selectivity of the *Bradyrhizobium japonicum* transcription factors in vivo and in vitro. Journal of Bacteriology **180:**2395-2401.
213. **Narberhaus, F., P. Krummenacher, H.-M. Fischer, and H. Hennecke.** 1997. Three disparately regulated genes for σ^{32}-like transcription factors in *Bradyrhizobium japonicum*. Molecular Microbiology **24:**93-104.
214. **Narberhaus, F., W. Weiglhofer, H.-M. Fisher, and H. Hennecke.** 1996. The *Bradyrhizobium japonicum* $rpoH_1$ gene encoding a σ^{32}-like protein is part of a unique heat shock gene cluster together with $groESL_1$ and three small heat shock genes. Journal of Bacteriology **178:**5337-5346.
215. **Natera, S. H. A., N. Guerreiro, and M. A. Djordjevic.** 2000. Proteome analysis of differentially displayed proteins as a tool for the investigation of symbiosis. Molecular Plant-Microbe Interactions **13:**995-1009.

216. **Neidhardt, F. C., and R. A. VanBogelen.** 1981. Positive regulatory gene for temperature-controlled proteins in *Escherichia coli*. Biochemical and Biophysical Research Communications **100:**894-900.
217. **Nemoto, T., Y. Ohara-Nemoto, M. Denis, and J. A. Gustafsson.** 1990. The transformed glucocorticoid receptor has a lower steroid-binding affinity than the nontransformed receptor. Biochemistry **29:**1880-6.
218. **Newman, G., and E. Crooke.** 2000. DnaA, the initiator of *Escherichia coli* chromosomal replication, is located at the cell membrane. Journal of Bacteriology **182:**2604-10.
219. **Niehaus, K., D. Kapp, and A. Pühler.** 1993. Plant defense and delayed infection of alfalfa pseudonodules induced by an exopolysaccharide (EPS I)-deficient *Rhizobium meliloti* mutant. Planta **190:**4125-25.
220. **Niner, B. M., and A. M. Hirsch.** 1998. How many *Rhizobium* genes, in addition to *nod*, *nif/fix*, and *exo*, are needed for nodule development? Symbiosis **24:**51-102.
221. **Nonaka, G., M. Blankschien, C. Herman, C. A. Gross, and V. A. Rhodius.** 2006. Regulon and promoter analysis of the *E. coli* heat-shock factor, sigma32, reveals a multifaceted cellular response to heat stress. Genes & Development **20:**1776-89.
222. **Ogawa, J.** 1993. *Rhizobium meliloti nod* gene regulation: a role for GroEL in the activation of *nod* gene expression. Ph.D. thesis. Stanford University, Stanford, CA.
223. **Ogawa, J., and S. R. Long.** 1995. The *Rhizobium meliloti groELc* locus is required for regulation of early *nod* genes by the transcription activator NodD. Genes & Development **9:**714-729.
224. **Ojha, A., M. Anand, A. Bhatt, L. Kremer, W. R. Jacobs Jr., and G. F. Hatfull.** 2005. GroEL1: a dedicated chaperone involved in mycolic acid biosynthesis during biofilm formation in Mycobacteria. Cell **123:**861-873.
225. **Oke, V., and S. R. Long.** 1999. Bacterial genes induced within the nodule during the *Rhizobium*-legume symbiosis. Molecular Microbiology **32:**837-850.
226. **Oke, V., B. G. Rushing, E. J. Fisher, M. Moghadam Tabrizi, and S. R. Long.** 2001. Identification of the heat shock sigma factor RpoH and a second RpoH-like protein in *Sinorhizobium meliloti*. Microbiology **147:**2399-2408.
227. **Ono, Y., H. Mitsui, T. Sato, and K. Minamisawa.** 2001. Two RpoH homologs responsible for the expression of heat shock protein genes in *Sinorhizobium meliloti*. Molecular & General Genetics **264:**902-912.
228. **Ott, T., J. T. van Dongen, C. Günther, L. Krusell, G. Desbrosses, H. Vigeolas, V. Bock, T. Czechowski, P. Geigenberger, and M. K. Udvardi.** 2005. Symbiotic leghemoglobins are crucial for nitrogen fixation in legume root nodules but not for general plant growth and development. Current Biology **15:**531-5.
229. **Peetermans, W. E., C. J. Raats, J. A. Langermans, and R. van Furth.** 1994. Mycobacterial heat-shock protein 65 induces proinflammatory cytokines but does not activate human mononuclear phagocytes. Scandinavian Journal of Immunology **39:**613-7.
230. **Perret, X., W. J. Broughton, and S. Brenner.** 1991. Canonical ordered cosmid library of the symbiotic plasmid of *Rhizobium* species NGR234. Proceedings of the National Academy of Sciences of the United States of America **88:**1923-7.

231. **Perret, X., C. Freiberg, A. Rosenthal, W. J. Broughton, and R. Fellay.** 1999. High-resolution transcriptional analysis of the symbiotic plasmid of *Rhizobium* sp. NGR234. Molecular Microbiology **32:**415-425.
232. **Perret, X., C. Staehelin, and W. J. Broughton.** 2000. Molecular basis of symbiotic promiscuity. Microbiology and Molecular Biology Reviews **64:**180-201.
233. **Peters, N. K., J. W. Frost, and S. R. Long.** 1986. A plant flavone, luteolin, induces expression of *Rhizobium meliloti* nodulation genes. Science **233:**977-980.
234. **Peters, N. K., and S. R. Long.** 1988. Alfalfa Root Exudates and Compounds which Promote or Inhibit Induction of *Rhizobium meliloti* Nodulation Genes. Plant Physiology **88:**396-400.
235. **Phillips, D. A., C. M. Joseph, and C. A. Maxwell.** 1992. Trigonelline and stachydrine released from alfalfa seeds activate NodD2 protein in *Rhizobium meliloti*. Plant Physiology **99:**1526-31.
236. **Platt, R., C. Drescher, S. K. Park, and G. J. Phillips.** 2000. Genetic system for reversible integration of DNA constructs and *lacZ* gene fusions into the *Escherichia coli* chromosome. Plasmid **43:**12-23.
237. **Porte, F., J. P. Liautard, and S. Köhler.** 1999. Early acidification of phagosomes containing Brucella suis is essential for intracellular survival in murine macrophages. Infection and Immunity **67:**4041-7.
238. **Putnoky, P., A. Kereszt, T. Nakamura, G. Endre, E. Grosskopf, P. Kiss, and A. Kondorosi.** 1998. The *pha* gene cluster of *Rhizobium meliloti* involved in pH adaptation and symbiosis encodes a novel type of K^+ efflux system. Molecular Microbiology **28:**1091-101.
239. **Rae, A. L., P. Bonfante-Fasolo, and N. J. Brewin.** 1992. Structure and growth of infection threads in the legume symbiosis with *Rhizobium leguminosarum*. The Plant Journal **2:**385-95.
240. **Raffaelle, M., E. I. Kanin, J. Vogt, R. R. Burgess, and A. Z. Ansari.** 2005. Holoenzyme switching and stochastic release of sigma factors from RNA polymerase in vivo. Molecular Cell **20:**357-66.
241. **Ramu, S. K., H. M. Peng, and D. R. Cook.** 2002. Nod factor induction of reactive oxygen species production is correlated with expression of the early nodulin gene *rip1* in *Medicago truncatula*. Molecular Plant-Microbe Interactions **15:**522-8.
242. **Reed, J. W., and G. C. Walker.** 1991. Acidic conditions permit effective nodulation of alfalfa by invasion-deficient *Rhizobium meliloti exoD* mutants. Genes & Development **5:**2274-87.
243. **Reeve, W. G., L. Brau, J. Castelli, G. Garau, C. Sohlenkamp, O. Geiger, M. J. Dilworth, A. R. Glenn, J. G. Howieson, and R. P. Tiwari.** 2006. The *Sinorhizobium medicae* WSM419 *lpiA* gene is transcriptionally activated by FsrR and required to enhance survival in lethal acid conditions. Microbiology **152:**3049-59.
244. **Reisenauer, A., C. D. Mohr, and L. Shapiro.** 1996. Regulation of a heat shock sigma32 homolog in *Caulobacter crescentus*. Journal of Bacteriology **178:**1919-27.
245. **Retzlaff, C., Y. Yamamoto, P. S. Hoffman, H. Friedman, and T. W. Klein.** 1994. Bacterial heat shock proteins directly induce cytokine mRNA and interleukin-1 secretion in macrophage cultures. Infection and Immunity **62:**5689-93.

246. **Revel, H. R., B. L. Stitt, I. Lielausis, and W. B. Wood.** 1980. Role of the host cell in bacteriophage T4 development. I. Characterization of host mutants that block T4 head assembly. Journal of Virology **33:**366-76.
247. **Riccillo, P. M., C. I. Muglia, F. J. de Bruijn, A. J. Roe, I. R. Booth, and O. M. Aguilar.** 2000. Glutathione is involved in environmental stress responses in *Rhizobium tropici*, including acid tolerance. Journal of Bacteriology **182:**1748-53.
248. **Roberts, R. C., C. Toochinda, M. Avedissian, R. L. Baldini, S. L. Gomes, and L. Shapiro.** 1996. Identification of a *Caulobacter crescentus* operon encoding *hrcA*, involved in negatively regulating heat-inducible transcription, and the chaperone gene *grpE*. Journal of Bacteriology **178:**1829-1841.
249. **Robertson, J. G., and P. Lyttleton.** 1984. Division of peribacteroid membranes in root nodules of white clover. Journal of Cell Science **69:**147-57.
250. **Robson, R. L., and J. R. Postgate.** 1980. Oxygen and hydrogen in biological nitrogen fixation. Annual Review of Microbiology **34:**183-207.
251. **Roche, P., F. Debellé, F. Maillet, P. Lerouge, C. Faucher, G. Truchet, J. Dénarié, and J. C. Promé.** 1991. Molecular basis of symbiotic host specificity in *Rhizobium meliloti*: *nodH* and *nodPQ* genes encode the sulfation of lipooligosaccharide signals. Cell **67:**1131-43.
252. **Rodríguez-Quiñones, F., M. Maguire, E. J. Wallington, P. S. Gould, V. Yerko, J. A. Downie, and P. A. Lund.** 2005. Two of the three *groEL* homologues in *Rhizobium leguminosarum* are dispensable for normal growth. Archives of Microbiology **183:**253-265.
253. **Roseman, A. M., S. Chen, H. White, K. Braig, and H. R. Saibil.** 1996. The chaperonin ATPase cycle: mechanism of allosteric switching and movements of substrate-binding domains in GroEL. Cell **87:**241-51.
254. **Roth, L. E., and G. Stacey.** 1989. Bacterium release into host cells of nitrogen-fixing soybean nodules: the symbiosome membrane comes from three sources. European Journal of Cell Biology **49:**13-23.
255. **Rubio, M. C., E. K. James, M. R. Clemente, B. Bucciarelli, M. Fedorova, C. P. Vance, and M. Becana.** 2004. Localization of superoxide dismutases and hydrogen peroxide in legume root nodules. Molecular Plant-Microbe Interactions **17:**1294-305.
256. **Rusanganwa, E., and R. S. Gupta.** 1993. Cloning and characterization of multiple *groEL* chaperonin-encoding genes in *Rhizobium meliloti*. Gene **126:**67-75.
257. **Ruvkun, G. B., V. Sundaresan, and F. M. Ausubel.** 1982. Directed transposon Tn*5* mutagenesis and complementation analysis of *Rhizobium meliloti* symbiotic nitrogen fixation genes. Cell **29:**551-559.
258. **Santos, R., D. Hérouart, A. Puppo, and D. Touati.** 2000. Critical protective role of bacterial superoxide dismutase in *Rhizobium*-legume symbiosis. Molecular Microbiology **38:**750-759.
259. **Santos, R., D. Hérouart, S. Sigaud, D. Touati, and A. Puppo.** 2001. Oxidative burst in alfalfa-*Sinorhizobium meliloti* symbiotic interaction. Molecular Plant-Microbe Interactions **14:**86-89.
260. **Sasse-Dwight, S., and J. D. Gralla.** 1990. Role of eukaryotic-type functional domains found in the prokaryotic enhancer receptor factor sigma 54. Cell **62:**945-54.

261. **Sauviac, L., H. Philippe, K. Phok, and C. Bruand.** 2007. An extracytoplasmic function sigma factor acts as a general stress response regulator in *Sinorhizobium meliloti*. Journal of Bacteriology **189:**4204-4216.
262. **Schlesinger, M. J., and C. Ryan.** 1993. An ATP- and hsc70-dependent oligomerization of nascent heat-shock factor (HSF) polypeptide suggests that HSF itself could be a "sensor" for the cellular stress response. Protein Science **2:**1356-60.
263. **Schroeder, B. K., B. L. House, M. W. Mortimer, S. N. Yurgel, S. C. Maloney, K. L. Ward, and M. L. Kahn.** 2005. Development of a functional genomics platform for *Sinorhizobium meliloti*: construction of an ORFeome. Applied and Environmental Microbiology **71:**5858-5864.
264. **Schulz, A., and W. Schumann.** 1996. *hrcA*, the first gene of the *Bacillus subtilis dnaK* operon encodes a negative regulator of class I heat shock genes. Journal of Bacteriology **178:**1088-1093.
265. **Scopio, A., P. Johnson, A. Laquerre, and D. R. Nelson.** 1994. Subcellular localization and chaperone activities of *Borrelia burgdorferi* Hsp60 and Hsp70. Journal of Bacteriology **176:**6449-56.
266. **Segal, G., and E. Z. Ron.** 1996. Heat shock activation of the *groESL* operon of *Agrobacterum tumefaciens* and the regulatory roles of the inverted repeat. Journal of Bacteriology **178:**3634-3640.
267. **Segal, G., and E. Z. Ron.** 1995. The *dnaKJ* operon of *Agrobacterium tumefaciens*: transcriptional analysis and evidence for a new heat shock promoter. Journal of Bacteriology **177:**5952-5958.
268. **Servant, P., C. Thompson, and P. Mazodier.** 1993. Use of new *Escherichia coli/Streptomyces* conjugative vectors to probe the functions of the two *groEL*-like genes of *Streptomyces albus* G by gene disruption. Gene **134:**25-32.
269. **Shaw, S. L., and S. R. Long.** 2003. Nod factor inhibition of reactive oxygen efflux in a host legume. Plant Physiology **132:**2196-204.
270. **Shi, Y., D. D. Mosser, and R. I. Morimoto.** 1998. Molecular chaperones as HSF1-specific transcriptional repressors. Genes & Development **12:**654-66.
271. **Sieberer, B. J., A. C. Timmers, and A. M. Emons.** 2005. Nod factors alter the microtubule cytoskeleton in *Medicago truncatula* root hairs to allow root hair reorientation. Molecular Plant-Microbe Interactions **18:**1195-204.
272. **Sigaud, S., V. Becquet, P. Frendo, A. Puppo, and D. Hérouart.** 1999. Differential regulation of two divergent *Sinorhizobium meliloti* genes for HPII-like catalases during free-living growth and protective role of both catalases during symbiosis. Journal of Bacteriology **181:**1634-1639.
273. **Slamti, L., J. Livny, and M. K. Waldor.** 2007. Global gene expression and phenotypic analysis of a *Vibrio cholerae rpoH* deletion mutant. Journal of Bacteriology **189:**351-62.
274. **Smith, D. F., and D. O. Toft.** 1993. Steroid receptors and their associated proteins. Molecular Endocrinology **7:**4-11.
275. **Soupène, E., M. Foussard, P. Boistard, G. Truchet, and J. Batut.** 1995. Oxygen as a key developmental regulator of *Rhizobium meliloti* N_2-fixation gene expression within the alfalfa root nodule. Proceedings of the National Academy of Sciences of the United States of America **92:**3759-3763.

276. **Sternberg, N.** 1973. Properties of a mutant of *Escherichia coli* defective in bacteriophage lambda head formation (*groE*). I. Initial characterization. Journal of Molecular Biology **76:**1-23.
277. **Straus, D., W. Walter, and C. A. Gross.** 1990. DnaK, DnaJ, and GrpE heat shock proteins negatively regulate heat shock gene expression by controlling the synthesis and stability of sigma 32. Genes & Development **4:**2202-9.
278. **Straus, D. B., W. A. Walter, and C. A. Gross.** 1987. The heat shock response of *E. coli* is regulated by changes in the concentration of σ^{32}. Nature **329:**348-351.
279. **Swanson, J. A., J. T. Mulligan, and S. R. Long.** 1993. Regulation of *syrM* and *nodD3* in *Rhizobium meliloti*. Genetics **134:**435-444.
280. **Szeto, W. W., J. L. Zimmerman, V. Sundaresan, and F. M. Ausubel.** 1984. A Rhizobium meliloti symbiotic regulatory gene. Cell **36:**1035-1043.
281. **Takano, T., and T. Kakefuda.** 1972. Involvement of a bacterial factor in morphogenesis of bacteriophage capsid. Nature: New Biology **239:**34-7.
282. **Tilly, K., N. McKittrick, M. Zylicz, and C. Georgopoulos.** 1983. The dnaK protein modulates the heat-shock response of *Escherichia coli*. Cell **34:**641-6.
283. **Tilly, K., H. Murialdo, and C. Georgopoulos.** 1981. Identification of a second *Escherichia coli groE* gene whose product is necessary for bacteriophage morphogenesis. Proceedings of the National Academy of Sciences of the United States of America **78:**1629-33.
284. **Tilly, K., J. Spence, and C. Georgopoulos.** 1989. Modulation of stability of the *Escherichia coli* heat shock regulatory factor σ^{32}. Journal of Bacteriology **171:**1585-1589.
285. **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.
286. **Tittabutr, P., W. Payakapong, N. Teaumroong, N. Boonkerd, P. W. Singleton, and D. Borthakur.** 2006. The alternative sigma factor RpoH2 is required for salt tolerance in *Sinorhizobium* sp. strain BL3. Research in Microbiology **157:**811-8.
287. **Tiwari, R. P., W. G. Reeve, M. J. Dilworth, and A. R. Glenn.** 1996. Acid tolerance in *Rhizobium meliloti* strain WSM419 involves a two-component sensor-regulator system. Microbiology **142 (Pt 7):**1693-704.
288. **Tomoyasu, T., J. Gamer, B. Bukau, M. Kanemori, H. Mori, A. J. Rutman, A. B. Oppenheim, T. Yura, K. Yamanaka, H. Niki, S. Hiraga, and T. Ogura.** 1995. *Escherichia coli* FtsH is a membrane-bound, ATP-dependent protease which degrades the heat-shock transcription factor s^{32}. The EMBO Journal **14:**2551-2560.
289. **Tomoyasu, T., A. Mogk, H. Langen, P. Goloubinoff, and B. Bukau.** 2001. Genetic dissection of the roles of chaperones and proteases in protein folding and degradation in the *Escherichia coli* cytosol. Molecular Microbiology **40:**397-413.
290. **Török, Z., P. Goloubinoff, I. Horváth, N. M. Tsvetkova, A. Glatz, G. Balogh, V. Varvasovszki, D. A. Los, E. Vierling, J. H. Crowe, and L. Vígh.** 2001. *Synechocystis* HSP17 is an amphitropic protein that stabilizes heat-stressed membranes and binds denatured proteins for subsequent chaperone-mediated refolding. Proceedings of the National Academy of Sciences of the United States of America **98:**3098-103.
291. **Török, Z., I. Horváth, P. Goloubinoff, E. Kovács, A. Glatz, G. Balogh, and L. Vígh.** 1997. Evidence for a lipochaperonin: association of active protein-folding GroESL

oligomers with lipids can stabilize membranes under heat shock conditions. Proceedings of the National Academy of Sciences of the United States of America **94:**2192-7.

292. **Tsvetkova, N. M., I. Horváth, Z. Török, W. F. Wolkers, Z. Balogi, N. Shigapova, L. M. Crowe, F. Tablin, E. Vierling, J. H. Crowe, and L. Vígh.** 2002. Small heat-shock proteins regulate membrane lipid polymorphism. Proceedings of the National Academy of Sciences of the United States of America **99:**13504-9.
293. **Tusher, V. G., R. Tibshirani, and G. Chu.** 2001. Significance analysis of microarrays applied to the ionizing radiation response. Proceedings of the National Academy of Sciences of the United States of America **98:**5116-5121.
294. **van Brussel, A. A., R. Bakhuizen, P. C. van Spronsen, H. P. Spaink, T. Tak, B. J. Lugtenberg, and J. W. Kijne.** 1992. Induction of Pre-Infection Thread Structures in the Leguminous Host Plant by Mitogenic Lipo-Oligosaccharides of *Rhizobium*. Science **257:**70-72.
295. **van der Vies, S. M., A. A. Gatenby, and C. Georgopoulos.** 1994. Bacteriophage T4 encodes a co-chaperonin that can substitute for *Escherichia coli* GroES in protein folding. Nature **368:**654-656.
296. **van Helden, J., B. André, and J. Collado-Vides.** 2000. A web site for the computational analysis of yeast regulatory sequences. Yeast **16:**177-187.
297. **Vasse, J., F. de Billy, S. Camut, and G. Truchet.** 1990. Correlation between ultrastructural differentiation of bacteroids and nitrogen fixation in alfalfa nodules. Journal of Bacteriology **172:**4295-4306.
298. **Vinuesa, P., F. Neumann-Silkow, C. Pacios-Bras, H. P. Spaink, E. Martínez-Romero, and D. Werner.** 2003. Genetic analysis of a pH-regulated operon from *Rhizobium tropici* CIAT899 involved in acid tolerance and nodulation competitiveness. Molecular Plant-Microbe Interactions **16:**159-68.
299. **Voroshilova, V. A., B. Boesten, V. E. Tsyganov, A. Y. Borisov, I. A. Tikhonovich, and U. B. Priefer.** 2001. Effect of mutations in *Pisum sativum* L. genes blocking different stages of nodule development on the expression of late symbiotic genes in *Rhizobium leguminosarum* bv. *viciae*. Molecular Plant-Microbe Interactions **14:**471-6.
300. **Wade, J. T., D. C. Roa, D. C. Grainger, D. Hurd, S. J. Busby, K. Struhl, and E. Nudler.** 2006. Extensive functional overlap between sigma factors in *Escherichia coli*. Nature Structural & Molecular Biology **13:**806-14.
301. **Walker, S. A., and J. A. Downie.** 2000. Entry of *Rhizobium leguminosarum* bv. *viciae* into root hairs requires minimal nod factor specificity, but subsequent infection thread growth requires *nodO* or *nodE*. Molecular Plant-Microbe Interactions **13:**754-62.
302. **Wallington, E. J., and P. A. Lund.** 1994. *Rhizobium leguminosarum* contains multiple chaperonin (*cpn60*) genes. Microbiology **140:**113-22.
303. **Wang, J. D., C. Herman, K. A. Tipton, C. A. Gross, and J. S. Weissman.** 2002. Directed Evolution of Substrate-Optimized GroEL/S Chaperonins. Cell **111:**1027-1039.
304. **Wang, Q. P., and J. M. Kaguni.** 1989. A novel sigma factor is involved in expression of the *rpoH* gene of *Escherichia coli*. Journal of Bacteriology **171:**4248-53.
305. **Wang, Q. P., and J. M. Kaguni.** 1989. DnaA protein regulates transcriptions of the *rpoH* gene of *Escherichia coli*. Journal of Biological Chemistry **264:**7338-44.
306. **Wasson, A. P., F. I. Pellerone, and U. Mathesius.** 2006. Silencing the flavonoid pathway in *Medicago truncatula* inhibits root nodule formation and prevents auxin transport regulation by rhizobia. Plant Cell **18:**1617-29.

307. **Wawrzynow, A., D. Wojtkowiak, J. Marszalek, B. Banecki, M. Jonsen, B. Graves, C. Georgopoulos, and M. Zylicz.** 1995. The ClpX heat-shock protein of *Escherichia coli*, the ATP-dependent substrate specificity component of the ClpP-ClpX protease, is a novel molecular chaperone. The EMBO Journal **14:**1867-77.
308. **Wells, D. H., and S. R. Long.** 2002. The *Sinorhizobium meliloti* stringent response affects multiple aspects of symbiosis. Molecular Microbiology **43:**1115-27.
309. **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. Applied and Environmental Microbiology **65:**374-80.
310. **Wood, D., J. Setubal, R. Kaul, D. Monks, J. Kitajima, V. Okura, Y. Zhou, L. Chen, G. Wood, N. J. Almeida, L. Woo, Y. Chen, I. Paulsen, J. Eisen, P. Karp, D. S. Bovee, P. Chapman, J. Clendenning, G. Deatherage, W. Gillet, C. Grant, T. Kutyavin, R. Levy, M. 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. Zhao, M. Dolan, F. Chumley, S. Tingey, J. Tomb, M. Gordon, M. Olson, and E. Nester.** 2001. The genome of the natural genetic engineer Agrobacterium tumefaciens C58. Science **294:**2317-2323.
311. **Wu, J., and A. Newton.** 1996. Isolation, identification, and transcriptional specificity of the heat shock sigma factor sigma32 from *Caulobacter crescentus*. Journal of Bacteriology **178:**2094-101.
312. **Wu, J., and A. Newton.** 1997. The *Caulobacter* heat shock sigma factor gene rpoH is positively autoregulated from a sigma32-dependent promoter. Journal of Bacteriology **179:**514-21.
313. **Xu, Z., A. L. Horwich, and P. B. Sigler.** 1997. The crystal structure of the asymmetric GroEL-GroES-(ADP)7 chaperonin complex. Nature **388:**741-50.
314. **Yamamori, T., and T. Yura.** 1982. Genetic control of heat-shock protein synthesis and its bearing on growth and thermal resistance in *Escherichia coli* K-12. Proceedings of the National Academy of Sciences of the United States of America **79:**860-4.
315. **Yang, W. C., C. de Blank, I. Meskiene, H. Hirt, J. Bakker, A. van Kammen, H. Franssen, and T. Bisseling.** 1994. *Rhizobium* nod factors reactivate the cell cycle during infection and nodule primordium formation, but the cycle is only completed in primordium formation. Plant Cell **6:**1415-26.
316. **Yeh, K.-C., M. C. Peck, and S. R. Long.** 2002. Luteolin and GroESL modulate in vitro activity of NodD. Journal of Bacteriology **184:**525-530.
317. **Yifrach, O., and A. Horovitz.** 1995. Nested cooperativity in the ATPase activity of the oligomeric chaperonin GroEL. Biochemistry **34:**5303-8.
318. **Yifrach, O., and A. Horovitz.** 1994. Two lines of allosteric communication in the oligomeric chaperonin GroEL are revealed by the single mutation Arg196-->Ala. Journal of Molecular Biology **243:**397-401.
319. **Young, J. P., L. Crossman, A. Johnston, N. Thomson, Z. Ghazoui, K. Hull, M. Wexler, A. Curson, J. Todd, P. Poole, T. Mauchline, A. East, M. 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 Biology **7:**R34.

320. **Yuan, Z. C., R. Zaheer, and T. M. Finan.** 2005. Phosphate limitation induces catalase expression in *Sinorhizobium meliloti*, *Pseudomonas aeruginosa* and *Agrobacterium tumefaciens*. Molecular Microbiology **58:**877-94.
321. **Yura, T., and K. Nakahigashi.** 1999. Regulation of the heat-shock response. Current Opinion in Microbiology **2:**153-158.
322. **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. Environmental Microbiology **5:**539-54.
323. **Zhao, K., M. Liu, and R. R. Burgess.** 2005. The global transcriptional response of *Escherichia coli* to induced σ^{32} protein involves σ^{32} regulon activation followed by inactivation and degradation of σ^{32} *in vivo*. The Journal of Biological Chemistry **280:**17758-17768.
324. **Zhou, Y.-N., N. Kusukawa, J. W. Erickson, C. A. Gross, and T. Yura.** 1988. Isolation and characterization of *Escherichia coli* mutants that lack the heat shock sigma factor σ^{32}. Journal of Bacteriology **170:**3640-3649.
325. **Zou, J., Y. Guo, T. Guettouche, D. F. Smith, and R. Voellmy.** 1998. Repression of heat shock transcription factor HSF1 activation by HSP90 (HSP90 complex) that forms a stress-sensitive complex with HSF1. Cell **94:**471-80.
326. **Zuber, U., and W. Schumann.** 1994. CIRCE, a novel heat shock element involved in regulation of heat shock operon *dnaK* of *Bacillus subtilis*. Journal of Bacteriology **176:**1359-63.
327. **Zweig, M., and D. J. Cummings.** 1973. Cleavage of head and tail proteins during bacteriophage T5 assembly: selective host involvement in the cleavage of a tail protein. Journal of Molecular Biology **80:**505-18.