Characterization of Sigma Factors of *Sinorhizobium meliloti* involved in Stress Response

Dissertation
zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften (Dr. rer. nat.)
der Fakultät für Biologie der Universität Bielefeld

vorgelegt von
Daniella Karine Cavalcanti de Lucena
aus Recife
Juli 2011

INDEX

A. SUMMARY

Alternative sigma factors play an important role in that they can respond with transcriptional activation to the presence of adverse conditions in their environment. The presence of multiple sigma factors is a ubiquitous feature in rhizobia. The genome of *Sinorhizobium meliloti* strain 1021 contains fourteen genes encoding for sigma factors, nine of which code for putative extracytoplasmic function sigma factors (RpoE) and two for heat shock sigma factors (RpoH). The purpose of this study was to gain detailed insight into the participation of sigma factors in the complex stress response system of *S. meliloti* 1021.

The RpoE sigma factors of *S. meliloti* share similarities in their predicted protein sequence and this could account for similar promoter selectivity and possible overlapping functions. Expression of sigma factors *rpoE5* and *rpoE1* was upregulated during oxidative stress. Induction of *rpoE5* expression was also observed under cold shock and heat shock, as well as under pH stress. A deficiency in growth was observed for sigma factor mutants for *rpoE2* and *rpoE5* in heat shock and oxidative stress conditions. However, an extreme growth deficiency phenotype was only observed for the *rpoH1* mutant at pH stress conditions.

At neutral pH, an upregulation of the rhizobactin biosynthesis operon was observed for the *rpoH1* mutant. Analyses of the promoter regions of the genes involved in rhizobactin biosynthesis showed that there are possible binding sites for RpoH1 and RpoE2 in the upstream regions of *rhbA*, the first gene in the rhizobactin biosynthesis operon. Also, a reduction in *rhbA* expression was observed for the *rpoH1 rpoE5* double mutant in comparison to the *rpoH1* single mutant. This indicates that alternative complementarity between different sigma factors could play a role in the regulation of rhizobactin at neutral pH.

The involvement of RpoH1 in pH stress response was further analysed by time-course microarray analyes, which lead to the classification of three groups of genes, which were transcriptionally regulated in an RpoH1-independent, an RpoH1-dependent or in a complex manner. Genes regulated in an RpoH1-dependent manner are known to be involved in stress and heat shock response, like *ibpA, grpE* and *groEL*. The promoter consensus binding site for RpoH1 was identified in a number of the genes classified as being RpoH1-dependent upon pH stress, as well as for the *rpoH1* gene itself. This study provided clear evidence that the sigma factor RpoH1 plays a major role in the pH stress response of *S. meliloti*.

B. INTRODUCTION

1. RHIZOBIA AND BIOLOGICAL NITROGEN FIXATION

1.1 BIOLOGICAL NITROGEN FIXATION

Nitrogen is essential for every living being, for it is a main component of nucleic acids and proteins. Even though nitrogen is among the most abundant elements on Earth, it is also a critical limiting element for growth of most plants due to its unavailability. In fact, when it comes to plant growth, only sunlight and water are more important than nitrogen [146]. Biological nitrogen fixation has commanded the attention of scientists concerned with plant mineral nutrition, and it has been exploited extensively in agricultural practice. All of the nutritional nitrogen required by humans is obtained either directly or indirectly from plants. Plants acquire nitrogen from the soil, through natural or industrial fertilizers; and from the atmosphere through biological nitrogen fixation [160].

Biological nitrogen fixation is a fundamentally important source of fixed nitrogen. Not only it makes atmospheric nitrogen available, but it also relieves ecosystem nitrogen limitation. Organisms that can fix nitrogen, that is, that can convert the stable nitrogen gas in the atmosphere into a biologically utilizable form, all belong to the prokaryote phylum. Those microorganisms catalyze biological nitrogen fixation with the enzyme nitrogenase, which has been highly conserved throughout evolution, and convert nitrogen to ammonia [160, 161]. The symbiotic nitrogen fixation by legumes is a major contributor to the combined nitrogen pool in the biosphere. Legumes are able to use nitrogen by establishing symbiosis with nitrogen-fixing bacteria, with the beneficial exchange of nutrients between the partners. In turn, the plant provides bacteria with carbon compounds that fuel the energy-expensive process of nitrogen fixation [89, 115]. Nitrogen-fixing symbiosis involves most of the 18,000 legume species with an expanding collection of α- and β-proteobacteria. The partnership of nitrogen-fixing symbiosis is of major ecological importance, as it occurs in all continents and accounts for a fourth of the nitrogen fixed annually on earth [89].

1.2 SYMBIOSIS BETWEEN RHIZOBIA AND LEGUME PLANTS

Rhizobial bacteria are soil based microorganisms which establish symbiotic relationships with legume plants. They currently belong to 12 genera and more than 90 species of α- and β-

proteobacteria [151]. All α-rhizobial genera belong to the Rhizobiales order, whereas β-rhizobial genera belong to the Burkholderiales order. To date, only one rhizobium, *Azorhizobium caulinodans*, is known to be a genuine diazotroph, able to grow ex planta at the expense of fixed nitrogen. Rhizobia represent a group that mostly comprises members from the genera *Allorhizobium*, *Azorhizobium*, *Bradyrhizobium*, *Mesorhizobium*, *Rhizobium* and *Sinorhizobium*. *Rhizobium* species exist as either free-living bacteria or in symbiosis with plant hosts, and the symbiotic interactions can be very diverse [89] (Figure 1). Many phenotypic variations regarding the localization, shape and anatomy of the plants, as well as the infection mode and differentiation status of endosymbionts are encountered in nature, most of which are under plant control (Figure 1). Moreover, various bacterial genetic strategies lead to symbiosis, thus supporting evidence that rhizobia have evolved largely independently. In this section, a description of symbiosis between rhizobia and legumes will be given, with emphasis in the formation of indeterminate plant nodules in the model bacterium *Sinorhizobium meliloti*.

The establishment of effective symbiosis depends on coordinated signal and nutrient exchanges between the symbiotic partners. The chemical signaling between legume plants and most rhizobial bacteria initiates a process called root nodule formation [39]. In the earliest stages of the interaction, plant roots liberate chemical compounds, such as flavonoids and betaines, which are sensed by their compatible rhizobial counterparts [27]. Those plant-derived compounds stimulate the bacteria to produce and secrete chemical compounds denominated Nod factors, which are then perceived by the plant partner. Bacteria are trapped and tunneled in the plant root by root hair curling via inward growth of the plant cell wall and finally enter the root hair tip by the formation of an infection thread, which is induced by the production not only of Nod factors, but also of exopolysaccharides (EPS) [12, 49]. This process results in the development of a root nodule that the bacteria colonize intracellularly. When an infection thread reaches a suitable plant cell, the bacteria are released into the plant cell cytoplasm where they undergo terminal differentiation into specialized cell types called bacteroids, which are able to convert dinitrogen into ammonium [47]. Root nodules are classified as indeterminate because they have a persistent meristem in which plant cells continue to divide throughout nodule development [15] (Figure 2A). Inside the plant cell, a structure called symbiosome is formed by endocellular nitrogen-fixing rhizobia surrounded by a plant-derived peribacteroid membrane. Some rhizobia that are adapted to aquatic or semi-

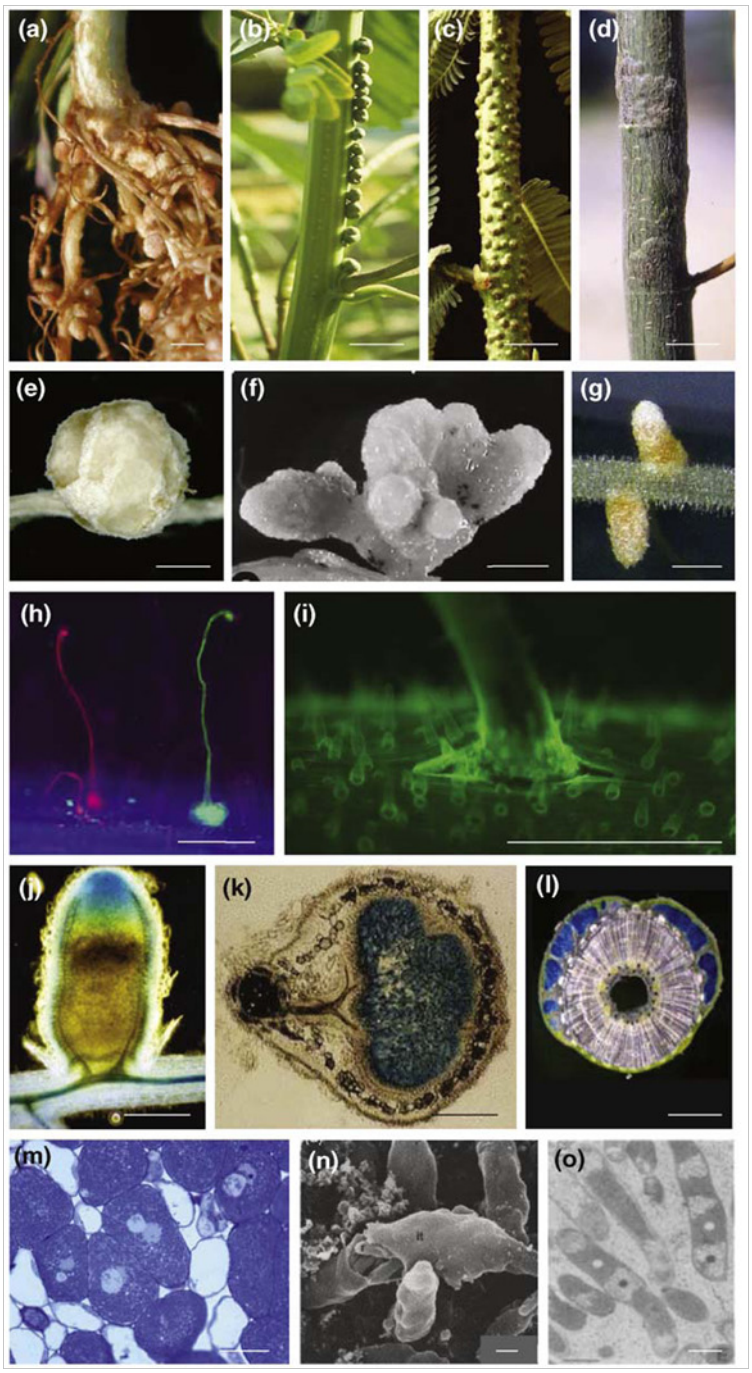


Figure 1. Morphological variations in rhizobium–legume symbiosis. Rhizobia in symbiosis with host plants form nodules on roots: a) *Cupriavidus taiwanensis-Mimosa pudica*; and on the stems of legumes: b) *Azorhizobium caulinodans-Sesbania rostrata*, c) *Bradyrhizobium sp.-Aeschynomene afraspera*, d) *Bradyrhizobium sp.-Aeschynomene sensitiva*. Nodules display various shapes: round e) *Sinorhizobium fredii*-soybean; coralloid f) *Methylobacterium nodulans-Crotalaria perrottetii*; or elongated g) *S. meliloti-Medicago sativa*. Infection proceeds via infection threads: h) *S. meliloti-M. sativa* or crack-entry at lateral roots: i) *Bradyrhizobium sp.–Aeschynomene indica*. Nodules are indeterminate j) as in *M. sativa* or determinate as in k) Lotus in *A. sensitiva* nodules *(legend continues)*

l) *Bradyrhizobium sp*. Fixation of N_2 occurs within infected nodule cells m) *C. taiwanensis–M.pudica* or fixation threads n) caesalpinioid legumes. (o) Dead-end highly differentiated *S. melilo*ti bacteroids in *M. sativa* nodules. Scale bars: (a–d) 1 cm, (e–g, l) 1mm, (i–k) 500 mm, (h, m) 50 mm, (n, o) 1 mm. Source: Masson-Boivin, 2009.

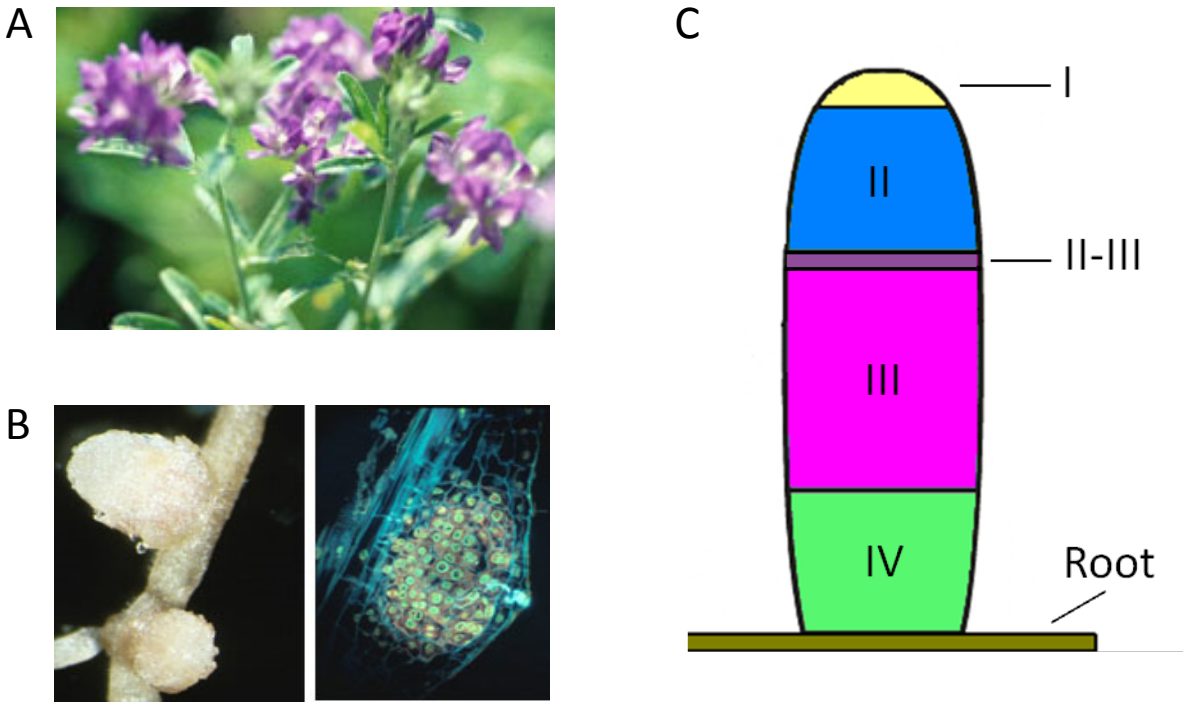


Figure 2. Symbionts *Medicago sativa* and *Sinorhizobium meliloti*. A) Picture of the host plant *Medicago sativa* (Alfalfa). B) Left : mature, nitrogen-fixing nodules on an alfalfa root. Right panel: Thin section of alfalfa nodule stained with DAPI and acridine orange. An infection thread is visible as a white-blue fluorescing structure in the outer layers (right side) [photo by Mark Dudley]. C) Representation of a root nodule with the different zones. Zone I is formed by meristem cells; zone II is the infection zone; zone III is the nitrogen fixation zone; and zone IV is the senescence zone. Source: Stanford.edu.

-aquatic tropical legumes do not penetrate the plants through root hair infection, but worm their way into emerging nodules at the loose cellular junctions of emerging lateral roots, in a process known as crack-entry (Figure 1i).

Moreover, the host plant manipulates the cell fate of the endosymbiotic bacteria, which undergo striking morphological changes, such as cell elongation coupled to genome amplification and the loss of reproductive capacity. In the legume *Medicago truncatula*, the bacteria undergo an irreversible differentiation mediated by plant factors called nodule-specific cysteine-rich peptides (NCR), which are targeted to the bacteria and enter the bacterial membrane and cytosol [145]. In the nodules, the endosymbionts reduce dinitrogen into ammonia that is subsequently assimilated by the host plant. All rhizobia rely on the most common form of nitrogenase, molybdenum-nitrogenase, for nitrogen fixation. This nitrogenase endures two major drawbacks: a high-energy requirement with a minimal stoichiometry of 16 mol of ATP for each mole of nitrogen reduced, and an extreme sensitivity to oxygen, which is in contradiction with the strict aerobic character of rhizobia. To overcome these limitations, legume plants fuel the endosymbiotic bacteria with plentiful photosynthetic derivatives and provide within nodule cells the nearly anoxygenic environment required for nitrogen fixation by producing leghemoglobin, a protein that binds oxygen with high affinity and reduces oxygen concentration around bacteroids [89, 112].

.

1.3 THE NITROGEN-FIXING SOIL BACTERIUM *SINORHIZOBIUM MELILOTI*

The focus of this research was the bacterium *Sinorhizobium meliloti*, which is the microbial symbiont of the agronomically important legume alfalfa (*Medicago* sp.). It is regarded as a model bacterium for a number of reasons, including the fact that it is a fast growing organism that is readily amenable to molecular studies. *S. meliloti* can undergo symbiosis with legumes from the genera *Medicago*, *Melilotus* and *Trigonella*. The diploid and autogamous lucerne *Medicago sativa* (Figure 1B) has been selected as a model legume, making the *S. meliloti*-*M. sativa* system a classic example for the study of *Rhizobium*-legume interactions. *S. meliloti* has three circular replicons: a chromosome (3.65 Mb) that encodes most of the housekeeping and essential genes [45] and two megaplasmids, pSymA (1.4 Mb) and pSymB (1.7 Mb) [7, 38]. Global changes in gene expression showed that *S. meliloti* adapts to abiotic and symbiotic environments in unique ways, and that the three bacterial replicons serve specialized roles in this adaptation [8].

In the symbiosis of *S. meliloti* with Medicago and with other legumes of the galegoid group, bacteroid differentiation is irreversible and the endosymbiotic bacteria do not survive nodule senescence [89, 92] (Figure 1n). In the nodules, four different zones contain bacteria in distinct stages of differentiation (Figure 2A). Zone I is formed by plant meristem cells and usually has no bacteria; zone II is the infection zone, where the bacteria enter the plant cells through infection threads; the interzone II-III is formed by elongated bacteroids that do not fix nitrogen, but already express the nitrogen fixation genes; nitrogen fixation takes place in zone III, which contains mature and fully differentiated bacteroids; and zone IV is the senescence zone in which degradation of the cells takes place [147].

2. ENVIRONMENTAL STRESS IN BACTERIA

All free-living microorganisms are frequently exposed to environmental stress, among them limitations in nutrient supply, sudden changes in osmolarity, and up- or downshifts in temperature. A rapid protective response to harmful conditions requires the coordinated induction of multiple genes [75]. Bacterial stress response can be defined as a cascade of alterations in gene expression and protein activity for the purpose of surviving extreme, rapidly changing and potentially damaging conditions sensed by the bacteria, and which results in the cells becoming broadly stress-resistant or eliminating the stress agent and mediating repair of cell injury [48]. Under stress, many bacteria exhibit global changes of gene expression that may result in altered metabolism and physiology, as well as enhanced resistance to multiple stresses. Common types of environmental stress conditions encountered by bacteria will be discussed in the following, with emphasis in how these stress conditions are faced by rhizobia.

2.1 HEAT SHOCK

Temperature is among the most important of the parameters that microbes monitor and their physiology needs to be readily readjusted in response to sudden temperature changes. When the ambient temperature rises to potentially harmful levels, cells mount a protective stress response called heat shock response. Often, the consequences of a sudden temperature shift are detected by the accumulation of denatured proteins, but primary thermosensors include DNA, RNA, proteins and lipids [75].

The heat shock response is classically defined as the cellular response to temperature increase. A major component of this response is the upregulation of a set of proteins termed heat shock proteins, whose function is to prevent heat-induced cellular damage. For instance, tertiary and quaternary structures of proteins are very susceptible to temperature changes and temperature shifts can easily shift a protein from an active to inactive conformation, or induce the disassembly of a dimer into a monomer. Heat shock proteins cover a wide range of cellular functions in maintaining protein and membrane homeostasis and nucleic acid topology at high temperatures. The heat shock proteins include molecular chaperones that aid in protein folding, and proteases that degrade unwanted or damaged proteins [43]. The rapid upregulation of chaperones and proteases during the heat shock response restores an appropriate protein-folding environment in the cell. Consistent with this idea, many other treatments that destabilize folded proteins or make it more difficult for nascent proteins to fold also activate this response [56].

The optimum temperature range for most rhizobia is 25–31ºC. However, high tolerance to heat shock is common in rhizobia. Their upper temperature limits are between 32º and 47ºC [1]. Microarray data in *S. meliloti* reported the upregulation of 169 genes after heat shock, including genes coding for chaperones and other heat shock proteins [129]. Rhizobia nodulating common bean are able to grow at temperatures up to 44ºC while the maximum growth temperature for chickpea rhizobia is 40ºC [32]. Also, rhizobia are the only bacteria known to induce a multitude of small heat shock proteins (sHsps) upon temperature upshift [136]. The main function of sHsps is to bind denatured proteins and to maintain them in a folding competent state. *Bradyrhizobium japonicum*, the nitrogen-fixing root nodule symbiont of soybean, responds to a heat shock with the induction of at least 10 sHsps [103].

2.2 COLD SHOCK

Cold is yet another temperature condition that microorganisms frequently encounter in nature. It causes physiological problems that are different from heat stress. The obstacles include reduced enzyme activity, decreased membrane fluidity, and RNA structures that interfere with translation. The cell envelope is the first cellular compartment to come into contact with the external temperature. Exposure to cold stress drastically alters membrane properties, and this must be counteracted quickly in order to maintain membrane integrity and the critical function

of membrane proteins [66]. Therefore, membranes may act as thermosensors per se, transducing the signal via membrane-integrated proteins. Sensor kinases are likely candidates for this purpose. The induction of cold shock proteins enables efficient translation and maintains membrane integrity after a temperature downshift, as indirect signals of cold stress are known to be stalled ribosomes [75].

Escherichia coli cells have evolved a mechanism to cope with the effects caused by sudden lowering of the environmental temperature. When an exponential culture of *E. coli* is transferred from 37 °C to a temperature below 20 °C, the cells transiently stop growing and enter an acclimation phase in which the synthesis of a small set of cold shock proteins is induced. The list of the *E. coli* cold shock-induced proteins, so far identified mainly by proteomic approaches, is essentially constituted by nucleic acid–binding proteins involved in different cellular processes like RNA degradation, transcription, DNA replication, translation, and ribosome maturation, as well as by five members of the Csp family. Global transcription profiling during cold shock has revealed that the level of transcripts encoding molecular chaperones also increases after cold shock. Overall data suggest that the cold shock response is intended for dealing with unfavorable secondary structures of nucleic acids, which are expected to hinder basic functions such as transcription, ribosome assembly, and translation; opposing the cold shock-induced decrease in membrane fluidity; accumulating sugars displaying a protective effect against the low temperature, such as trehalose; and helping protein folding at low temperatures [48].

Some rhizobia display a large diversity in their tolerance to temperature stress and are able to grow better at low temperature than at high temperature [1]. In temperate climates, where low temperature limits the efficiency of the symbiosis, different species of *Rhizobium* show different thermoadaptation characteristics and produce heat and cold shock proteins at temperatures outside their normal growth ranges. Arctic and temperate strains of rhizobia respond similarly to cold shocks by synthesizing proteins under their minimal growth temperatures at freezing temperatures as low as -10 °C [26]. Rhizobia in cold shock temperatures reduce protein synthesis in comparison with that obtained under optimal temperatures, probably due to the inhibition of the initiation of translation and the accumulation of ribosomal particles [26].

2.3 OXIDATIVE STRESS

Reactive by-products of oxygen, such as superoxide anion radical, hydrogen peroxide, and the highly reactive hydroxyl radicals are generated continuously in cells grown aerobically. In *E. coli*, the respiratory chain can account for as much as 87% of the total H_2O_2 production. The production of harmful superoxide and hydroxyl radicals cause toxic effects which damage all macromolecules, including proteins, lipids and DNA [21]. A primary effect of lipid peroxidation is a decrease in membrane fluidity, which alters membrane properties and can disrupt membrane-bound proteins significantly. DNA is also a main target; active species attack both the base and the sugar moieties producing single- and double-strand breaks in the backbone and lesions that block replication [43]. Bacteria respond to oxidative stress by invoking two distinct stress responses, the peroxide stimulon and the superoxide stimulon, depending on whether the stress is mediated by peroxides or the superoxide anion. Those bacterial genetic responses to oxidative stress coordinate the expression of multiple genes and are controlled by two major transcriptional regulators (OxyR and SoxRS) [21]. The SoxR and SoxS proteins constitute a sensor-regulator system that senses superoxide and modulates gene expression. The SoxRS regulon responds to the intracellular accumulation of superoxide, triggered by redox-cycling agents. The redox-cycling agents produce superoxide at the expense of the oxidation of NADPH, decreasing the reducing capacity of the cell []. The hydrogen peroxide is scavenged in most microrganisms by peroxidases and catalases [21].

In rhizobia, oxidative stress may interfere at several steps in symbiosis and the maintenance of nitrogenase activity is subject to a delicate equilibrium. A high rate of respiration is necessary to supply the energy demands of the nitrogen reduction process in the nodules, but oxygen and reactive oxygen species irreversibly inactivate the nitrogenase complex [121]. A diffusion barrier in the cortex of nodules greatly limits permeability to oxygen, and the necessary oxygen is delivered by the plant oxygen carrier, leghaemoglobin, present exclusively in the nodule. Despite these strategies ensuring a low free oxygen concentration, the high rate of respiration inevitably results in there being large amounts of reactive oxygen species in the root nodule [57]. Reactive oxygen is also generated by bacteroids through the processes of respiration and nitrogen fixation and in plant cells by the oxidation of leghemoglobin. Therefore, efficient protection against oxidative stress is necessary for efficient nitrogen fixation and to delay senescence. In rhizobia, oxidative stress resistance proteins such as superoxide dismutase play a critical role in protecting the nitrogen fixation process [128].

2.4 pH STRESS

Bacterial cell function is extremely dependent on the maintenance of an appropriate intracellular pH. Although the precise relationship between pH homeostasis and acid tolerance in bacteria is not well understood, genetic and biochemical studies point to a central role for the maintenance of internal pH in bacteria that achieve acid tolerance [109].

Soil acidity is a major factor limiting legume growth and nitrogen fixation. It affects the exchange of molecular signals between rhizobia and their host, reducing nodulation, and hinders symbiotic development [109, 118, 158]. Some of the proteins localized in the symbiosome have an acidic pH optimum, which also suggests that bacteroids are located within an acidic compartment [15, 91]. Different species of rhizobia show different levels of tolerance to acidic pH. In general, rhizobia have a light alkaline intracellular pH [52, 109]. Moreover, rhizobia that present slow growth rates also show increased tolerance to low pH levels, in comparison to fast-growing rhizobia [52]. In *S. meliloti,* the intracellular pH has to be at least 6.5 to allow for cell growth [109] and the response to acidic pH stress is characterized by the differential expression of genes associated with various cellular functions, such as exopolysaccharide I biosynthesis and chemotaxis [61]. In *Rhizobium tropici* and *S. medicae*, the *lpiA* gene is transcriptionally up-regulated in response to acid shock, but is not essential for growth in acidic conditions [118]. In *S. medicae* und *R. leguminosarum* the ActS-ActR system is essential for growth in acidic pH stress [37, 143, 144].

2.5 OTHER STRESSES

2.5.1 Nutritional stress

The stringent response is caused by unfavorable nutrient conditions such as amino acid starvation and induces a global metabolic shift for the bacteria to adapt to changes in their environment. Guanosine tetraphosphate (ppGpp) molecules are rapidly synthesized in response to nutritional stress. Specifically, nutrient stress triggers the ribosome-associated RelA protein to catalyse the formation of guanosine pentaphosphate, pppGpp, using ATP. pppGpp is then hydrolysed to ppGpp, which associates directly with the RNA polymerase holoenzyme to control nutritional stress response [152]. In *E. coli*, the enzymes RelA and SpoT are triggered by different starvation signals to produce (p)ppGpp. In many Gram-positive bacteria this is carried out by RelA and two small homologs. An *S. meliloti relA*

mutant, which cannot produce ppGpp, cannot form nodules on plants and overproduces exopolysaccharides [133].

2.5.2 Osmotic Stress

Salt stress is one of the major environmental stress conditions and constitutes a severe problem facing the establishment of bacteria in their natural environment. Survival and growth in saline environments are the result of adaptive processes, such as ion transport and compartmentation, osmotic solute synthesis and accumulation. Several chemical substances from micro-organisms and plants are reported to be accumulated under salt stress to protect them against this stress. These include polyamines, polysaccharides, amino acids such as proline, serine, and glutamine, organic solutes, soluble sugars, and inorganic cations [78]. High salt concentration affects the growth of rhizobia and their nitrogen fixation capacity. Salt stress can directly impair the interactions between rhizobia and the host-plant, inhibiting nodule formation, or it can reduce the leghaemoglobin content inside the nodules, leading to the formation of ineffective nodules [17, 160]. Salt-tolerant rhizobia exposed to increased salinity can maintain equilibrium across the membrane by exclusion of salts and via accumulation in cytoplasm of compatibles and non-toxic solutes called osmoprotectants [2]. The growth of rhizobia under osmotic stress is linked to the ability to develop intrinsic mechanisms leading to the conservation of cell integrity. Salt tolerance in *S. meliloti* strain A5 is associated mainly with up-regulation of anoxidoreductase and heat shock proteins, and induction of a putative oligopeptide ABC transporter in salty medium [101].

3. SIGMA FACTOR FAMILIES AND THEIR ROLE IN TRANSCRIPTIONAL REGULATION

3.1 GENERAL REMARKS

Transcription is a primary indicator of the state of differentiation for any organism. It is the most fundamental process needed for regulation of cellular adaptation and it is carried out by DNA-dependent RNA polymerases (RNAPs). RNAPs are unambiguously closely related in structure and function across all kingdoms of life and also have conserved mechanisms [46]. Prokaryotic core RNAP has a molecular mass of 400 kDa and is composed of four distinct conserved subunits: β, β', ω and an α dimer (Figure 3). The RNAP core enzyme possesses the

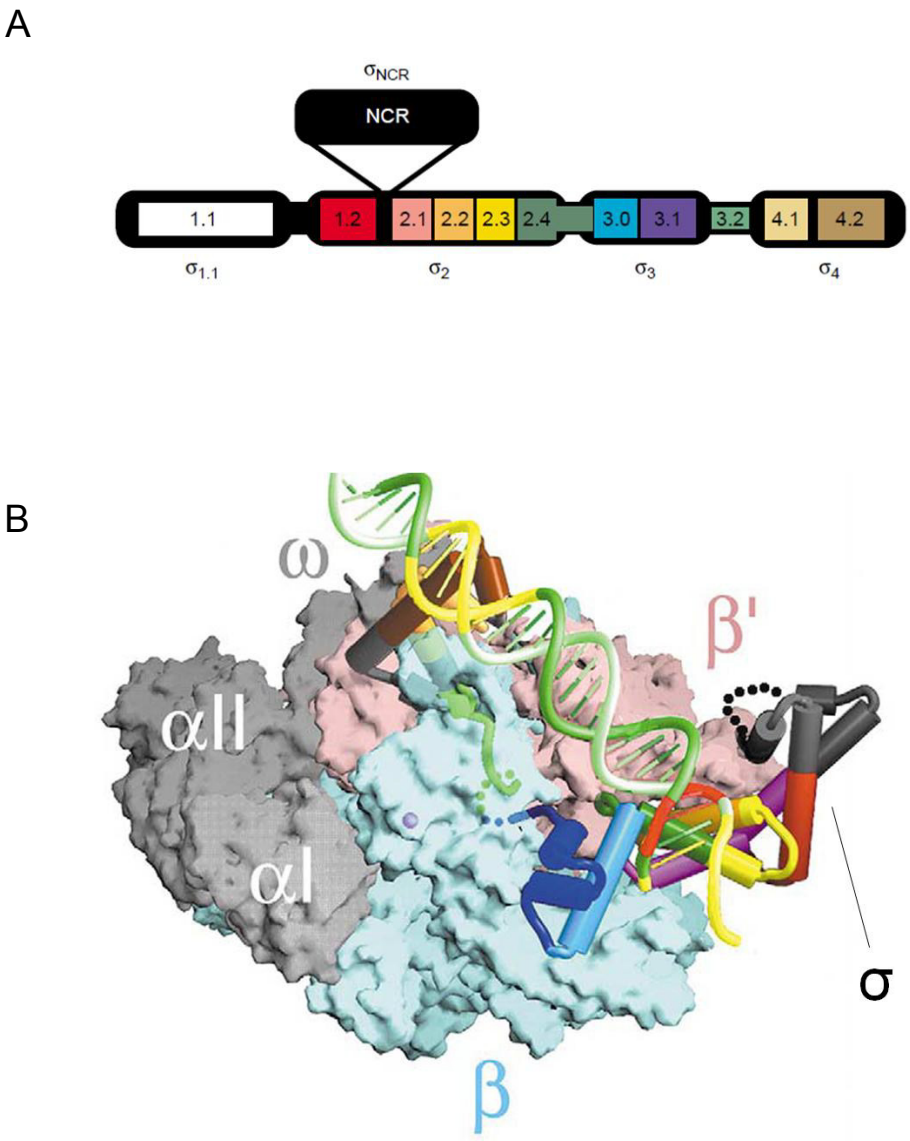


Figure 3. Structure of sigma factors and RNA polymerase. A) Domains of sigma factors. The conserved regions are indicated and color coded. Structured domains are indicated beneath. NCR stands for non-conserved region. B) Structure of the RNA polymerase holoenzyme and the DNA. The sigma factor is shown as a C backbone with α helices shown as cylinders, colored according to the conserved regions shown in A. Source: Nakamura and Darst, 2003.

catalytic activity, which includes the synthesis of RNA complementary to the DNA template in the presence of nucleoside triphosphates. The core enzyme by itself is incapable of recognizing specific promoter DNA sequences, or of melting the DNA to initiate transcription. To carry out these functions, it must bind one of several specificity factors, called, sigma (σ), to form a holoenzyme [69]. Bacterial sigma factors confer upon the core RNAP the ability to distinguish a promoter sequence within the DNA [108]. Association of different sigma factors with the core enzyme makes it possible for the holoenzyme to recognize different promoters and express different sets of target genes. Sigma factors thus provide efficient mechanisms for simultaneous regulation of large numbers of genes [18, 69]. Since sigma factors all compete for the same core RNAP, the repertoire of sigmas in several organisms demonstrates there are diverse ways of organizing transcriptional space [55].

The distribution of the RNAP core complexes to the sigma factors is the major level of regulation in a bacterium and understanding this process is a prerequisite to understanding bacterial life [108]. Since the sigma subunit associates with core RNA polymerase only during trancription initiation and is then released during elongation of the transcript, the use of specialized subunits is an effective method of transcriptional regulation [110]. Thus, multiple sigma factors in a single organism, each with different promoter selectivity, often play a key role in controlling gene expression [97]. The number of sigma factors encoded in a genome is quite variable and ranges from a minimum of one in *Mycoplasma* sp. [65] to a maximum of 63 in *Streptomyces coelicolor* [13, 65].

Sequence alignments of the sigma factors revealed that they have four regions of high conservation, σ1- σ4, and they can be further divided in subregions (Figure 3A). The first region does not have a much conserved sequence but acidity of the amino acid residues is preserved. One major function of the σ1.1 subregion is self-inhibition, preventing any unwanted promoter recognition for free floating σ factors. This inhibition is deactivated when the σ factor binds to the core region of RNAP [6]. Region 2, specifically subregions σ2.4 and σ2.3 are involved in promoter melting and -10 sequence recognition [6]. The third conserved region of the σ factor is comprised of three alpha helices. Region 3 also recognizes the extended -10 element and this interaction is believed to stabilize the holoenzyme open complex and promote initiation. The fourth conserved region of the σ factor is comprised of hydrophobic amino acid residues. This region binds to a β-1 flap of the RNAP as well as to

the -35 element of DNA. The two interactions alter the upstream DNA route, bringing it in closer proximity to RNAP [81, 104, 113].

The identification, characterization, and sequence analysis of bacterial sigma factors have revealed that they can be categorized into two structurally unrelated families , the σ70 and the σ54 families [113], with little if any sequence identity between them [81]. Although both sigma types bind the same core RNAP, their holoenzymes differ markedly in their control. Members of the σ70 family of sigma factors are components of the RNA polymerase holoenzyme that direct the bacterial core RNA polymerase to specific promoter elements that are usually 5-6 base-pairs (bp) in length and are centred 10 and 35 bp upstream (positions -10 and -35) of the transcription initiation site [59]. In contrast, σ54 proteins recognize the -12 and -24 sequences of promoters and require an activator protein [154]. The housekeeping sigma factor (RpoD or σ70 in *E. coli*) directs transcription of genes essential for basic cellular processes and required for cell survival under normal growth conditions [55]. However, bacteria also maintain a set of alternative sigma factors (from the σ70 family) that bind to consensus sequences that differ from the consensus sequence recognized by the housekeeping sigma factor. Transcriptional regulation in terms of regulons controlled by alternative sigma factors is a late evolving phenomenon and different bacterial species exhibit large differences in the number of alternative sigma factor encoding genes, presenting therefore huge flexibility in their transcriptional regulatory patterns. Alternative sigma factors direct transcription of genes required for specialized functions in response to changes in environmental stimuli, such as stress responses, as well as growth transitions and morphological changes [23].

3.2 THE σ54 FAMILY

Although no sequence conservation exists between the σ^{54} and σ^{70} families of sigma factors, both types bind to the core RNA polymerase. Nonetheless, the holoenzyme formed with σ54 sigma factors has different properties than a σ70 holoenzyme. The C-terminus of σ54 enables DNA binding, but σ54 factors require a separate activator protein along with the core RNAP to form an open promoter complex. The σ54 N-terminus inhibits isomerization in the absence of the appropriate activator [93]. Also, promoter structures recognized by σ54 differ from those recognized by σ70. σ54 promoters are short and highly conserved sequences that are located at -24 and -12 upstream of the transcription initiation site, in contrast to σ70 promoter sites, which are generally located at -35 and -10 upstream of the transcription start site.

3.3 THE σ70 FAMILY

The σ70 family includes the primary sigma factor present in all bacteria examined to date, as well as related alternative sigma factors [113]. The principal sigma factor in *E. coli* is sigma 70, so called because the protein is 70 kDa in size. *E. coli* also has six alternative sigma factors that are used in special circumstances (Table 1) [62]. Alternative sigma factors within the σ70 family are further categorized by the physiological processes they control. Within the σ70 family of sigma factors is a large, phylogenetically distinct subfamily called the extracytoplasmic function (ECF) factors. These sigma factors are responsible for regulating a wide range of functions, all involved in sensing and reacting to conditions in the membrane, periplasm, or extracellular environment [62]. Besides the primary sigma factor RpoD, the most studied sigma factors belonging to the sigma 70 family are RpoH, RpoS, RpoE, FecI and RpoF. RpoD-like sigma factors alone (without the RNAP) are not able to bind to promoters. This is prevented by sigma region 1, which inhibits contact between region 4 and the –35 site [22]. Functions of the σ70 family have differentiated in two directions: (1) the response to stress conditions endangering the biochemical transactions of life inside the cell (intracytoplasmic stress), and (2) the response to stressors that can be sensed and perhaps dealt with already at the cell wall, or that are injurious to this cellular compartment (extracytoplasmic stress) [108].

3.4 TYPES OF SIGMA FACTORS AND STATE OF THE ART IN *S. MELILOTI*

Fourteen putative sigma factor genes have been identified by genome sequence analysis of *S. meliloti* [45], and they are likely to be important regulators of the endosymbiotic process. The fourteen sigma factors are annotated as *rpoD*, *rpoN*, *rpoE1* to *rpoE9*, *fecI*, *rpoH1* and *rpoH2*. *S. meliloti* maintains multiple copies of genes that code for sigma sigma factors which are implicated in stress response: nine putative extracytoplasmic function (ECF) sigma factors (RpoE1-RpoE9), and two heat shock sigma factors (RpoH1 and RpoH2) [45]. The presence of multiple sigma factors might imply that response to environmental stress is highly regulated and crucial to the survival of this organism, in free-living conditions and even in its symbiotic interactions [110]. Some sigma factors have already been characterized in *S. meliloti*, namely

the genes for the RpoD and RpoN sigma factors [122, 124], as well as RpoE2 [10, 40] and two RpoH genes [110].

3.4.1 RpoD (σ^{70}) is the sigma factor for housekeeping functions

Sigma 70, the principal sigma factor, which is responsible for the housekeeping functions in growing cells, is called RpoD in proteobacteria and SigA in Gram-positive and other bacteria. Typically, most transcription in exponentially growing cells is initiated by RNAP holoenzyme carrying a housekeeping sigma factor similar to *E. coli* sigma 70 [108]. All other sigma factors (non-RpoN) are related to sigma factor RpoD.

The primary sigma factor in *S. meliloti* is very similar to other σ^{70} members. The *S. meliloti sigA*-homologous gene, coding for the primary sigma factor, was cloned and expressed *in vitro*. The deduced amino acid sequence of the *S. meliloti* SigA protein is very similar to those of *E. coli* σ^{70} subunit RpoD, as well as to other members of the σ^{70} family. The nucleotide sequence of the gene revealed 81% identity to the *E. coli rpoD* gene [124] and it is located in the *S. meliloti* chromosome rather than in one of the symbiotic megaplasmids. A putative transcription termination signal found in *sigA* also suggests that in *S. meliloti* this gene is not in an operon with downstream genes [124].

3.4.2 RpoE is the sigma factor for extracytoplasmic function

The most abundant class of alternative sigma factors is composed of structurally related proteins called extracytoplasmic function (ECF) sigma factors. ECF factors control the expression of genes encoding proteins that answer to physical, chemical or biological stress conditions [116]. In contrast to RpoH and RpoS, which react to intracytoplasmic stressors, ECF sigma factors react to disturbances already at the bacterial surface or within the bacterial cell wall, e.g., in the periplasmic space of Gram-negative bacteria [116]. RpoE and all other ECF sigma factors contain only the conserved domains 2 and 4 (Figure 3), which are essential for promoter binding. Most ECF sigma factors share the important property of specifically interacting with a protein called anti-sigma factor, which plays a role in the control of sigma factor activity. In the absence of a stimulus, the ECF sigma factor is kept inactive by

interaction with its cognate anti-sigma factor In the presence of a stimulus, the anti-sigma factor gets inactivated and the complex becomes available for transcription initiation [20, 62, 96, 113]. Bacteria with a profound ability to deal with a multitude of environmental stressors contain many different ECF sigma factors [96, 108]. σE in *E. coli*, can work as a second heat shock sigma factor [105]. ECF sigma factors can contribute to regulation of virulence and virulence-associated genes in a number of bacteria, including *Salmonella typhimurium*, *Pseudomonas aeruginosa*, and *Mycobacterium tuberculosis*. ECF sigma factors are also involved in pathogenesis [73].

The presence many different ECF sigma factors in *S. meliloti* suggests that this bacterium is quite resilient and might bear a profound ability to deal with diverse environmental conditions. Only one of the six ECF sigma factors has been studied so far. RpoE2 is activated by various stresses, including heat, salt and stationary phase, and controls the expression of other stress response genes such as *katC*, which codes for a catalase, and *rpoH2*, which codes for a heat shock sigma factor [10, 40, 41]. 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 [129]. Given the number of ECF sigma factors in *S. meliloti*, it is possible that there is functional overlap.

3.4.3 RpoH (σ^{32}) is involved in heat shock response

Sigma factor RpoH is required for the expression of heat shock genes, as well as other stress response genes. RpoH-like sigma factors contain all 4 conserved regions (Figure 3) [106], but occur only in proteobacteria. With the exception of members of the ε-group, all sequenced proteobacteria contain at least one copy of RpoH [108]. Within the RpoH sigma factor sequences there is a nine amino acid conserved segment that is called the RpoH box. This segment is involved in the control of the translation of a chaperone complex that binds denatured proteins generated under stress conditions. The RpoH box is also involved in the stability of the sigma factor and in its binding to the core RNA polymerase [108]. In databases, the RpoH sequences that represent confirmed RpoH-acting proteins usually contain a perfect match for this conserved sequence. Altered RpoH boxes are observed in organisms with multiple *rpoH* genes and may have functional significance for the action of these proteins.

The first RNA thermometer that acts via the melting mechanism was found in *the E. coli rpoH* gene [100]. RpoH levels in response to heat shock are controlled largely at the level of translation. Translational repression is mediated by RNA secondary structure that forms at the 5' end of the *rpoH* mRNA transcript [100]. Two segments in the open reading frame of *rpoH* form an RNA structure that blocks the entry of the ribosome to the Shine-Dalgarno (SD) sequence. Disruption of the structure at heat shock temperatures liberates the ribosome binding site and enhances translation of the sigma factor, resulting in the induction of the heat shock response. [99]. RpoH is important under non-stress conditions in *E. coli*, because the *rpoH* mutant is unable to grow above 20 °C [162]. The requirement for RpoH during growth is tied to the expression of genes that encode chaperones, in particular the *groESL* operon that encodes the subunits of the GroEL-GroES molecular chaperone machine. In *E. coli*, *rpoH* transcription is very sophisticated. Thereare five promoters upstream of the *rpoH* gene, recognizedby sigma 70, sigma E and sigma 54 [70].

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 *Agrobacterium tumefaciens* [105], and the *rpoH* genes in both *A. tumefaciens* and *Caulobacter crescentus* are autoregulated from RpoH-dependent promoters [120, 157]. Commonly, bacterial genomes contain a single *rpoH* gene, but all of the root-nodulating members of the α-proteobacteria whose genomes have been completely sequenced maintain multiple copies of *rpoH.* Multiple *rpoH* genes have been found in root-nodulating rhizobia: *Bradyrhizobium japonicum* [107], *Mesorhizobium loti* [71], *Rhizobium etli* [50], *Rhizobium leguminosarum* biovar *viciae* [156], *Sinorhizobium fredii* [86] and *S. meliloti* [45]. All of these species have two *rpoH* genes except *B. japonicum*, which has three [72, 107]. Two *rpoH* genes have been identified in *Brucella melitensis* and *Rhodobacter sphaeroides* [54]. Each of the *rpoH* genes of *B. japonicum*, *S. meliloti*, *Brucella melitensis* and *Rhodobacter sphaeroides* is able to complement, totally or partially, the temperature-sensitive phenotype of an *E. coli rpoH* mutant, thus suggesting that they are functionally similar to *E. coli* RpoH [29, 54, 107, 111]. *R. etli* also possesses two *rpoH* genes. The *rpoH1* gene is mainly involved in heat shock and oxidative responses, while *rpoH2* participates in osmotic tolerance. The *R. etli rpoH2* mutant is able to nodulate plant cells and to fix atmospheric nitrogen (Nod^+Fix^+ phenotype), while *rpoH1* and *rpoH2* mutants were able to nodulate *Phaseolus vulgaris* plants, causing a Nod^+ phenotype [87]. The presence of multiple copies of genes for a particular alternative sigma factor suggests that rhizobia may

contain multigene sigma families in order to respond more specifically to changes faced in either their symbiotic or free-living state [110].

S. meliloti has two genes that code for RpoH-like proteins: *rpoH1* and *rpoH2* [110], both of which are located on the chromosome [45]. In contrast to *E. coli*, the *rpoH1* and *rpoH2* single mutants and the *rpoH1 rpoH2* double mutant grow at normal temperatures [110, 111], but the *rpoH1* mutant is sensitive to high temperatures. The *rpoH1* mutant also exhibited increased sensitivity to various agents, suggesting that RpoH is required to protect the bacterial cell against other environmental stresses encountered within the host [97]. Synthesis of several heat shock proteins (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 [111]. Another study showed that the RpoH1 regulon controls expression of genes coding for proteases and chaperones such as *clpB*, *groESL5*, and *lon* during heat shock [97]. RpoH2 did not control the expression of any of these genes under the conditions tested. In symbiosis with alfalfa, an *rpoH1* mutant exhibits a nitrogen fixation defect (Fix^- phenotype), and an *rpoH1 rpoH2* double mutant exhibits a nodule formation defect (Nod^- phenotype). Electron micrographs showed that *rpoH1* mutant cells undergo senescence right after infecting plant cells and no apparent phenotype was found for the *rpoH2* single mutant either in culture or in symbiotic situations [97, 111].

3.4.4 RpoS (σ^{38}, σ^S) is required for the expression of stationary phase genes

The σ^S (RpoS) subunit of RNA polymerase is the master regulator of the general stress response in *E. coli* [64]. In this organism, RpoS is necessary for transcription of all genes whose products are required during stationary phase, as well as for transcription of a large number of stress-response genes [51, 95]. RpoS is regulated at the levels of transcription, translation, proteolysis and protein activity. It responds to starvation and many other stress conditions, such as high osmolarity, acidic or alkaline pH values, and low or high temperature [63]. While RpoS is smaller than RpoD, it contains all 4 conserved regions (Figure 3), which are necessary for this sigma factor proper function. Like RpoD, wild-type RpoS is not able to bind to promoters in the absence of RNA polymerase [108]. In *E. coli*, RpoS is also a major regulator of the general starvation response. σ^S is called sigma B in Gram-positive bacteria and also plays a key role in protecting from different environmental stress conditions, including starvation, hyperosmolarity, oxidative damage, and reduced pH [95]. Through

enhancing environmental survival, as well as by directly activating virulence genes, σ^B and σ^S have important roles in bacterial pathogenesis.

No *rpoS* homologue could be found in any of the completely sequenced genomes of α-proteobacteria. Even though an RpoS factor has not yet been characterized in *S. meliloti*, nor any sequence in the annotated genome seems to share homology to the RpoS sigma factors of other bacteria, an *in vivo* study has shown that in *S. meliloti*, sigma factors which are functionally homologous to RpoS are able to recognize the promoter sequences of *E. coli* stationary-phase genes [95]. Plasmids containing RpoS-dependent growth phase-regulated promoters of *E. coli* were mobilized into *S. meliloti* and transcriptional activation of the promoters was growth phase dependent, as it is in *E. coli*. The results suggest that a putative stationary-phase-specific RNA polymerase with functional homology to the RpoS of *E. coli* exists in *S. meliloti* [95].

3.4.5 RpoN (σ^{54}) is a unique sigma factor involved in nitrogen metabolism

RpoN is a unique sigma factor. It differs considerably in sequence and structure from the primary sigmas and is responsible for the transcription of specific regulons [81]. These sigma factors are not related to the RpoD-like proteins; rather, they are used by the bacterial cell to initiate transcription from promoters that do not bind to RpoD [108]. RpoN recognizes a GG at position –24 and a GC at position –12 promoter sequences [9]. These differences may mirror the fact that RpoN-containing RNA polymerase holoenzyme alone cannot generate the open complex state of transcription initiation and is therefore unable to initiate transcription by itself. Instead, the open complex state is accomplished by the ATP hydrolysis activity of an activator protein, called Enhancer-Binding Protein (EBP). EBPs generally bind to enhancer sites situated 100 base pairs or more upstream of the transcription initiation site and contact the σ^{54}-holoenzyme through DNA looping. Because each EBP is controlled by its own signal transduction pathway, different sets of RpoN-dependent genes can be transcribed under diverse conditions [98]. Whereas sigma 70 holoenzymes carry out this process on their own, sigma 54 holoenyzmes require both an enhancer and ATP to perform this process [137].

RpoN occurs only in proteobacteria and a few Gram-positive bacteria. RpoN controls operons that have to remain absolutely silent when not needed. Most of them are connected to nitrogen metabolism and encode proteins that mediate a high turnover of energy. The processes controlled by RpoN are not essential for cell survival and growth under favorable conditions, with the exception of *Myxococcus xanthus* [74]. Although most eubacteria contain multiple members of the sigma 70 family, they usually have no more than one representative of the sigma 54 family. Moreover, the distribution of sigma 54 family members has no known representatives in any high-GC, Gram-positive bacteria or in cyanobacteria [137].

In the symbiotic group of the Rhizobiales, RpoN is best known as the sigma factor enabling transcription of the nitrogen fixation genes. However, rhizobia have also recruited RpoN for the expression of other symbiotic genes [33]. In *S. meliloti*, RpoN-binding sites were found upstream of genes involved in common RpoN-dependent functions, such as assimilation of ammonium and uptake of C4-dicarboxylic acids. In this organism, the σ^{54}-dependent positively acting regulatory enhancer-binding proteins NifA, NtrC, DctA and DctD have been characterized. NifA is used for activating the expression of nitrogen fixation genes during symbiosis; NtrC is involved in ammonium transport and assimilation [122]; DctA and DctD are required for activation of promoters involved in C4-dicarboxylate metabolism and transport [79]. Furthermore, several other symbiosis-related genes are reported to be regulated by RpoN [33]. Sequence analysis predicts that *S. meliloti* encodes yet four additional EBPs [8].

3.4.6 FecI is the sigma factor for iron metabolism regulation

FecI is another type of extracytoplasmic function sigma factor. In *E. coli,* FecI takes part in a signaling cascade that starts directly at the cellular surface and proceeds to the nucleoid in the cytoplasm [19]. This cascade induces the genes that mediate ferric citrate uptake. FecI-dependent RNAP holoenzyme initiates transcription of the *fec* operon, additionally under control of an iron regulator called Fur [19]. Proteins that are closely related to *E. coli* FecI occur only in proteobacteria and, as an exception, they are also present in *Streptomyces coelicolor*. Due to their high similarity, these close relatives of *E. coli* FecI may also play a role in cellular iron homeostasis [108].

In *S. meliloti*, *fecI* is most likely in an operon with *fecR*, which is a putative iron transport regulator transmembrane protein. The FecI predicted protein contains conserved sigma factor regions 3 and 4 [45].

3.4.7 RpoF is the sigma factor for motility functions

FliA-like proteins are the sigma factors responsible for motility and differentiation functions. In *E. coli*, FliA is the sigma factor required for flagellum synthesis and chemotaxis. This group also includes Gram-positive factors SigB, D, E, F, G, I and K. FliA-like sigma factors are related to RpoD and RpoE, which contain the conserved regions 2, 3, and 4, but not region 1 (Figure 3) [108]. These proteins might be able to bind alone to promoters and wait there for RNAP complexes to be recruited, if this is not prevented by anti-sigma factors. They occur frequently in bacteria, even in non-motile ones [68]. In *S. meliloti*, no sigma factor has yet been identified as being the sigma factor for motility functions.

4. OBJECTIVE

It is a scientific challenge to learn how sigma factors interact with pathways, activators, and repressors to form the highly complicated regulatory network that allows a cell to survive in specific environments or to perform intricate tasks of cellular differentiation. In a complex interaction between two organisms, such as the *Rhizobium*-legume symbiosis, the ability to assess global transcription patterns and transcriptional regulation by sigma factors is of utmost interest. Besides, stress response is paramount for the survival of rhizobia, both in the soil and in planta. For this reason, the demand for a more effective utilization of biologically fixed nitrogen in agricultural systems has prompted studies on rhizobia diversity and tolerance to biotic and abiotic factors, such as pH and temperature.

The purpose of the present study was to gain detailed insight into sigma factor regulation and the complex stress response regulatory system of *S. meliloti* using different kinds of stress conditions as effectors. The aim was likewise to provide a basis for understanding the molecular mechanisms of sigma factor regulation and to identify genes involved in stress response whose expression is sigma factor-dependent. Because of the fact that pH constitutes

a major stressor for *S. meliloti*, both in soil and during symbiosis, the main focus of this work was on the characterization of sigma factor regulation under pH stress, which, unlike heat shock and other stresses, had not been previously described in linkage to sigma factor regulation. Also, because the regulation of gene expression is a dynamic process, special attention was granted to the characterization of changes in gene expression over time, with the aid of global transcription profiling analyses.

5. PRESENTATIONS

5.1 PUBLICATIONS

Daniella de Lucena, Alfred Pühler, and Stefan Weidner. The role of sigma factor RpoH1 in the pH stress response of *Sinorhizobium meliloti*. BMC Microbiology. 2010 18;10:265.

5.2 CONFERENCE POSTERS

July 2006. 7th European Nitrogen Fixation Conference, University of Aarhus - Denmark

October 2007. Functional Genomics and Systems Biology, Wellcome Trust Genome Campus, Hinxton - UK

September 2009. 7th European Nitrogen Fixation Conference, University of Gent - Belgium

B. MATERIAL AND METHODS

1. MATERIAL

1.1 BACTERIAL STRAINS AND PLASMIDS

1.1.1 Bacterial Strains

Escherichia coli

Strain	Characteristics	Reference
DH5α MCR	F$^-$ *endA1 supE44 thi-1* λ$^-$ *recA1 gyrA96 relA1 deoR* Δ(*lacZYA-argF*)*U169* φ80d*lacZ*ΔM15 *mcrA* Δ(*mrr hsdRMS mcrBC*)	[53]
S17-1	*E. coli* 294 ::RP4-2(Tc::Mu)(Km::Tn7) *pro res _recA*, Tpr	[134]

Sinorhizobium meliloti

Rm1021	Spontaneous mutant of wild type strain RU47, Sm^r*; NxR;* wild type	[90]
Rm1021Δ*rpoE1*	Rm1021 derivative, *rpoE1* mutant, Sm^r	This study
Rm1021Δ*rpoE2*	Rm1021 derivative, *rpoE2* mutant, Sm^r	This study
Rm1021Δ*rpoE5*	Rm1021 derivative, *rpoE5* mutant, Sm^r	This study
Rm1021Δ*rpoH1*	Rm1021 derivative, *rpoH1* mutant, Sm^r	This study
Rm1021Δ*fecI*	Rm1021 derivative, *fecI* mutant, Sm^r	This study
Rm1021Δ*rpoE3*	Rm1021 derivative, *rpoE1*	This study

	mutant, Sm^r	
Rm1021Δ*rpoE7*	Rm1021 derivative, *rpoE7* mutant, Sm^r	This study
Rm1021Δ*rpoH1*Δ*rhrA*	Rm1021 derivative, *rpoH1 rhrA* mutant, Sm^r	This study
Rm1021Δ*rpoH1*Δ*rpoE2*	Rm1021 derivative, *rpoH1 rpoE2* mutant, Sm^r	This study
Rm1021Δ*rpoH1*Δ*rpoE3*	Rm1021 derivative, *rpoH1 rpoE3* mutant, Sm^r	This study
Rm1021Δ*rpoH1*Δ*rpoE5*	Rm1021 derivative, *rpoH1 rpoE5* mutant, Sm^r	This study

1.1.2 Plasmids

Plasmid	**Characteristics**	**Reference**
pK18mobsacB	pUC18 derivative, *sacB lacZ*α Km^r, mobilizable	[130]
pK18Δ*rpoE1*	pK18mobsacB derivative, *sacB lacZ*α Km^r, mobilizable	This study
pK18Δ*rpoE2*	pK18mobsacB derivative, *sacB lacZ*α Km^r, mobilizable	This study
pK18Δ*rpoE5*	pK18mobsacB derivative, *sacB lacZ*α Km^r, mobilizable	This study
pK18Δ*rpoE7*	pK18mobsacB derivative, *sacB lacZ*α Km^r, mobilizable	This study
pK18Δ*rpoE8*	pK18mobsacB derivative, *sacB lacZ*α Km^r, mobilizable	This study
pK18Δ*rpoH1*	pK18mobsacB derivative, *sacB lacZ*α Km^r, mobilizable	This study
pK18Δ*fecI*	pK18mobsacB derivative, *sacB lacZ*α Km^r, mobilizable	This study
pJN105	pJN105 *araC-PBAD* cassette cloned in pBBR1MCS5*;*	This study

	GmR;	
pJN105*rpoH1*	pJN105derivative, *araC-PBAD* cassette cloned in pBBR1MCS5*;GmR*;	This study

1.2 OLIGONUCLEOTIDES

Name	**Sequence (5'→3')**
check_fecI1A	GATCGTGCGCCACATCGAAG
check_fecI1B	GCCGTGCCGAGAACTTCGATT
check_rpoE1A	TGCGATCCTGTTGCGGATA
check_rpoE1B	TGGTGACGGAACTCTCCAT
check_rpoE2A	TCGGCTTCGTCGTCTCGTA
check_rpoE2B	CCTCGTCCTTCAGCTTGAACA
check_rpoE5A	GGTTAACGCACGAAGGTAGAAG
check_rpoE5B	CGACAACGAATAGAGCGAAGGA
check_rpoH1A	GAAGAACGATGCCGCACACC
check_rpoH1B	TAGAGCCGCCGAACGCCAAT
check_fecI2A	AGTGGCTTGGCAACGCAACC
check_fecI2B	CATCATCAAGACCGGCATCG
check_rpoE3A	TCCGGATCACCTGGCGATTA
check_rpoE3B	GATGACGGTCCGATCATCAC
compl_rpoH1_A	GCGAGAATTCGCCGCACTTATCTATTAGCC
compl_rpoH1_B	GATCGGATCCTTAAGCGCCTTCAACCAC
del _fecI1_B	GGTGCCGCAGGTACATGTGA
del_fecI1_A	CGCGCATTGGTCGTGCGATT
del_fecI1_C	TCACATGTACCTGCGGCACCAGGCCTCGACCATGACGAAT
del_fecI1_D	GATCGTGCGCCACATCGAAG
del_rpoE7_A	AGCTGTCGACTGTAAGCGGCGATGCCTTCC

del_rpoE7_B	CAAGAACCGCCTCGGTCAGACGGTTGTGCGCCTGCATGAT
del_rpoE7_C	ATCATGCAGGCGCACAACCGTCTGACCGAGGCGGTTCTTG
del_rpoE7_D	AGCTGTCGACCGATCGCCACCTGGATGGTT
del_rpoE1_A	AGTAGGATCCGCGATCAGGAGGTCAT
del_rpoE1_B	GTCCTTCATCGCTTCGGCAACCGGCATCAATTCCAG
del_rpoE1_C	CTGGAATTGATGCCGGTTGCCGAAGCGATGAAGGAC
del_rpoE1_D	AGTCGGATCCACGATCCTCTGCGTTGAAGC
del_rpoE2_A	ATCGGAATTCGCTCGTCCTCGATGAT
del_rpoE2_B	AACGAAGGCACGCGAGGTGACACGCTTGAACTCTTGG
del_rpoE2_C	CCAAGAGTTCAAGCGTGTCACCTCGCGTGCCTTCGTT
del_rpoE2_D	AGCGGAATTCAACCGCGACGGTTCCTATC
del_rpoE3_A	ATTAGGATCCGAGGCGGTTACCGAATGGCT
del_rpoE3_B	CAACATGGCTGCGAGCCTGAGGCGCATTGCTCTTCCAGTG
del_rpoE3_C	CCGCGTAACGAGAAGGTCACTCAGGCTCGCAGCCATGTTG
del_rpoE3_D	GCATAAGCTTCGCGGATGGTCTCGCTGAAT
del_rpoE5_A	GCGCAAGCTTCTGCAGGATGGAAGCGATT
del_rpoE5_B	CTCGTCCGCTCAGTTCAATTGTCGCGATGCGTGACC
del_rpoE5_C	GGTCACGCATCGCGACAATTGAACTGAGCGGACGAG
del_rpoE5_D	ACGTAAGCTTGCCGACCAGAACCGTAA
del_rpoH1_A	CGAAGACAGCGACGATGCAC
del_rpoH1_B	ACCAGCCAATCCTGCCACTGCTCGAACTTCTTGACCGCCT
del_rpoH1_C	AGGCGGTCAAGAAGTTCGAGCAGTGGCAGGATTGGCTGGT
del_rpoH1_D	TATGAAGAGAGGCTCGGCCA

1.3 ENZYMES AND MARKERS

Enzyme	**Supplier**
Restrictionendonucleases	Fermentas
T4-DNA Ligase	Roche
Taq-DNA-Polymerase	Peq-Lab
Pfu DNA Polymerase	Fermentas
RNaseA	Qiagen
DNaseA	Qiagen

Superscript II Reverse Transcriptase	Invitrogen
DNA molecular weight marker	Roche
DNase I	Qiagen

1.4 CHEMICALS

Chemical	**Supplier**
2,2-dipyridyl	Roth
Aceton	Roth
Agar	GIBCO
antibiotics	Serva, Sigma-Aldrich
Bromephenol blue	Serva
dNTPs	Qiagen Amersham Biosciences
Ethanol	Roth
ethidium bromide	Serva
HCl	Roth
IPTG(isopropyl-beta-D- thiogalactopyranoside)	Serva
Isopropanol	Roth
LB Base/Agar	GIBCO
Tris ICN	Biomedicals
Tryptone	Oxoid
X-Gal	Roth
Yeast Extract	Oxoid Difco
B-Mercaptoethanol	Roth

1.5 KITS

Kit	**Supplier**
QIAprep Miniprep	Qiagen
QIAquick PCR Purification Kit	Qiagen

Rneasy Minikit	Qiagen
Rnase-Free DNAse Set	Qiagen
CyScribe Purification Kit	Amersham Biosciences
DNA Labeling and Detection Kit	Boehringer

1.6 ADDITIONAL MATERIAL

Material	**Supplier**
Eppendorf tubes	Greiner/Star Lab/Brand
Glassware	Schott
Cuvettes	Brand
Gloves	Ansell
Parafilm	American
PCR-Stripes	Biozym
Petri dishes	Greiner
Pipette tips	Greiner/Star Lab
Sterile filters	Schleicher & Schuell
Plastic tubes	Greiner
Cellulose acetate filters	Sartorius

1.7 SOFTWARE AND INTERNET SOURCES

Program	**Reference / Company**
BioEdit 7.0	[58]
Clone Manager 6.0	SciEdCentral Software
BLAST	[3]
EMMA	[34]
Genesis	[138]
Excel	Microsoft
ClustalW	[142]
Excel Sheet for RT RT-PCR analysis	[102]

Protein BLAST	[85]
GenDB	[94]

1.8 CULTURE MEDIA AND SUPPLEMENTS

Media

LB-medium (Luria-Bertrani Broth) [127]

10 g/l	Tryptone
5 g/l	Yeast Extract
5 g/l	NaCl

TY-medium [14]

5 g/l	Tryptone
3 g/l	Yeast extract
0.4 g/l	$CaCl_2$

Vincent minimal medium (VMM) [11, 148]

Solution A:	2.56 g/l	K_2HPO_4
	1.56 g/l	KH_2PO_4
	0.246 g/l	$MgSO_4 \times 7\ H_2O$
	1 g/l	NH_4Cl
	1.62 g/l	Na_2-succinate

Ingredients were dissolved in 800 ml Millipore and pH value was adjusted KOH or H_3PO_4. Water was added to fill 1 liter and the solution was autoclaved.

Solution B:	67 g/l	$CaCl_2$

Autoclaved

Solution C: 10 g/l $FeCl_3 \times 6\ H_2O$
Sterile filtered

Solution D: 3 g/l H_3BO_3
2.23 g/l $MnSO_4 \times 4\ H_2O$
0.287 g/l $ZnSO_4 \times 7\ H_2O$
0.125 g/l $CuSO_4 \times 5\ H_2O$
0.065 g/l $CoCl_2 \times 6\ H_2O$
0.12 g/l $NaMoO_4 \times 2\ H_2O$
Sterile filtered

1 ml of each solution 2, 3, 4 and 5 were added one by one to 1 liter of solution 1.

SOB Medium [127]

Solution A: 20 g Tryptone
5 g Yeast extract
0.5 g NaCl
Dissolved in 980 ml H_2O

Solution B: 18.65 g/l KCl

Solution C: 406.6 g/l $MgCl_2 \times 6\ H_2O$

Solution D: 493 g/l $MgSO_4 \times 7\ H_2O$

This medium was used for preparation of the competent *E. coli* cells. 10 ml of solution B were added to solution A, and then they were autoclaved. Solutions C and D were autoclaved separately, and 5 ml of each were added to the A and B solutions.

SOC medium
SOB medium supplemented with 20 ml of an 180 g/l solution of glucose.

Hogness Freezing Medium [153]

Stock 1:	520 ml	Glycerol (87%)
	4.99 g	Na_2-citrate × 2 H_2O
	9 g	$(NH_4)_2SO_4$
	0.99 g	$MgSO_4$ × 7 H_2O
	Filled-up with water till 800ml and autoclaved	
Stock 2:	6.2 g	K_2HPO_4
	1.796 g	KH_2PO_4
	Filled-up with water till 200ml and autoclaved	

After autoclaving solutions were cooled down and mixed together.
All the media, if not indicated otherwise, were dissolved in distilled water and autoclaved.

Supplements

Supplements for growth media

Agar
For solid bacterial media, 16 g/l were added.

Sucrose
To select sucrose-resistant *S. meliloti* 1021 colonies (for loss of *sacB* gene) 10 g/l of sucrose were added directly to TY medium before autoclaving. To test *E. coli* sucrose-sensitivity 5 g/l were added to LB medium before autoclaving.

Antibiotics

Kanamycin (Km)
For the selection of kanamycin-resistant *E. coli* clones 50 µg/ml of antibiotic were added to solid media; 30 µg/ml were added to liquid media to prevent the loss of plasmids. For the

selection of kanamycin-resistant *S. meliloti* clones 200 µg/ml of antibiotic were added to both solid and liquid media.

Streptomycin (Sm)
For the selection of *S. meliloti* strain 1021, 600 µg/ml of antibiotic were added to both solid and liquid media.

Gentamycin (Gm)
For the selection of Gentamycin-resistant *E. coli* clones, 8-10 mg/l of antibiotic were added to solid media.

Neomycin (Nm)
For the selection of *S. meliloti* clones who had lost the pK18mobsac plasmid, 80 mg/l of Neomycin antibiotic were added to solid media.

The antibiotics were sterile filtered and 1 ml of each stock solution was added to 1 l of medium, when indicated.

Concentrations of growth media antibiotics used for *E. coli* and *S. meliloti*.

Antibiotic	***E. coli***	***S. meliloti***
Gentamicin	10 mg/l	30 mg/l
Kanamycin	50 mg/l	-
Neomycin	-	80 mg/l
Streptomycin	-	600 mg/l

1.9 BUFFERS AND SOLUTIONS

DNA and enzyme buffers

E1-Lysis solution 25 % (w/v) Sucrose

	10 % (w/v) Ficoll	
	After autoclaving:	
	0,4 mg/ml RNAse A	
	1 mg/ml Lysozyme	
Na-Acetate	3 M Na-Acetate pH 5.3	
EDTA	25 mM EDTA pH 8	
RNase-Lysozyme-Solution	1 mg/ml RNAse A	
	2 mg/ml Lysozyme	
	Dissolved in TE buffer	
Sucrose solution	20 % (w/v) Saccharose in TE buffer	
Proteinase K solution	20 mg/ml Proteinase K in TE buffer	
TE-buffer	10 mM	Tris-HCl
	1 mM	EDTA
	pH 7.5	
10 × TA- restriction buffer	660 mM	K-Acetate
	330 mM	Tris-HCl
	100 mM	Mg-Acetate
	5 mM	Dithiothreitol (DTT)
	1 mg/ml	Bovine serum albumin (BSA)
	pH 7.5	(adjust with acetic acid)
10 mM dNTP-mix	10 mM	dATP
	10 mM	dCTP
	10 mM	dGTP
	10 mM	dTTP
RNase A	20 mg/ml	

DNA electrophoresis buffers

TAE buffer	40 mM	Tris-HCl
	10 mM	Na-Acetate
	1 mM	EDTA
	pH 7.8	(adjust with acetic acid)
BPB loading buffer	80 ml	Glycerol
	10 ml	TAE buffer
	2.5 g	Bromophenol blue
Agarose gel	0,5-2 % (w/v)	Agarose in TAE buffer

Solutions for microarray hybridizations

20 × SSC	3 M	NaCl
	0.3 M	Na_3-Citrate
	pH 7.4	
Rinsing solution 1	250 ml	H_2O
	250 µl	Triton X100
	Dissolved at 80 °C for 5 min; cooled down to the room temperature	
Rinsing solution 2	500 ml	H_2O
	50 µl	32 % HCl
Rinsing solution 3	225 ml	H_2O
	25 ml	1 M KCl
Blocking solution	150 ml	H_2O

	40 µl	32 % HCl
	50 ml	4 × QMT Blocking solution
	MilliQ/HCL mix is pre-warmed to 50 °C. The 4 × QMT Blocking solution was added 5 min before use and the complete Blocking solution was pre-warmed to 50 °C for at least 5 min.	
Washing buffer 1	2 × SSC	
	0.2 % SDS	
	30 °C	
Washing buffer 2	0.5 × SSC	
	20 °C	

Buffer for the preparation of competent *E. coli* cells

$CaCl_2$ buffer	10 mM	HEPES
	15 mM	$CaCl_2 \times 2\ H_2O$
	55 mM	$MnCl_2 \times 2\ H_2O$
	250 mM	KCl
	All components, except $MnCl_2$, were mixed and the pH was adjusted to 6.7 with KOH. Then the $MnCl_2$ was added and the mixture was filter sterilized.	

2. METHODS

2.1 CULTIVATION OF BACTERIA

Bacteria were grown using solid media, as well as liquid media. In special cases, the media were supplemented with additives listed above (section 1.8). The growth of bacteria in solid medium was achieved by spreading of a single colony and incubation at the adequate temperature. The cultivation of bacteria in liquid media was made by inoculating a single colony from a plate followed by incubation with shaking at 180 rpm. *E. coli* cells were cultivated at 37 °C in LB. *S. meliloti* cells were cultivated at 30 °C in TY or Vincent minimal medium (VMM). Liquid cultures were shaken at 180 rpm.

2.2 DETERMINATION OF VIABLE TITLE

Viable titles were determined before preparing serial 10-fold dilutions, from 10^{-1} to 10^{-5}. 100 µl of the appropriate dilutions were then plated on non selective plates using sterile handles. Plates were then incubated at the appropriate temperature until colonies were grown. Bacterial titers were also measured photometrically at 600 nm against a blank containing culture medium. An optical density (OD_{600}) of 0.1 corresponds to 2 x 10^7 *E. coli* cells/ml or 1 x 10^8 *S. meliloti* cells/ml.

2.3 PROCEDURES FOR CONTINUOUS PH AND PH SHIFT GROWTH EXPERIMENTS

S. meliloti strains were grown in VMM at 30 °C at either pH 7 or pH 5.75 for growth tests at continuous pH values. Triplicate samples were measured for optical density at 580 nanometers, twice a day, for 7 days. For pH shift experiments cells of three independent cultures were grown in 30 ml of VMM with pH 7.0 to an $o.D._{580}$ of 0.8. Cell cultures of each flask were then centrifuged (10,000×g, 2 min, 30 °C) and the supernatant was discarded. The cell pellets were resuspended in 30 ml VMM with pH 5.75 or 30 ml VMM with pH 7.0 (control) and incubated at 30 °C. The pH of the medium was set by addition of HCl or NaOH. At six time points cell suspension probes of 5 ml were harvested from each flask and immediately centrifuged (10000×g, 1 min, 4 °C). The resulting pellets were instantly frozen in liquid nitrogen for later RNA preparation. Cell suspension probes were harvested at 0, 5, 10, 15, 30, and 60 minutes following the pH shift.

2.4 STORAGE OF BACTERIAL STRAINS

Bacteria were grown in the liquid medium overnight until the optical density (OD_{600}) of 1 was achieved. 100 µl of bacterial culture were mixed with 120 µl of 87 % (v/v) glycerol and stored at -20 °C. For storage at -80 °C, 180 µl of bacterial culture were mixed with 20 µl of Hogness Freezing Medium.

2.5 ISOLATION, MANIPULATION AND ANALYSIS OF NUCLEIC ACIDS

2.5.1 Isolation of genomic DNA from *S. meliloti*

The whole genomic DNA was isolated from *S. meliloti* cells using NucleoSpin Tissue kit from Macherey-Nagel. With the NucleoSpin Tissue method, lysis is achieved by incubation of the samples in a solution containing SDS and proteinase K at 56 °C. DNA binds to the silica membrane of the NucleoSpin Tissue columns and binding is enhanced by addition of chaotropic ions and ethanol to the lysate. Contaminations are removed by washing buffer and genomic DNA is finally eluted in elution buffer under low ionic strength conditions.

Protocol:

- Centrifuge up to 1 ml culture with $o.D._{600}$ = 0.8 for 5 min at 8,000 × g. Remove supernatant.
- Resuspend the pellet in 180 µl buffer T1 by pipeting up and down. Add 25 µl of proteinase K. Vortex vigorously and incubate with shaking at 56 °C until complete lysis is obtained (at least 1h).
- Add 20 µl RNase A solution and incubate for an additional 5 min at room temperature.
- Vortex the samples. Add 200 µl of buffer B3, vortex vigorously and incubate at 70 °C for 10 min. Vortex briefly.
- Add 210 µl of ethanol (96 – 100 %) to the sample and vortex vigorously.
- For each sample, place one NucleoSpin® Tissue column into a 2 ml collecting tube. Apply the sample to the column. Centrifuge for 1 min at 11,000 × g. Discard the flow-through and place the column back into the collecting tube.

- Add 500 µl buffer BW. Centrifuge for 1 min at 11,000 × g. Discard flow-through and place the column back into the collecting tube.
- Add 600 µl buffer B5 to the column and centrifuge for 1 min at 11,000 × g. Discard flow-through and place the column back into the collecting tube.
- Centrifuge the column for 1 min at 11,000 × g.
- Place the NucleoSpin Tissue column into a 1.5 ml microcentrifuge tube and add 100 µl prewarmed elution buffer BE (70 °C). Incubate at room temperature for 1 min. Centrifuge 1 min at 11,000 × g.

2.5.2 Isolation of Plasmid DNA

Plasmid DNA was isolated from bacteria grown overnight in liquid medium. Bacterial culture was added to a 1.5 ml Eppendorf tube and cells were sedimented by centrifugation at 6000 rpm for 10 minutes. Supernatant was then discarded. Then plasmid extraction was carried out using the QIAprep Spin Miniprep Kit (250) (QIAGEN) and the buffers (P1, P2, N3 and PE) delivered with the kit following the manufacturers protocol (centrifugation steps were carried out at room temperature for 1 minute at 13,000 rpm). This method combines alkaline lysis with silica-gel membrane columns that bind DNA under high salt conditions and elute at low salt conditions.

Protocol:

- The sedimented cells were resuspended in 250 µl of Buffer P1.
- 250 µl of Buffer P2 were added and the Eppendorf tubes were inverted 3-5 times. Alkaline lysis was carried out until the cell suspension cleared. 350 µl of Buffer N3 were then added to neutralize the lysate pH and to introduce high salt concentrations. The tubes were inverted 3-5 times.
- Samples were centrifuged for 10 minutes. The supernatant was applied to the provided QIAprep Spin Column by decanting.
- To bind the plasmid DNA to the QIAprep Column membrane the columns were centrifuged and the flow-through was discarded.
- The plasmid DNA was washed by adding 750 µl of Buffer PE to the columns and then centrifuging them. The flow-through was discarded and the columns were centrifuged for another minute to remove residual buffer PE.

- To elute the DNA, 40 µl of autoclaved MilliQ H_2O were added to the center of the column. The columns were transferred to new, sterile and labeled Eppendorf tubes and centrifugation for achievement of eluted DNA was performed for one minute.

2.5.3 RNA extraction and purification:

2.5.3.1 RNA extraction

Cells were harvested by pipetting 1.5 ml of culture in four eppendorf tubes each. The suspensions were centrifuged for 1 min at 13,000 rpm, supernatant was discarded and the cells centrifuged for an additional 30 sec. Then the supernatant was completely removed with a pipette and the cells were immediately frozen in liquid nitrogen until use. All following steps were carried out on ice. To extract the RNA, components of the RNeasy Mini Kit (50) (QIAGEN) were used (RLT Buffer, RWI Buffer,RPE buffer, RNeasy mini columns). Cell pellets were resuspended in 200-400 µl of cold 10 mM Tris-HCl pH 8.8. Cells (maximum volume 200 µl) were then transferred into Fast Protein Tubes containing 700 µl RLT Buffer and glass powder. The cells were cracked open in the Ribolyzer for 30 sec at level 6.5 and immediately incubated for 3 min on ice. Insoluble compounds were removed by centrifugation (13,000 rpm, 4 °C, 3 min) and 675 µl of each supernatant were transferred into a new Eppendorf tube.

The following steps were carried out at room temperature. 375 µl of ethanol 100 % were added to each RNA extract, mixed and 700 µl were transferred immediately into an RNeasy mini column. The RNA sample was centrifuged at 8000 x g for 15 sec. The flow-through was discarded. The columns were washed by adding 700 µl of RWI Buffer to each column and centrifugation at 8000 x g for 15 sec. The columns were placed in new collection tubes and 500 µl of RPE buffer was added to each column. The columns were centrifuged at 8000 g for 15 sec and the flow through was discarded. This step was repeated with a centrifugation at 8000 x g for 2 min to dry the membrane and ensure successful elution without buffer contaminations. The columns were transferred into new, RNase free Eppendorf tubes and the RNA was eluted with 40 µl of RNase-free water, which was applied to the center of the membrane. The columns were centrifuged for 2 min at 8000 x g for elution.

2.5.3.2 DNase treatment

To remove DNA traces in the RNA extracts, 20 µl RDD buffer (QIAGEN) and 10 µl of DNase (QIAGEN) were added to 80 µl of each RNA extract and incubated for 1 hour at 30 °C.

2.5.3.3 RNA purification

To remove the DNase in the RNA extracts, another RNA purification with the QIAGEN RNeasy kit was performed as follows: 350 µl of Buffer RLT were added to the 110 µl of RNA extract from the DNase treatment. After the addition of 250 µl 100% ethanol, each sample was mixed and loaded immediately onto a RNeasy mini column and centrifuged for 15 sec at more than 8000 x g. The columns were placed in new collection tubes and 500 µl of RPE buffer was added to each column. The columns were centrifuged at 8000 x g for 15 sec and the flow through was discarded. This step was repeated with a centrifugation at more than 8000 x g for 2 min to dry the membrane and ensure successful elution without buffer contaminations. The columns were transferred into new, RNase free Eppendorf tubes and the RNA was eluted with 35 µl of RNase free H_2O, which was applied to the center of the membrane. Then the columns were centrifuged for 2 min at more than 8000 x g.

2.5.3.4 DNA-check-PCR and RNA integrity

To make sure that the DNase I treatment was successful a 20 µl-scale PCR was carried out using 1 µl of each RNA extract as templates (any primer pair that performs well in PCR can be used here). 10 µl of each PCR product and 1 µl of each RNA extract were loaded on a 1.5 % (w/v) agarose gel.

2.5.4 Visualization and quantification of DNA

2.5.4.1 Agarose gel electrophoresis

DNA is negatively charged at pH 7.0 and therefore migrates to the positive pole of an electrical field. Agarose gels are used to separate DNA molecules by their length. With the help of a DNA marker, the size of the studied DNA fragment can be extrapolated.

Protocol:
Horizontal gels (7 x 10 cm) with 1 % (w/v) agarose in TA buffer (40 mM Tris acetate, 10 mM Na acetate, 1 mM EDTA, pH 7.8) were used to analyse PCR products, plasmid extracts and restriction digestion products. Gels were run at 80 V (~11.43 V/cm) for 30 to 45 minutes, stained with the DNA intercalating agent ethidium bromide (~1 mg/l) for 5 minutes, washed in deionized water for 10 minutes and photographically documented under UV light on the transilluminator.

2.5.4.2 Quantification of DNA and RNA

The concentration of DNA and RNA was determined using ND-1000 Spectrophotometer (NanoDrop).

2.5.5 Cloning

2.5.5.1 Restriction digestion of DNA

Restriction enzymes are sequence specific endonucleases which cleave DNA, resulting in blunt ended DNA or DNA with single strand overhangs at the 3´-OH or 5´-phosphate end. MilliQ H_2O, 2 µg DNA, 10 x One Phor All buffer (final concentration 1 x or 2 x; depending on restriction enzyme properties) and 1 µl of each restriction enzyme (10 U/µl) were mixed in a total volume of 20 µl. Incubation was carried out at 37 °C for 2 hours. Restriction enzymes were inactivated for 20 min at 65 °C.

2.5.5.2 Ligation of restriction digestion products

DNA molecules, which were digested by the same restriction enzyme or an isoschizomer, can be ligated using the ATP dependent T4 ligase. The ligase catalyses the formation of the phosphodiester bond between the 3´-OH and 5´-phosphate ends of two DNA molecules.

Protocol:
Purified restriction digested DNA fragments and purified restriction digested plasmid DNA were used in one ligation mix. Ligations were performed by mixing plasmid and insert (molar

ratio 1 : 2) with 1 U of T4 DNA ligase (Promega) and its appropriate buffer. The tube containing the ligation mix was placed in an open thermos bottle containing water. The thermos bottle was incubated over night is incubated overnight at a gradient from 24 °C to 4 °C. The ligation mix was then used for transformation or stored at –20 °C until further use.

2.5.6 DNA Transfer

2.5.6.1 Transformation of *E. coli*

Bacterial transformation is the process by which bacterial cells take up naked DNA molecules. In the transformation procedure, the cell walls of bacteria are treated in a way that the DNA can pass into the cytoplasm. *E. coli* is not naturally competent, which means it cannot be transformed with plasmid DNA without prior treatment. *E. coli* cells can be made competent for transformation with the treatment with $CaCl_2$ solution.

- Preparation of *E. coli* DH5α chemically competent cells

The protocol followed was that of Sambrook and Maniatis [126]. 10 ml of LB medium were inoculated with *E. coli* DH5α or S17-1 and grown with agitation overnight at 37 °C. 0.2 ml of the overnight culture was subcultured into a reaction tube with 10 ml of fresh LB medium and grown to an OD_{600} of 0.4-0.6. The cell culture was centrifuged for 5 min at 6000 rpm at 4 °C. The supernatant was discarded, the pellet resuspended in 10 ml ice cold 100 mM $CaCl_2$ and incubated on ice for 30 min. The cells were then pelleted at 6000 rpm, 5 min at 4 °C. The supernatant was discarded and the pellet resuspended in 1 ml 100 mM $CaCl_2$ and 200 µL of glycerol. 100 µL aliquots of this cell suspension were pipetted into Eppendorf tubes, frozen in liquid nitrogen and stored at –80 °C until needed.

- Transformation of chemically competent cells

50 ng of either purified plasmid or plasmid in a ligation mixture were mixed with 100 µl $CaCl_2$ competent *E. coli* cells on ice. The transformation system was then incubated for 30 min on ice, then 60 sec at 42 °C in a water bath, followed by 2 more minutes on ice. Thereafter, 1 ml of SOC medium was added and the cells were incubated at 37 °C for 60 min. 100 µl of the transformed cell suspension were plated on agar plates containing antibiotics

selective for the transformed plasmid. The rest of the cell suspension was centrifuged at 5000 rpm for 5 min and 800 µl of the supernatant were discarded, the cells were resuspended in the remaining volume and then plated. The plates were incubated at 37 °C for 18 hours. Single colonies of the transformants were inoculated in liquid medium, incubated at 37 °C for 12 hours and analysed by polymerase chain reaction (PCR) or restriction digestion of extracted plasmid DNA. Positive samples were sent for sequencing, for confirmation of correct cloning and accurate insert DNA sequence.

2.5.6.2 Conjugation of *E. coli* S17-1 and *S. meliloti*

Conjugation is the direct transfer of mobilizable plasmid DNA from one bacterium to another. In this process species barriers can be overcome by the use of appropriate mobilizable vector plasmids. Since *S. meliloti* cannot be transformed with plasmid DNA, plasmids can only be introduced by conjugational transfer from an *E. coli* strain containing an appropriate plasmid. To mobilize plasmids from *E. coli* to *S. meliloti* the *E. coli* strain S17-1 can be used as a donor [134]. It contains an RP4-derivative (*oriT*RP4) integrated into the genome. This allows the conjugational transfer of plasmids containing the *oriT*RP4 but lacking the mobilization genes. Therefore, these plasmids will not be mobilized again, if they are passed on to a recipient, which does not contain the RP4 mobilization genes.

Protocol:
E. coli S17-1 and *S. meliloti* were grown to stationary phase in liquid cultures at appropriate conditions. 0.1 ml of the over night S17-1 culture were subcultured into 10 ml fresh selective medium and grown at 37 °C to an OD_{600} of 0.5. Then 0.5 ml of the logarithmically growing donor culture were mixed with 0.8 ml of stationary recipient culture in an Eppendorf tube, spun down at 6000 rpm for 30 sec and the supernatant discarded by decanting. The pellet was resuspended in the flow back and the cell suspension was applied on a cellulose acetate filter (Sartorius) placed on TY plate without antibiotic. The plate was incubated overnight at 30 °C. After growth overnight at 30 °C the filter was taken from the plate and cells were washed from it using 700 µl of 0.85 % NaCl. 100 µl of a series of dilutions of the suspension, ranging from 10^{-2} to 10^{-5}, were then plated on plates with the required antibiotics.

2.6 POLYMERASE CHAIN REACTION (PCR) TECHNIQUES

2.6.1 Amplification of DNA using polymerase chain reaction

Polymerase chain reaction is a technique for enzymatic replication of DNA *in vitro*. PCR uses two primers that anneal to the sense and antisense strands at the ends of the DNA fragment which has to be amplified. The cycles of DNA denaturing - primer annealing - second strand DNA synthesis lead to the amplification of the chosen DNA fragment. Termostable Pfu polymerase (Fermentas) was used in all PCR reactions in this work, according to the manufacturer's instructions. After the PCR, the products were checked in an agarose gel and if applicable purified using the QIAquick PCR Purification Kit (250) (QIAGEN) PCR amplification was performed as following. For a PCR system, 20 pmol of each primer, 0.5 U of Pfu and 1X Pfu buffer, as well as 0.2 mM of dNTPs were mixed with 100 ng of bacterial DNA. Sterile distilled H_2O was added to a final volume of 25 µl. Amplification was carried out as follows, 95 °C for 2 minutes, then 30 cycles of denaturation at 95 °C for 15 seconds, annealing (adequate annealing temperature of the primer) for 30 seconds, extension at 72 °C for a time period related to the length of the product (1 minute for 1 Kb). Final steps comprised those of 72 °C for 8 minutes and 8 °C until the system was moved and stored at -20 °C.

2.6.2 Gene SOEing

The method of Gene SOEing allows for the combination of DNA sequences. Through gene SOEing it is possible to create marker-free deletion mutants in *S. meliloti* [67]. In a first PCR regions up- and downstream of the desired deletion were amplified, and then they were fused in a second PCR. The deletion constructs obtained were subsequently cloned into the suicide vector pK18mobsacB, which allows sucrose selection for vector loss [130]. The resulting plasmids were conjugated into *S. meliloti* via *E. coli* S17-1 to introduce deletions by allelic exchange. Production of mutant strains was confirmed by PCR reactions with check primers designed to amplify DNA fragments spanning the gene of interest.

2.6.3 Two step quantitative real-time reverse transcription PCR (RT RT-PCR)

The changes in transcriptional activity of a gene can be measured with reverse transcription real-time PCR. In the first step, extracted and purified DNA-free RNA is reverse transcribed by a reverse transcriptase using random hexameric primers. It is crucial to use the same amount of RNA from each sample in real-time RT-PCRs, since the amplification factor of the very sensitive PCR can significantly influence the results. The cDNA obtained in the first step is then used for PCR reactions with gene specific primers. During real-time PCR, the amount of PCR product is measured after each cycle and a Ct value (the point at which the amount of PCR product increases linearly) can be obtained. To normalize the amount of RNA from each extract, sets of primers for genes have to be included in the real-time PCR that show no change in their expression level under the experimental conditions. These are often primers for housekeeping genes.

Reverse transcription

The RNA concentrations of all extracts were measured and ~9 µg RNA were used during reverse transcription (30 µl-scale). ~9 µg RNA were mixed with 1 µl 5 µg/µl of random hexamers and filled up with RNase-free H_2O to a final volume of 18 µl. These mixes were first incubated for 10 min in a 65 °C water bath and then for 5 min on ice. Condensed water was spun down briefly. Real-time PCR was carried out using the QIAGEN Quantitect SYBR Green reverse transcription-PCR kit Green kit according to the manufacturer's instructions. 15 µl 2 x Quantitect Master Mix, 0.5 µM forward primer and 0.5 µM reverse primer, 1 µl cDNA and MilliQ H_2O to 30 µl were mixed and loaded onto a 96 well plate for the continuous fluorescence detector (DNA Engine OPTICON, MJ Research).

The program for real-time PCR was the following:

Step	Temperature / Time
1 initial DNA melting	95 °C / 15 min
2 cycle DNA melting	95 °C/ 20 sec
3 annealing	56 °C / 30 sec
4 elongation	72 °C / 45 sec
5 scan plate	

6 cycling: go to step 2 for 40 times
7 perform melting curve from 60 to 90 °C;
read every 1 °C; hold for 1 sec between reads
8 end

Analysis of Real-time PCR results

The Ct values of all samples were loaded into a Microsoft Excel file and processed with a normalization algorithm [102]. The resulting mean normalized expression values were than used to calculate ratios of up or down regulation of the assayed genes under the conditions tested. The housekeeping gene *Smc02641*, which codes for a glucose dehydrogenase, was used for normalization.

2.6.4 Purification of PCR products with the QIAquick PCR Purification Kit

Purification of PCR products with the *QIAquick PCR Purification Kit* was carried out according to the manufacturer's protocol for purification using a microcentrifuge (centrifugation steps were carried out at room temperature for 1 minute at 13,000 rpm). DNA fragments ranging from 100 bp to 10 kilobases (kb) were purified from primers, nucleotides, polymerases, and salts on a QIAGEN unique silica-gel membrane using special buffers provided with the kit.

Protocol:
- 5 volumes of Buffer PB were added to 1 volume of PCR sample and mixed by pipetting up and down.
- To bind the DNA, the samples were applied to QIAquick columns and centrifuged. The flow through was discarded
- To wash the DNA, 750 µl of Buffer PE was added to the QIAquick columns and centrifuged. The flow through was discarded and the columns spun for another minute to remove residual ethanol from Buffer PE.
- To elute the DNA, 30-50 µl autoclaved MilliQ H_2O were added to the center of the membrane. The columns were transferred to new Eppendorf tubes. To increase the yield of the elution the columns with the applied water were incubated 1 min at room temperature before centrifugation.

2.7 MICROARRAY ANALYSES

2.7.1 Transcriptional profiling using the SM14kOligo whole genome microarray

For microarray hybridization, three independent bacterial cultures from each condition were prepared as biological replicates for RNA isolation. Accordingly, for each time point, dual-fluorescence-labelled cDNA probes were prepared to hybridize with three slides, respectively. For each preparation of Cy3 and Cy5 labelled cDNAs, 10 µg of total RNA were used [34]. To each microarray, the cDNA of the pH 7.0 and pH 5.75 grown cultures were mixed and hybridised. Slide processing, sample hybridization, and scanning procedures were performed applying the Sm14kOligo microarray, that carries 50 mer to 70 mer oligonucleotide probes directed against coding regions and intergenic regions [135]. Analysis of microarray images was carried out applying the ImaGene 6.0 software (BioDiscovery) as described previously [77]. Lowess normalization and significance test (fdr) were performed with the EMMA software [34, 36]. M-values ($\log_2$ experiment/control ratio), P-values (*t* test) and A-values were also calculated with EMMA. The M-value represents the logarithmic ratio between both channels. The A-value represents the logarithm of the combined intensities of both channels. Fluorescently labeled purified PCR products were were lyophilized and resuspended in 110 µl of DIG Easy Hyb solution. Hybridization was carried out at 36 °C for 1 h in the HS4800 Hybridization Station (Tecan). Before applying the hybridization sample to the microarray, it was denatured for 3 minutes at 95 °C. Following hybridization, the arrays were washed twice in 2 × SSC, 0.2 % SDS for 5 minutes at 30 °C and subsequently twice in 0.5 × SSC for 2 minutes at 20 °C [77, 123].

Slide processing

Protocol:
- place the slides in a plastic rack and carry out the processing by transferring the racks from one container to the other, lift the rack up and down during washing
- wash slides for 5 min at room temperature in 250 ml of rinsing solution 1
- wash slides for 2 min at room temperature in 250 ml of rinsing solution 2, twice
- wash slides for 10 min at room temperature in 250 ml of rinsing solution 3

- wash slides for 1 min at room temperature in 250 ml of MilliQ H_2O
- incubate slides for 15 min at 50 °C in 200 ml prewarmed blocking solution in a glass container, with constant shaking.
- wash slides for 1 min at room temperature in 250 ml of MilliQ H_2O
- place rack on an 12 × 8 cm plastic microplate cover (Genomics Solutions) containing 2 Kim-wipes and immediately spin in the microplate centrifuge at 1200 rpm for 3 min. Use a stack of 3 used glass slides at every side of the plastic dish to lift up the rack with the slides, this avoids precipitation artifacts at the side of the slide. Be sure to counter-balance using an appropriate balance.

Checking fluorescently labeled targets on agarose gels prior to microarray hybridizations

Protocol:
- combine 2 µl of the labeled target with appr. 4 µl of 80 % (v/v) glycerol
- run a 0.8 % (w/v) agarose gel in TA buffer for 20 min at 80 V
- load BPB-marker as a positive control
- place the agarose gel in -90° orientation on the Typhoon Imager
- start the "Typhoon Scanner Control" software
- load the Cy5/Cy3 medium template file with the acquisition for fluorescence orientation: 90° pixel size: 200 µm focal plane: 3 mm.
- scan the gel at 633 and 532 nm to detect Cy5-red and Cy3-green labeled target cDNA.

2.7.2 Microarray data analysis

Pre-processing

Image processing was performed with ImaGene (version 6.0.1). For each spot the background corrected spot intensities were calculated using the means of all chosen pixels for background and signal. Negative spots or spots that were flagged as empty or having bad quality were removed. The mean intensity (a-value) was calculated for each spot using the standard formula $a_i=\log_2(R_iG_i)^{0.5}$. $R_i=I_{ch1i}-Bg_{ch1i}$ and $G_i= I_{ch2i}-Bg_{ch2i,}$ where I_{ch1i} or I_{ch2i} is the intensity of a spot in channel 1 or channel 2 and Bg_{ch1i} or Bg_{ch2i} is the background intensity of a spot in channel 1 or channel 2, respectively. The logarithm to the base 2 of the ratio of intensities

($\log_2$ experiment/control ratio or m-value) was calculated for each spot using the formula $m_i=\log_2(R_i/G_i)$ [11].

Normalization and filtering

Normalization englobes the dealing with systematic differences between the two fluorescent dyes caused by eventual differences in labelling and detection efficiencies. Because of the differences, the data distribution is not centered at zero although often most of the genes should have similar values in experiment and reference. In a normalization procedure, a normalization factor *c* is calculated and used to move the center of the distribution to zero, so that $\text{m-value}_{\text{norm}}(i) = \text{m-value}(i) - c$ for spot *i.* The data from the pilot verification experiments was normalized using locally weighted regression (LOWESS) [25]. This method accounts for differences between the two dyes that depend on intensity (a-values) [155]. Data used for the calculation is assumed to have an m-value of zero.

Clustering analysis of the microarray data

K-means clustering analysis of the microarray time-course data was performed with the aid of the Genesis software [138]. The Euclidean distance between two data points x and y were applied in the clustering algorithm:

Euclidean distance : $d(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$

Uncentered correlation: $d(x,y) = 1 - r_u$, where $r_u = \dfrac{\sum_{i=1}^{n}(x_i)(y_i)}{\sqrt{\sum_{i=1}^{n}(x_i)^2}\sqrt{\sum_{i=1}^{n}(y_i)^2}}$

After normalization, only genes with at least threefold change in expression (M-value of ≥ 1.4 or $\leq$ -1.4) in at least one point of time in the wild type microarrays were considered for clustering analysis. Genes that did not present an evaluable expression value for at least 5 of the 6 points of time (missing values on the microarray flagged as empty spots) were not

considered. K-means clustering was used for distributing differentially regulated genes into 6 groups, both with the wild type and with the *rpoH1* mutant microarray data. The data matrix for clustering analysis after filtering of absolute m-values consisted of 6 time points. The algorithm was repeated 100 times for clustering. The solution with the smallest sum of distances within clusters was chosen as the final result. In the *k*-means clustering, number of clusters *k* was set to 6. In the experiments, the distance was calculated based on uncentered Pearson correlation.

Uncentered correlation: $d(x,y) = 1 - r_u$, where $r_u = \frac{\sum_{i=1}^{n}(x_i)(y_i)}{\sqrt{\sum_{i=1}^{n}(x_i)^2}\sqrt{\sum_{i=1}^{n}(y_i)^2}}$

2.7.3 Microarray data accession numbers

The entire set of microarray data has been deposited in the ArrayLims database [36].

2.8 CAS SIDEROPHORE ASSAY

CAS assay is a highly sensitive chemical colorimetric method for the detection of siderophores, which is based on their affinity for iron(III) and is therefore independent of the structure. The following chemical equation explains the principle:

FeDye3-’ + L”- * FeL3-” + Dye’-

A strong ligand L (e.g., a siderophore) is added to a highly colored iron dye complex. When the iron ligand complex is formed, the release of the free dye is accompanied by a change in color. Chrome azurol S (CAS) assay mixtures for siderophore detection were prepared as described by Schwyn and Neilands [132]. A 6-ml volume of 10 mM HDTMA solution was placed in a l00 ml volumetric flask and diluted with water. A mixture of 1.5 ml iron(III) solution (1 mrvr FeCl, * $6H_2O$, 10 mM HCl) and 7.5 ml 2 mM aqueous CAS solution was slowly added under stirring. Supernatants of *S. meliloti* cultures grown in VMM were mixed 1:1 with a CAS assay solution. After equilibrium was reached, the absorbance at 630

nanometers was measured. The relative siderophore activity was determined by measuring optical density ratios of different cultures.

C. RESULTS

1. CHARACTERIZATION OF *SINORHIZOBIUM MELILOTI* 1021 AND SIGMA FACTOR DELETION MUTANTS UNDER STRESS CONDITIONS

Alternative sigma factors play an important role in that they can respond with transcriptional activation of specific genes under environmental stress conditions. Considering the presence of multiple genes coding for alternative sigma factors in the genome of *S. meliloti [44]*, sequence analyses, as well as experimental characterization of *S. meliloti* wild type and sigma factor mutants were foremostly pursued. The aim of this first approach was to gain detailed insight into which sigma factors most actively take part in specific stress conditions, in order to start to unravel the complex stress response system of *S. meliloti*.

1.1 SEQUENCE ANALYSIS OF ALTERNATIVE SIGMA FACTORS

Alternative sigma factors of *S. meliloti* belonging to the RpoE and RpoH families were analysed for the conservation in their amino acid sequences. The nine *S. meliloti* RpoE sigma factors are phylogenetically related and possess the 2.1, 2.4, 4.1, and 4.2 subregions, which are typical domains of the ECF sigma factors conserved among α-proteobacteria. Alignments of the amino acid sequences of *S. meliloti* RpoE sigma factors showed high conservation among the nine proteins (Figure 4). The sequences of the RpoE sigma factors of *S. meliloti* have highest amino acid conservation at the 2.1, 4.1 and 4.2 subregions. They all also share at least 50% sequence similarity to *E. coli* RpoE.

The RpoH sigma factors are also well conserved among rhizobia (Figure 5). RpoH1 of *S. meliloti* shows 86% of amino acid identity to *A. tumefaciens*; 60%, 70% and 50% identity to RpoH1, RpoH2 and RpoH3 of *B. japonicum*, respectively; and 58% and 46% to *R. etli* RpoH1 and RpoH2. Surprisingly, there is only 43% identity between the amino acid sequences of RpoH1 and RpoH2 of *S. meliloti*, even less identity than to RpoH sequences of other bacteria. The RpoH box, a sequence of nine amino acid residues, is conserved among all the predicted proteins [105]. This region is reported to be involved in the interaction of the sigma factor with the core enzyme [111]. Additionally, phylogenetic analysis of RpoH sigma factors of rhizobial species was performed with the aid of ClustalW TreeView programs.

C. crescentus RpoH was included for comparison. The analysis revealed that the tree based on the amino acid sequences of RpoH sigma factors is similar to the known phylogenetic tree of rhizobial species based on 16S rRNA sequences (Figure 5). Interestingly, the RpoH2 sigma factor of *S. meliloti* is more similar to the RpoH sigma factors of *R. etli*, whereas the RpoH1 sequence shares more similarity with the RpoH sequences of *A. tumefaciens* RpoH and *B. japonicum* RpoH1.

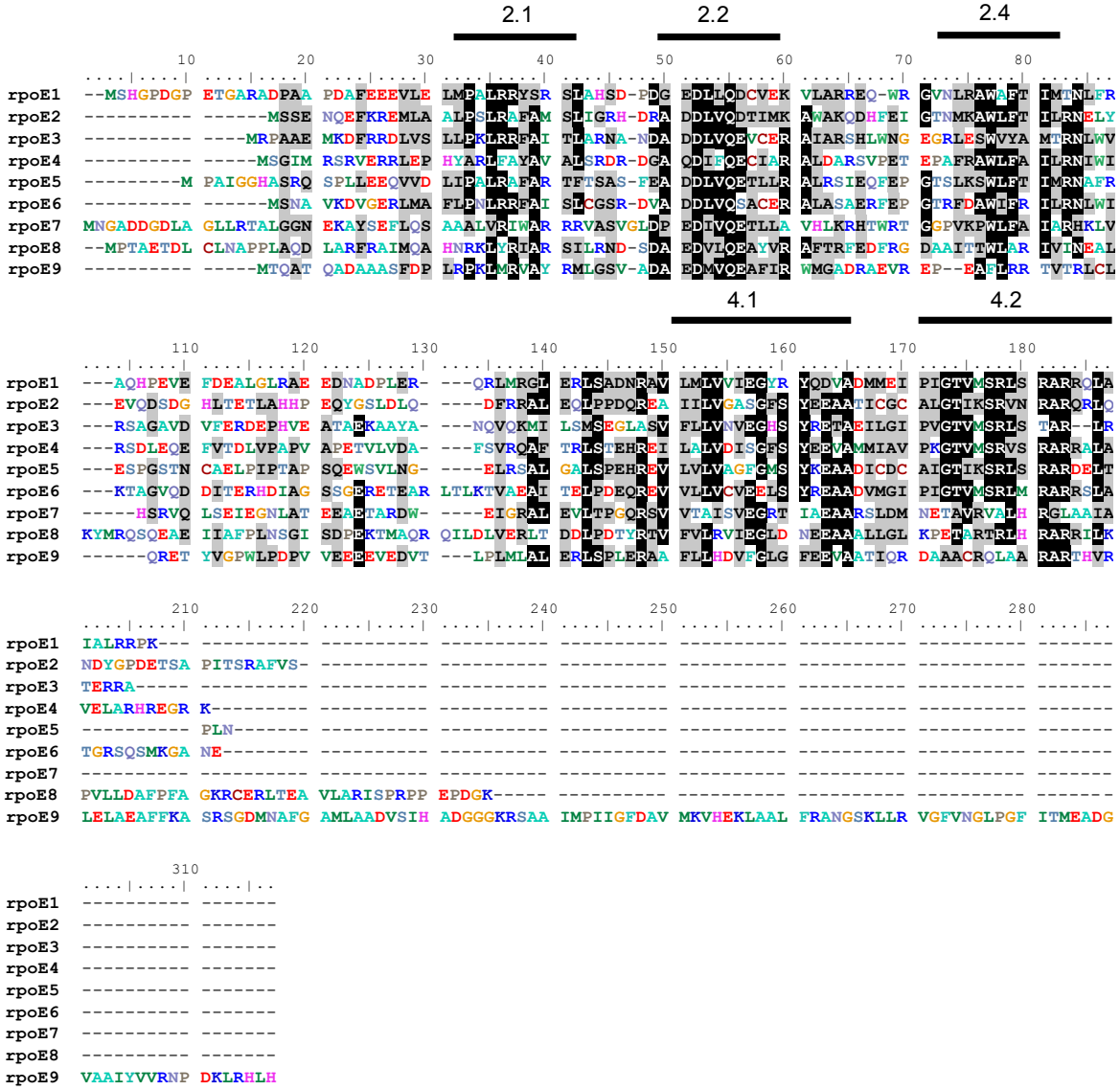


Figure 4. Alignment of predicted *S. meliloti* RpoE sigma factors. The conserved regions (and subregions) 2 (2.1, 2.2 and 2.4) and 4 (4.1 and 4.2) are indicated. Alignments were performed with ClustalW and BioEdit. Numbers on top indicate the amino acid position relative to the start of the proteins.

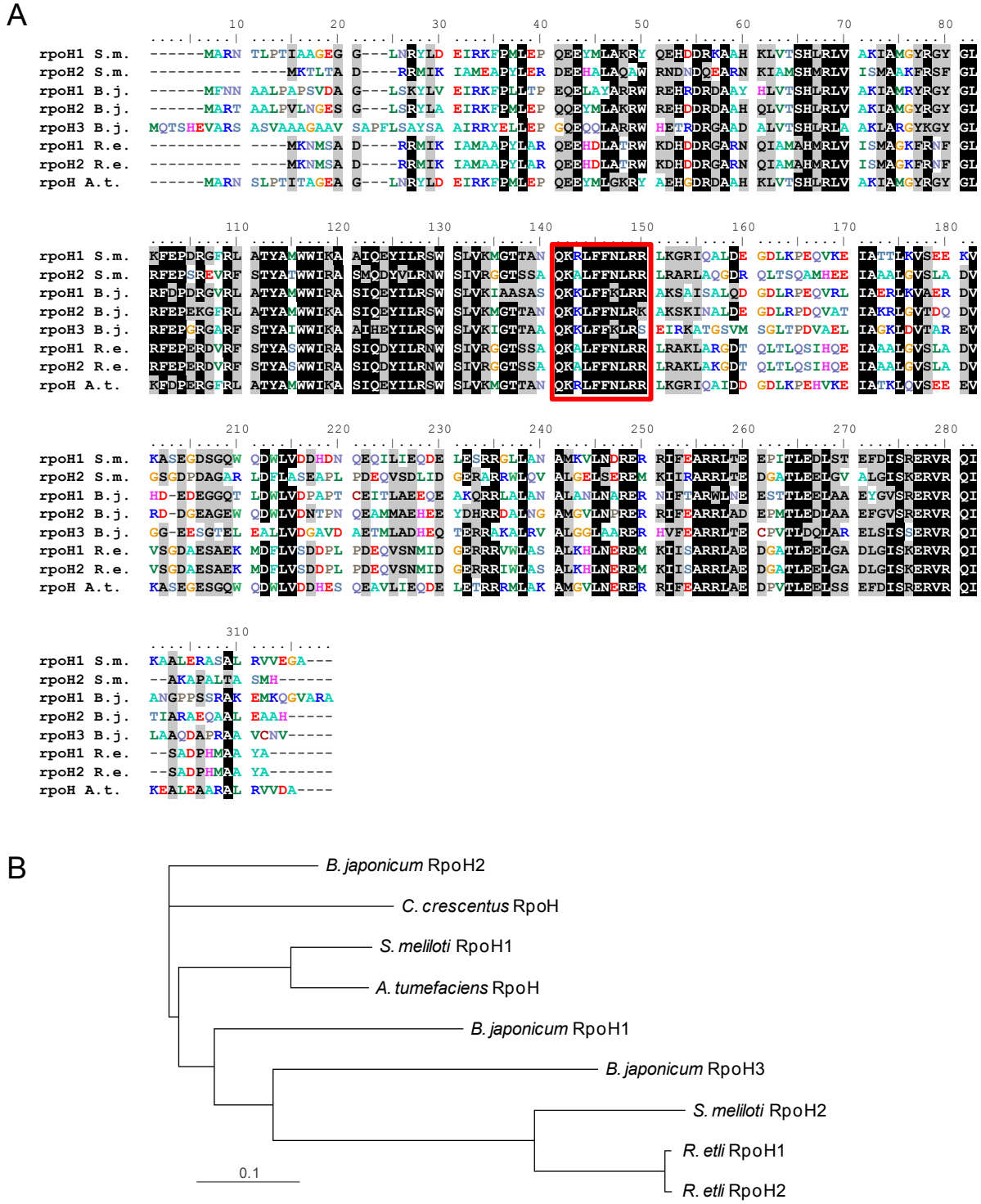


Figure 5. Rhizobial RpoH sigma factors. A) Alignment of 8 amino acid sequences of RpoH proteins from *S. meliloti*, *B. japonicum*, *R. etli* and *A. tumefaciens*, made with ClustalW. Numbers on top indicate the amino acid position relative to the start of the proteins. The RpoH box is shown in red. B) Phylogenetic tree of RpoH sigma factors. The phylogram was constructed with Tree View PHYLYP 1 after alignment with ClustalW. RpoH of *C. crescentus* was included for comparison.

1.2 EVALUATION OF SIGMA FACTOR GENE EXPRESSION UNDER DIFFERENT STRESS CONDITIONS

In order to evaluate which of the *rpoE and rpoH* homologs were implicated in the responses to environmental stress conditions, real time RT-PCR experiments were performed to assess the differential expression of sigma factor genes under stress. For that purpose, the expression of four *S. meliloti* sigma factor genes, *rpoE1, rpoE2, rpoE5* and *rpoH1*, was analysed under different stress conditions, namely, cold shock (15 °C), heat shock (38 °C), oxidative stress (3 mM H_2O_2) and pH stress (pH 5.75), at the time points 5, 10 or 20 minutes after the stress condition was applied. *S. meliloti* strain 1021 (also called Rm1021) was used as wild type in this and in all following experiments. Quantitative real time RT-PCR revealed that there was an overall increase in the expression levels of the extracytoplasmic function sigma factors under the assessed stress conditions, with the exception of *rpoE2*, which showed no significant expression levels under the studied environmental stress conditions (Figure 6). *rpoE1* showed little change in expression at the conditions tested, with a stronger response to oxidative stress. For the sigma factor gene *rpoH1*, the most significant expression levels were observed for heat shock and pH stresses. For the sigma factor *rpoE5*, expression levels increased more strongly under cold and heat shock, though an increase in expression was observed in all analysed conditions. The quantitative real time RT-PCR of the sigma factor genes showed that there is indeed a differential regulation of sigma factor genes in response to unfavourable environmental conditions and the genetic circuits under the regulation of those sigma factors are likely to enable the cells to handle environmental stress.

1.3 CONSTRUCTION OF SIGMA FACTOR DELETION MUTANTS

The following approach was to generate knockout mutants of *S. meliloti* sigma factor genes with the aim of identifying the physiological roles of those sigma factors in different stress conditions. Bearing this in mind, sigma factors which were already known to respond to heat or cold shock, as well as pH and oxidative stress were chosen to be mutated. Mutations of seven sigma factor genes, *rpoE1*, *rpoE2*, *rpoE3*, *rpoE5, rpoE7 rpoH1* and *fecI* were performed with the utilization of the Gene SOEing technique [67] (Figure 7). The mutant constructions include representatives of the three main functional classes of alternative sigma factors, namely extracytoplasmic function, heat shock and iron metabolism and control.

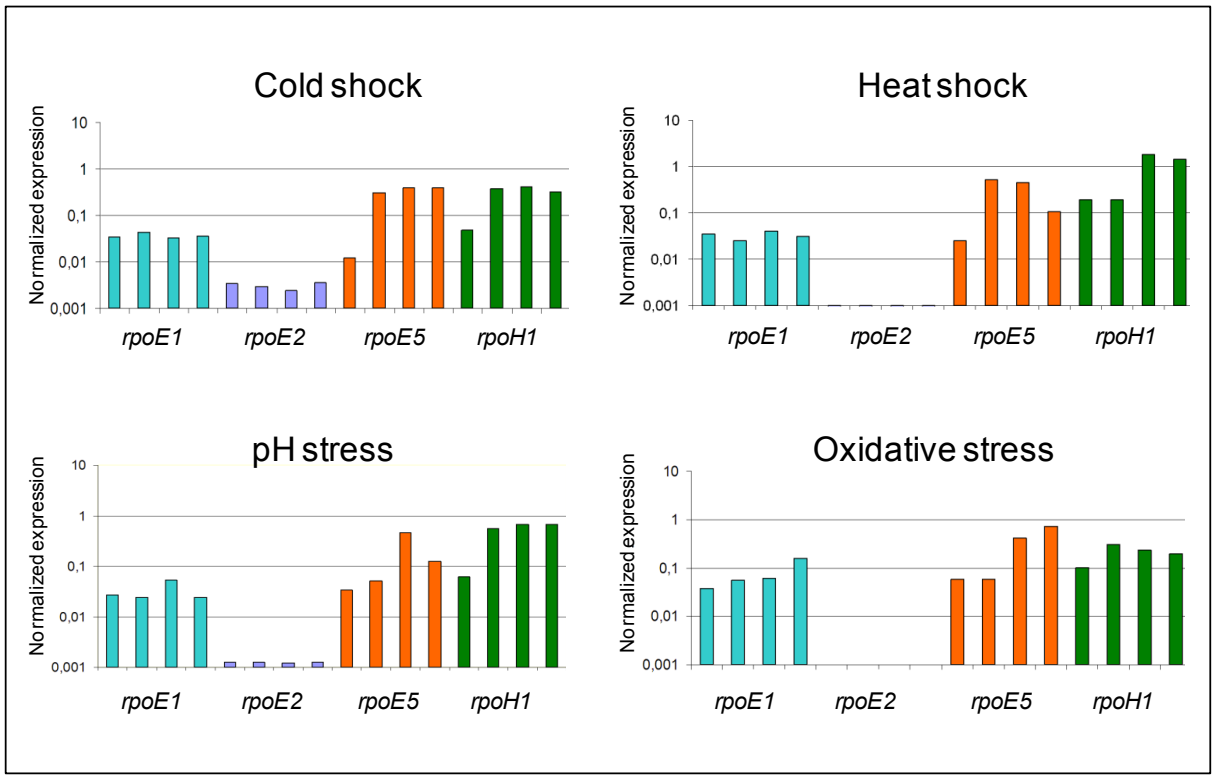


Figure 6. Quantitative real time RT-PCR of the sigma factor genes *rpoE1*, *rpoE2*, *rpoE5* and *rpoH1* under the environmental stress conditions cold shock (15°C), heat shock (38°C), oxidative stress (3 mM H_2O_2) and pH stress (pH 5.75), at different time points after the stress condition was applied. The housekeeping gene *Smc02641*, which codes for glucose oxidase, was used for normalization.

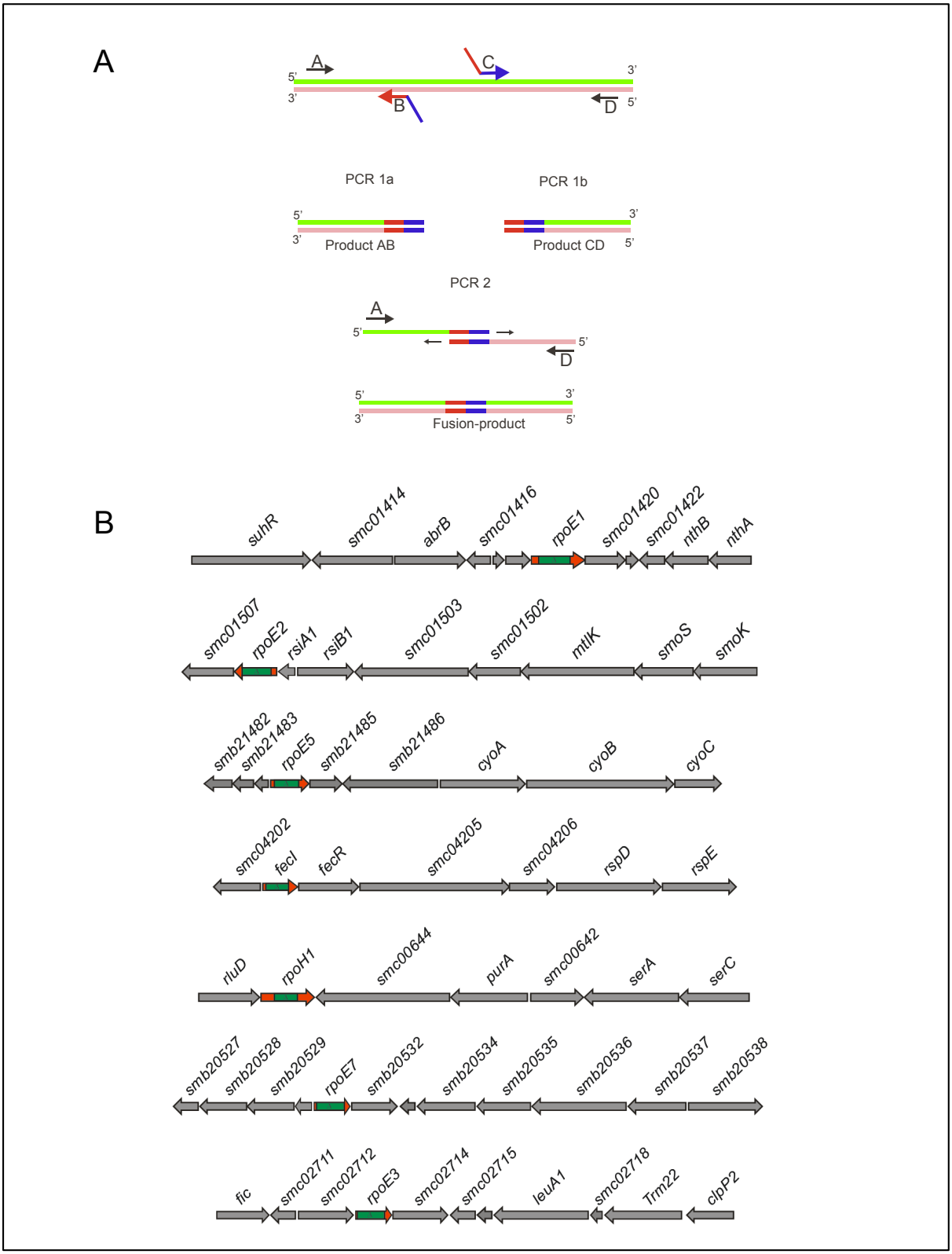


Figure 7. **Construction of the sigma factor deletion mutants.** A) Gene SOEing: regions up and downstream the gene were amplified and fused together in a second PCR. The deletion constructs were cloned into pK18mobsacB and the resulting plasmids were conjugated into *S. meliloti* to introduce deletions by allelic exchange (Schäfer *et al.*, 1994). B) Map of *S. meliloti* sigma factor genes with adjacent genes. Deleted regions are shown in green.

Polymerase Chain Reaction (PCR) experiments were carried out in order to get the mutated cassette with the deleted genes of interest and the obtained deletion constructs were ligated to the cloning vector pK18mobsacB [130]. The recombinant vectors were introduced into the *E. coli* strain DH5α and the success of the transformation was confirmed by diagnostic restriction endonuclease digestions of the plasmids purified from potentially recombinant bacteria. The plasmids bearing the construct were then transferred to the *E. coli* strain S17-1 for conjugation into *S. meliloti* in order to introduce the deletions by allelic exchange. The produced *S. meliloti* sigma factor knockout mutants were finally verified through PCR experiments with check primers located up and downstream of each deleted region, flanking therefore the deletion cassettes, and subsequent sequencing of the PCR fragments. Deletion mutations were established for all seven aforementioned sigma factor genes.

1.4 PHENOTYPICAL ANALYSES OF SIGMA FACTOR DELETION MUTANTS UNDER STRESS CONDITIONS

1.4.1 Heat stress

For the characterization of deletion mutants for sigma factor genes during heat stress, growth tests were performed with the deletion mutants and the wild type *S. meliloti* 1021, in TY medium, at 30 °C and 38 °C. The samples were measured for optical density at 580 nanometers, at different time periods, for five days. In those experiments it was clearly observed that the sigma factor deletion mutants grew slightly more poorly than the wild type at heat shock conditions. As expected, the mutant for the gene coding for the sigma factor RpoH1 did not grow at 38 °C (Figure 8). The mutants for the *rpoE2* and *rpoE5* genes also showed deficient growth in comparison to the wild type. The *rpoE2* gene has been described in the literature as responding to heat shock stress [129], though it had not been characterized in continuous growth at a high temperature. Also, *rpoE2* was not differentially expressed in the real time RT-PCR analysis at heat shock conditions.

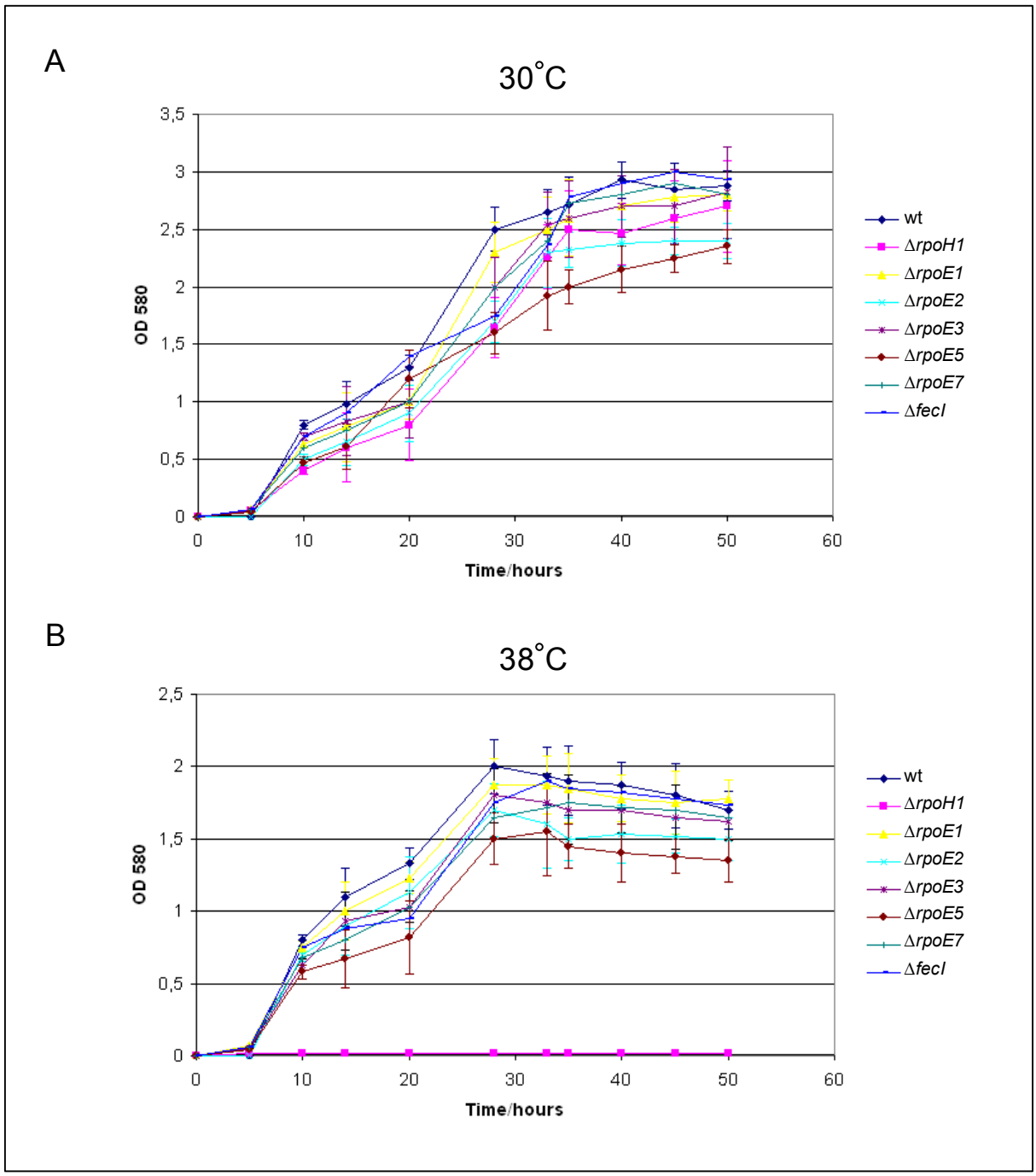


Figure 8. Characterization of *S. meliloti* wild-type and sigma factor mutants under heat stress. The growth of *S. meliloti* 1021 (blue) and mutants in TY medium at 30 °C (A) and 38 °C (B) was measured for optical density at 580 nm, at different time points, for 50 hours. Each panel shows the data from three representative experiments.

1.4.2 Cold stress

S. meliloti 1021 (wild type) and sigma factor mutants were analysed during continuous growth at 20 °C. In those experiments, though a slighltly deficient growth was observed for the sigma factor mutants in relation to the wild type, no significant difference was detected between the mutants and their wild-type counterpart, as it was seen for the *rpoH1* mutant under heat stress. A reduction in growth was observed for all strains as of 38 hours of growth and all behaved similar to the *S. meliloti* wild type (Figure 9). The strains also exhibited similar doubling times in both rich and in minimal media.

1.4.3 Oxidative stress

Rhizobia induce an oxidative burst when they invade plant roots [128]. In its free-living form, oxidative molecules are generated endogenously as by-products of aerobic metabolism. Thus, *S. meliloti* mutants were tested for sensitivity to 3 mM of H_2O_2. After reaching an optical density of 0.8 in Vincent minimal medium, cells were subjected to 3 mM of H_2O_2 for one hour, then plated overnight at 30 °C for analysis of viability by the number of colonies. The surviving fraction was calculated as the number of viable cells after treatment with or in the presence of the compound, divided by the number of viable cells in the absence of stress. In general, the wild type strain was more resistant to H_2O_2 than the sigma factor mutants (Figure 10). A viability reduction with respect to the wild type was observed when the *rpoE2* and *rpoH1* mutants were exposed to H_2O_2, with survival rates of 62% and 63%, respectively. The remaining sigma factor mutants did not generate significant differences in response to H_2O_2. The survival rates in relation to the wild type were 88% for the *rpoE1* mutant, 83% for the *rpoE3* mutant, 81% for the *rpoE7* mutant, 79% for the *fecI* mutant and 75% for the *rpoE5* mutant. In addition to its known role in protection against heat stress, RpoH has already been implicated in the oxidative stress response in *E. coli* and in *R. etli* [31, 87].

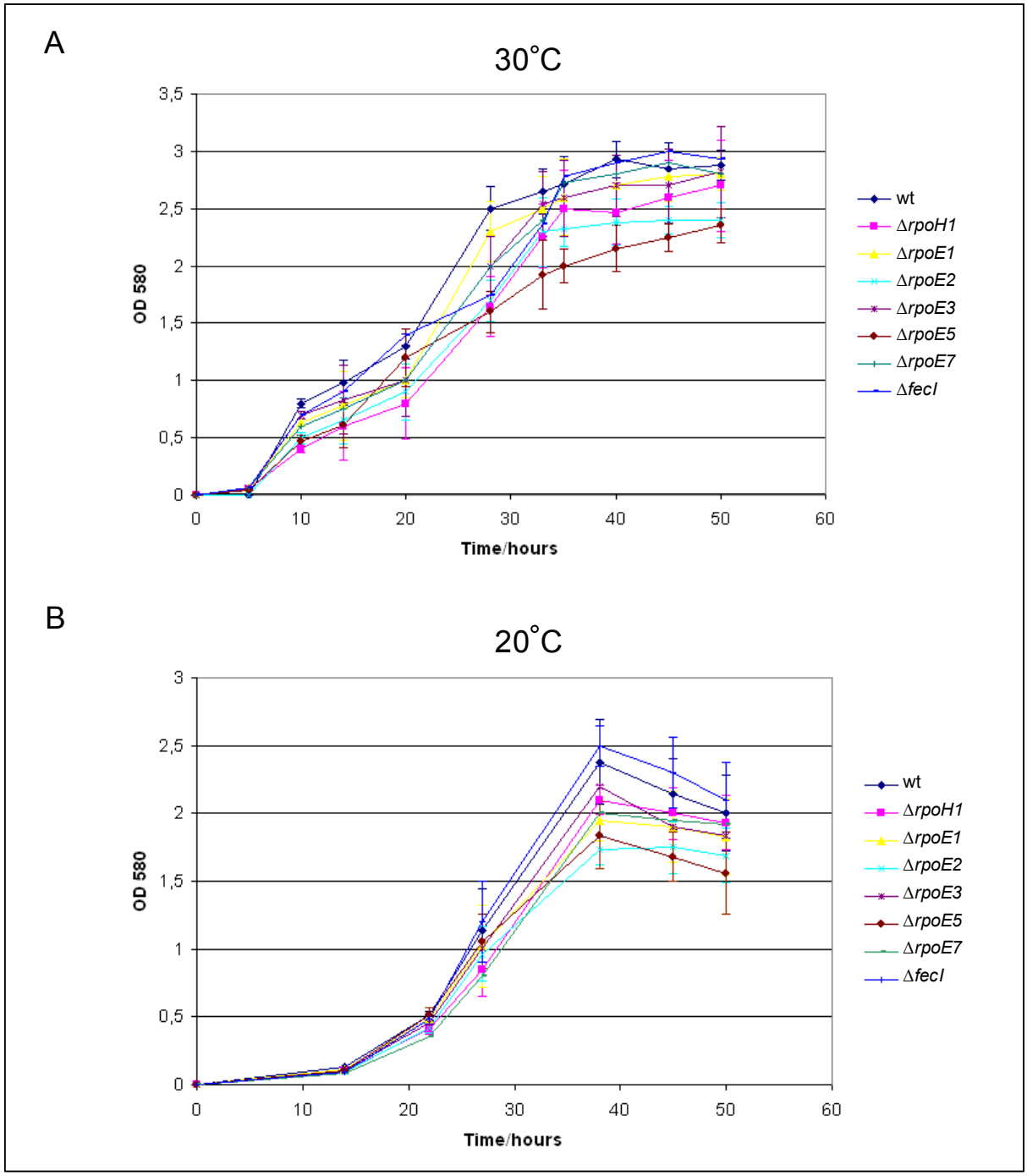


Figure 9. Characterization of *S. meliloti* wild-type and sigma factor mutants under cold stress. The growth of *S. meliloti* 1021 (blue) and mutants in TY medium at 30 °C (A) and 20°C (B) was measured for optical density at 580 nm, at different time points, for 50 hours. Each panel shows the data from three representative experiments.

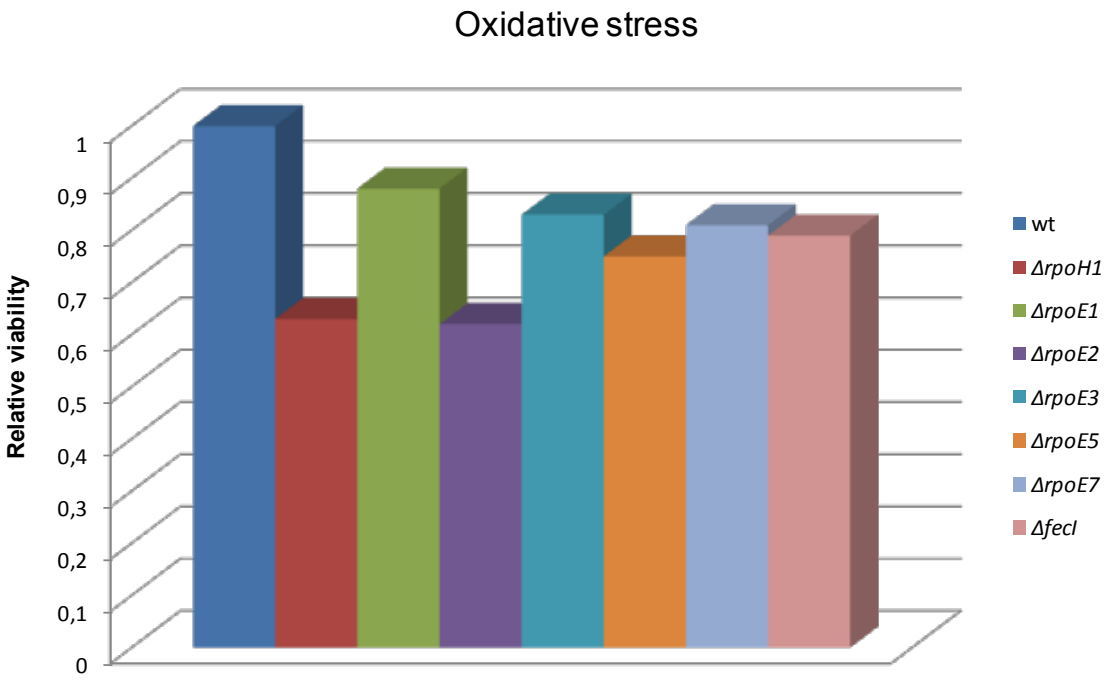


Figure 10. Viability of *S. meliloti* sigma factors mutants in response to oxidative stress. The viabilities of the sigma factor mutant strains were calculated in relation to the wild type after treatment with H_2O_2. Exponential-phase cultures were incubated with 3 mM H_2O_2 for 60 minutes and columns represent the surviving fractions of the mutant strain populations divided by the surviving fraction of the wild type population. The data presented are the averages of results from two independent experiments.

1.4.4 pH stress

In previous tests by Hellweg et al. [61], it was determined that the optimal pH stress condition for *S. meliloti*, that is, the pH value for which the cells are stressed but nevertheless survive, is pH 5.75. In order to test for a role for sigma factors in pH stress, the growth of sigma factor mutants was monitored in comparison to the wild type in VMM medium at pH 7 and at pH 5.75. The samples were measured for optical density at 580 nm, for eight days. In those experiments it was clearly observed that the sigma factor mutants reveal deficient growth behaviour in comparison to the wild type at low pH values. All five sigma factor mutants analysed grew more poorly in comparison to the wild type at pH 5.75. Strikingly, growth was severely impaired for the *rpoH1* mutant at pH 5.75. This mutant presented no growth at low pH (Figure 11). Complementation of the *rpoH1* mutant phenotype and further investigations for elucidating the role of RpoH1 in pH stress response are described in chapter 3.

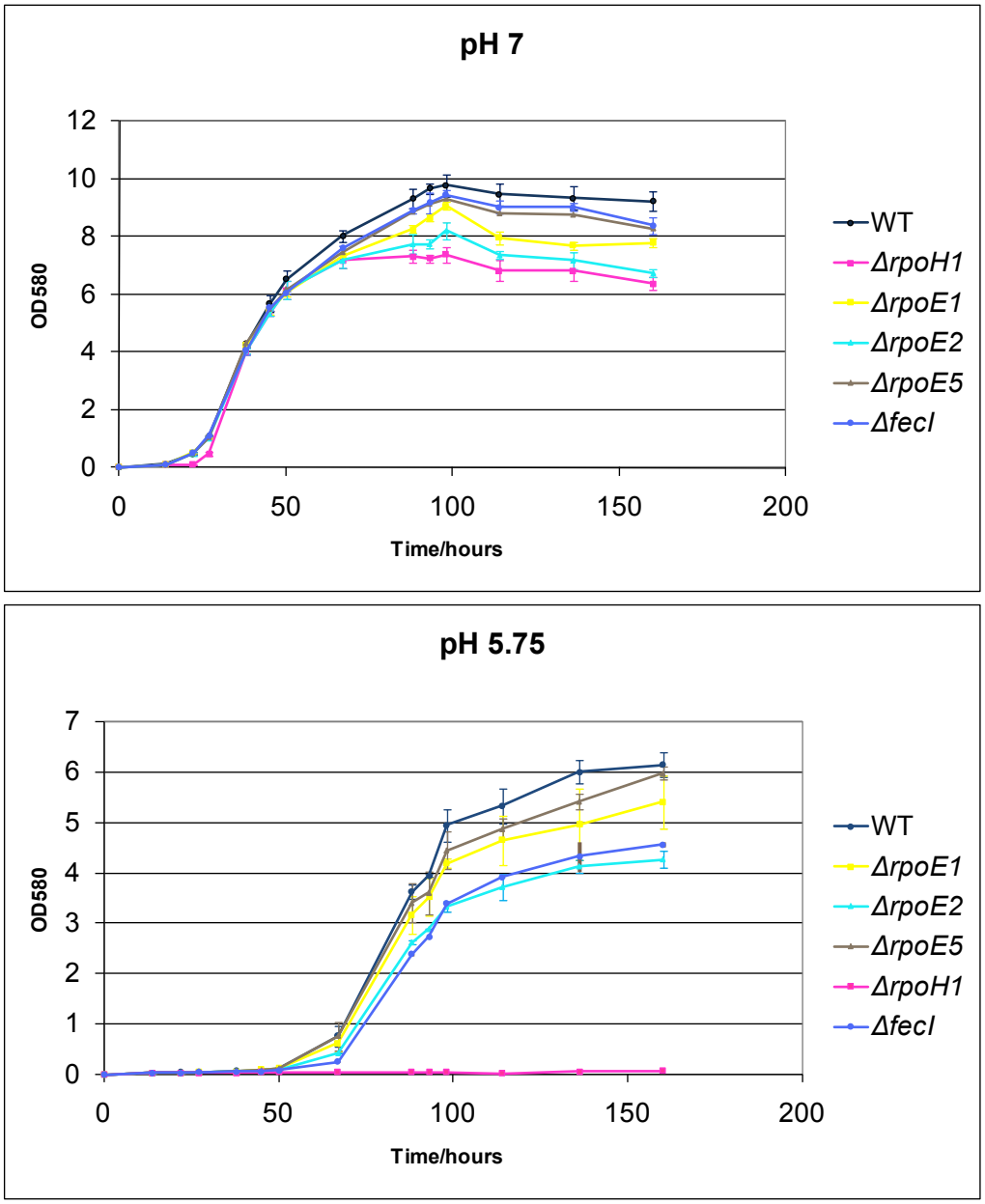


Figure 11. Growth curves of *S. meliloti* 1021 wild type strain and mutant strains for sigma factor genes at neutral and acidic pH. *S. meliloti* 1021 and mutant strains for sigma factor genes *rpoE1*, *rpoE2*, *rpoE5*, *fecI* and *rpoH1* were grown in VMM medium at 30°C at either pH 7.0 (A) or pH 5.75 (B). Each panel shows the data from three representative experiments. The error bars indicate the standard deviation calculated from three independent cultures.

2. RPOH1 AND THE REGULATION OF IRON UPTAKE

2.1 TRANSCRIPTION PROFILING OF THE *RPOH1* MUTANT VERSUS WILD TYPE AT NEUTRAL PH

Among all the sigma factors analysed, the *rpoH1* mutant was the only one that responded to almost all stress conditions tested. Moreover, among all the sigma factor mutants produced, the *rpoH1* mutant showed the most peculiar phenotypes, specially at high temperatures and low pH values, in which it presented no growth at all. For this reason, this mutant was selected to be more deeply investigated in this study. At first, transcription profiling experiments were performed for examining the differential expression of genes in the sigma factor *rpoH1* mutant in comparison to the wild type at neutral pH values. Both *S. meliloti* wild type strain 1021 and *rpoH1* mutant were cultivated at pH 7.0. After reaching an optical density of 0.8 at 580 nm, cell suspension probes were harvested and immediately centrifuged at 10000×g for 2 min at 30°C. The cell pellets were instantly frozen in liquid nitrogen for RNA preparation. Slide processing, sample hybridization, and scanning procedures were performed applying the Sm14kOligo microarray, that carries 50 mer to 70 mer oligonucleotide probes directed against coding regions and intergenic regions [135]. Analysis of microarray images was carried out applying the ImaGene 6.0 software (BioDiscovery) and Lowess normalization and significance test were performed with the EMMA software [34]. M-values ($\log_2$ experiment/control ratio), P-values (*t* test) and A-values were also calculated with EMMA. Only genes with a twofold difference in expression (M-value of ≥ 1 or ≤ -1) were considered significant. At neutral pH, the rhizobactin biosynthesis operon was observed among the significant differentially expressed genes (Figure 12). Rhizobactin is a well-known *S. meliloti* siderophore, a low molecular weight ligand that binds to ferric iron with high affinity [83]. All genes for the rhizobactin biosynthesis operon, *rhbABCDEF*, were upregulated, as well as the rhizobactin transporter gene *rhtA*. The gene for the rhizobactin activator *rhrA*, however, was downregulated in the mutant. Except for the genes involved in the rhizobactin siderophore biosynthesis and regulation, basically no other genes were differentially expressed in the *rpoH1* mutant at pH 7.0, in comparison to the wild type.

2.2 ASSESSMENT OF SIDEROPHORE PRODUCTION BY THE *RPOH1* DELETION MUTANT

In order to assess the production of siderophores by the *rpoH1* deletion mutant in comparison to the *S. meliloti* wild type, a CAS test was performed. The CAS test is a chemical assay for the detection of siderophore production. It is based on the the removal of ferric iron from an intensely pigmented complex by a competing ligand such as a siderophore. When a siderophore forms a complex with the ferric ion, the release of the free dye is indicated by a color change from blue to yellow [132]. Therefore, the CAS reagent provides a non-specific test for iron-binding compounds and the reaction rate is a direct indicator of the siderophore-concentration (Figure 13A).

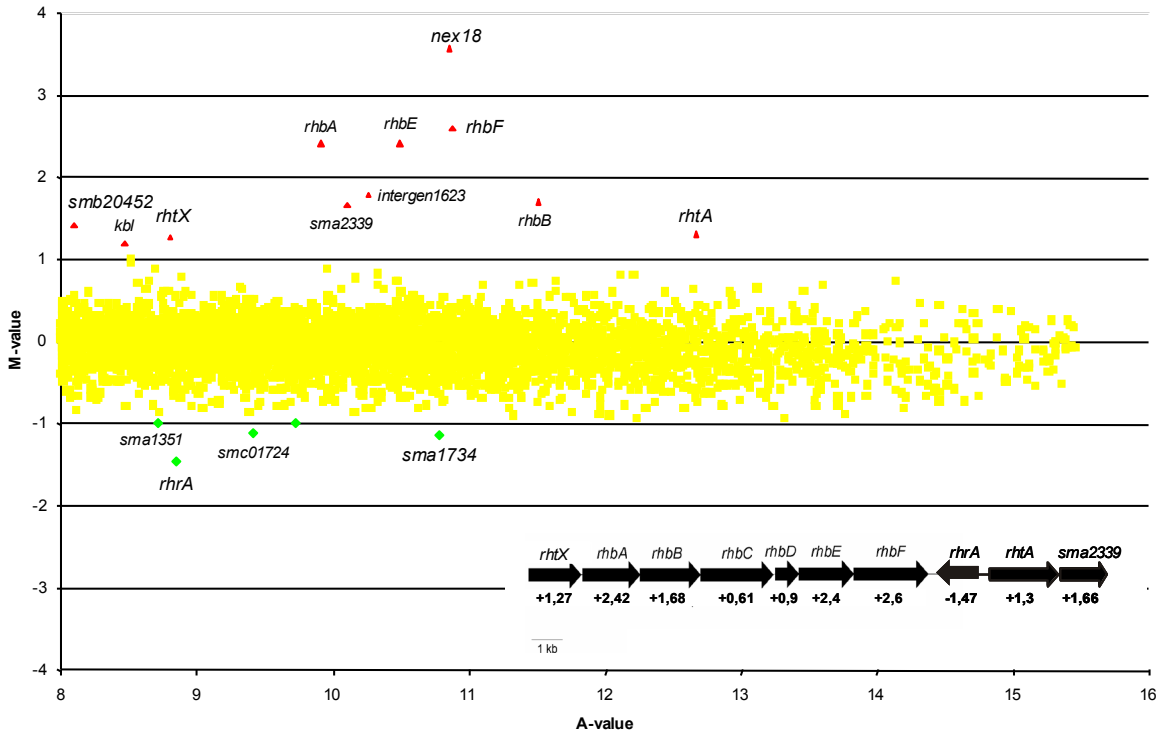


Figure 12. Scatter plot of the microarray analysis of the *S. meliloti rpoH1* mutant versus wild type at pH 7.0. The plot shows the $\log_2$ ratio (M-value) versus the mean signal intensity (A-value) obtained by comparison of the transcriptomes of *S. meliloti rpoH1* mutant versus *S. meliloti* wild type strain 1021. Genes with the greatest changes in expression values ($-1 \leq$ M-value ≥ 1) are indicated. On the low right corner is an illustration of the genetic map for the operon coding for proteins involved in rhizobactin 1021 biosynthesis and uptake. The numbers below the genes indicate the $\log_2$ expression ratios of the genes obtained through the transcriptome analysis.

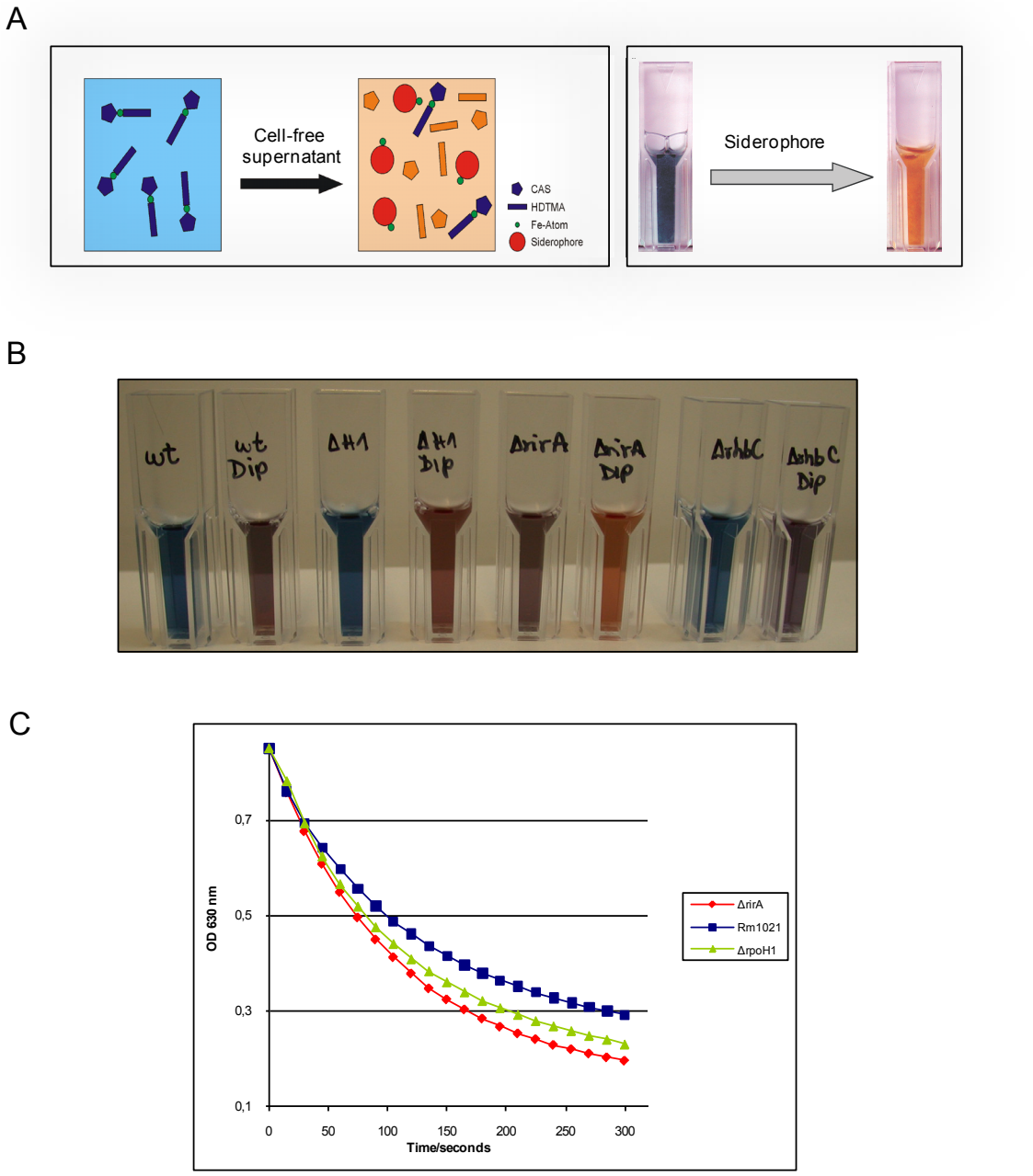


Figure 13. CAS Assay. A) Principle of CAS-assay functioning. The formation of the iron-siderophore complex results in a color change from blue to orange when the CAS dye is released. B) CAS-assay performed with *S. meliloti* wild type, *rpoH1* mutant, *rirA* mutant and *rhbC* mutant; Dip-2,2`-Dipyridyl (iron chelator). C) CAS time-course assessment of siderophore production. The time-course experiment was performed by measuring the optical density of the CAS-assay supernatant at 630 nm for five minutes, in 15-second intervals.

The increase in the expression of the rhizobactin operon in the *rpoH1* deletion mutant, in comparison to the *S. meliloti* wild type, was confirmed by the CAS assay (Figures 13B and C). Besides the *rpoH1* mutant and the wild type *S. meliloti* strain 1021, mutants for the iron global regulator gene, *rirA* [24], as well as mutants for a siderophore biosynthesis gene, *rhbC,* were also analysed for siderophore production in high and low iron conditions, as positive and negative controls of siderophore production, respectively. The low iron conditions were produced by growing the cells in medium containing 2,2-dipyridyl, which is a strong chelator with high specifity to iron. As expected, the *rhbC* mutant did not produce any siderophores, whereas in the wild type cells, siderophore production was observed in the iron deficient environment. Reflecting the absence of iron uptake regulation, siderophore production by the *rirA* mutant was observed in a high rate both high and low iron conditions. The color change observed in the assay with the *rpoH1* mutant indicates that more siderophores are produced by this mutant in comparison to the wild type production (Figure 13). Moreover, a time-course experiment was performed by measuring the optical density of the CAS-assay supernatant at 630 nm for the first five minutes after addition of the CAS solution, in 15-second intervals. The results showed that the *rpoH1* mutant produces siderophores in an intermediary rate between the wild type and the *rirA* mutant (Figure 13C). The results obtained by the CAS-assay corroborate the microarray results that the rhizobactin operon is indeed upregulated in the *rpoH1* mutant cells.

2.3 Regulation of the rhizobactin biosythesis operon

The *rhrA* gene codes for an activator of the rhizobactin biosynthesis. Even though there was a confirmed increase in siderophore production by the *rpoH1* mutant, the *rhrA* gene expression was downregulated in the microarray analysis. Therefore, with the aim of testing the hypothesis that another gene, other than the RhrA regulator, plays a role in the activation of the rhizobactin biosynthesis operon, a double mutant for the genes *rpoH1* and *rhrA* was produced. Such secondary activation would explain the fact that, even though the *rhrA* activator is downregulated in *rpoH1* mutant cells, the entire rhizobactin biosynthesis operon is upregulated. For this purpose, total RNA of *S. meliloti* wild type, *rpoH1* single mutant and *rpoH1 rhrA* double mutant cells was isolated for a real time RT- PCR experiment in which the expression levels of the rhizobactin biosyntheis operon gene *rhbA* were analysed. The

results obtained with this experiment showed that the expression of the *rhrA* regulator gene seems to be crucial for the expression of the rhizobactin biosynthesis operon, since basically no expression of *rhbA* was observed in the *rpoH1 rhrA* double mutant (Figure 14).

2.3.1 Upregulation of iron uptake is not the reason for the lack of growth of the *rpoH1* mutant at low pH values

Besides the *S. meliloti* wild type strain and the *rpoH1* mutant the *rpoH1 rhrA* double mutant was also analysed for its growth phenotype. All samples were grown in Vincent minimal medium and measured as triplicates, twice a day, for five days. (Figure 15). No significant difference was seen between te growth of the *rpoH1 rhrA* double mutant and the *rpoH1* single mutant. However, both strains present deficient growth in comparison to *S. meliloti* 1021.

To test whether the upregulation of iron uptake was the reason for the lack of growth of the *rpoH1* mutant at low pH values (Figure 11), in which iron becomes more soluble, a growth test was performed with variations in the iron concentration in the medium (0,37 or 37µM $FeCl_3$), as well as in the pH value (pH 7 or 5.75). The mutant cells did not grow even in medium at low iron concentration and low pH, indicating that the increase iron uptake is unlikely to be the reason for the lack of growth of *rpoH1* mutant cells at low pH values (Figure 16).

2.3.2 Expression analysis of the rhizobactin operon in *rpoH1* mutant cells after pH shock

With the aim of assessing the participation of the rhizobactin operon at low pH conditions, real time RT-PCR experiments were performed. A pH shock experiment was chosen for this analysis because the *rpoH1* mutant does not grow at all at continuously low pH conditions. Total RNA of *S. meliloti* wild type and *rpoH1* mutant grown at pH 7, as well as of *rpoH1* mutant cells grown at pH 7 followed by 10 minutes of pH shock (pH 5.75) were isolated for a real time RT-PCR experiment. The expression levels of the rhizobactin biosyntheis operon gene *rhbA* and the rhizobactin activator gene *rhrA* were then analysed.

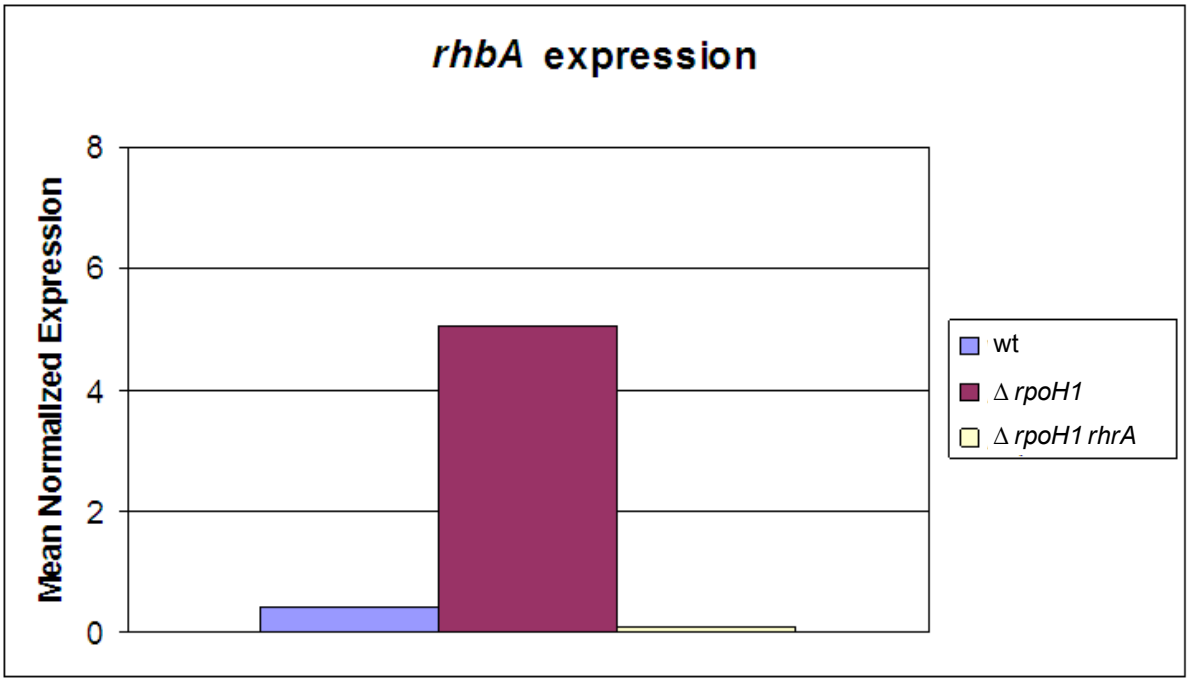


Figure 14. Quantitative real time RT-PCR of the rhizobactin biosynthesis operon gene *rhbA*. The expression levels were measured in wild type *S. meliloti* cells, as well as in the *rpoH1* single mutant and the *rpoH1 rhrA* double mutant. The housekeeping gene *Smc02641*, which codes for glucose oxidase, was used for normalization.

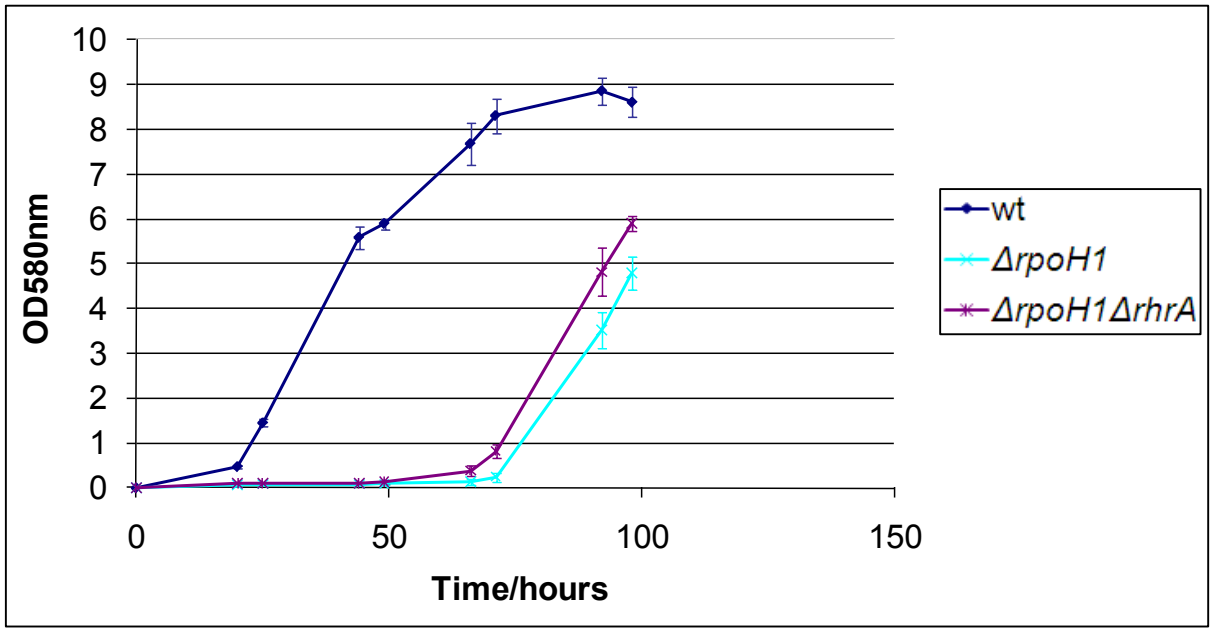


Figure 15. Growth cultures of the *S. meliloti* wild type strain 1021 (wt), the *rpoH1* mutant (Δ*rpoH1*) and the *rpoH1 rhrA* double mutant (Δ*rpoH1* Δ*rhrA*). *(legend continues)*

(legend continued) Strains were grown in Vincent minimal medium and measured for optical density at 580 nm at different time points, for five days. The error bars indicate the standard deviation calculated from three independent cultures.

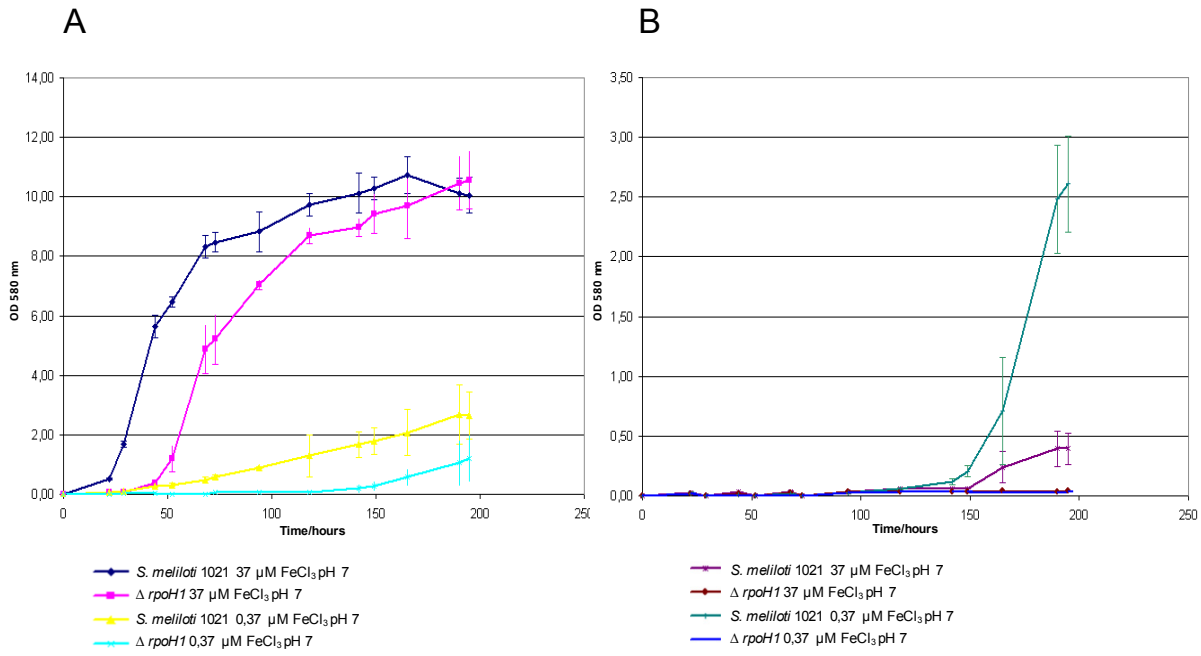
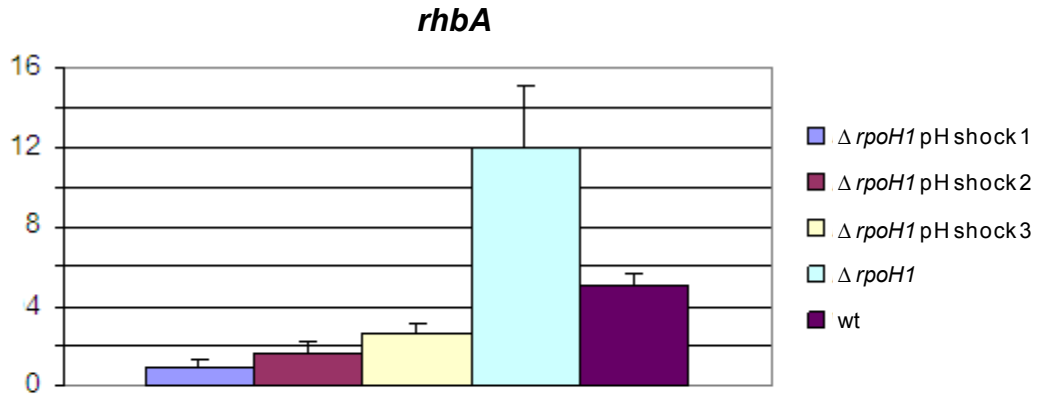


Figure 16. Growth cultures of the *S. meliloti* wild-type strain 1021 and the sigma factor *rpoH1* mutant at different pH values and iron concentrations. Strains were grown in medium at pH 7.0 (A) or pH 5.75 (B), with either 0,37µM or 37µM $FeCl_3$. Samples were measured for optical density at 580 nm at different time points, for seven days. The error bars indicate the standard deviation calculated from three independent cultures.

The results obtained in the real time RT-PCR experiment showed that the expression of both the *rhbA* and the *rhrA* genes in *rpoH1* mutant cells after pH shock is quite lower than in cells at pH 7. This result was observed in three independent *rpoH1* mutant cultures (Figure 17) and it suggests that the rhizobactin biosynthesis operon is unlikely to play a role in pH stress response.

A



B

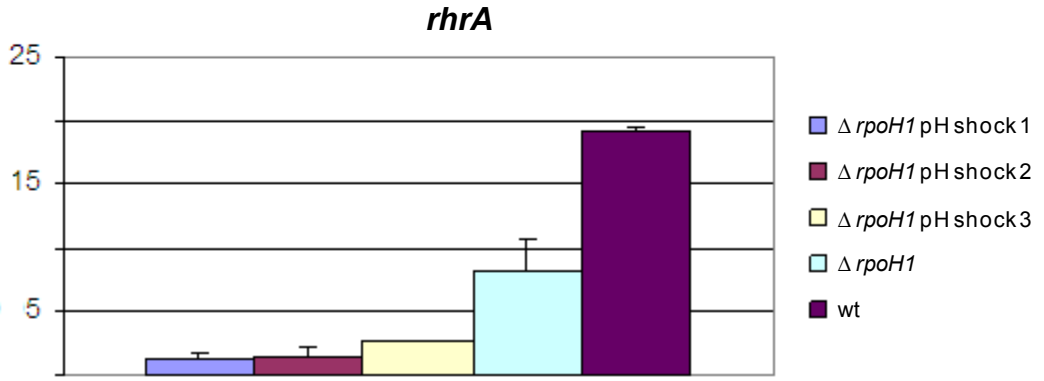


Figure 17. Quantitative real time RT-PCR for detection of rhizobactin expression levels after pH shock. Quantitative real time RT-PCR of the rhizobactin biosynthesis operon gene *rhbA* (A) and the rhizobactin regulator gene *rhrA* (B). The expression levels were measured in 3 biological controls for *rpoH1* mutant cells after 10 minutes of pH shock at pH 5.75, as well as for Δ*rpoH1* and wild type *S. meliloti* cells grown at pH 7. The housekeeping gene *Smc02641*, which codes for a glucose dehydrogenase, was used for normalization.

2.3.3 Possible extracytoplasmic sigma factor involvement in the regulation of the rhizobactin biosynthesis operon

In the absence of RpoH1, it is possible that the the first gene of the rhizobactin biosynthesis operon (*rhbA*) is transcribed by RNAP containing another sigma factor in a manner that is not observed in the wild type, either because a secondary more complex regulation of the rhizobactin operon is missing in the *rpoH1* mutant and/or because RpoH1 is responsible for the transcription of the rhizobactin operon itself. In that case, the alternative sigma factor aiding to promote transcription of the *rhb* operon could be for instance one of the RpoE sigma factors. To identify the probable binding sites for sigma factors that could control the expression of the rhizobactin operon genes, 500 nucleotides of the 5' region of *rhbA* were analysed with PatScan [35] for the presence of the consensus binding site for RpoE2 and RpoH1, which are the two consensus binding sites already described in the literature for *S. meliloti* sigma factors [111, 129]. Interestingly, possible binding sites for both RpoE2 and RpoH1 were found at 96 and 90 nucleotides upstream of the *rhbA* start codon, respectively (Figure 18A). This lead to the hypothesis that the upregulation of rhizobactin biosynthesis in the *rpoH1* mutant could be due to the transcription by an RpoE2-RNA polymerase, that would bind to the promoter of the *rhbA* gene in the absence of RpoH1, leading therefore to a stronger expression. To test whether an RpoE2-containing RNA polymerase could be responsible for the expression of the rhizobactin biosynthesis operon in the *rpoH1* mutant, an *rpoH1 rpoE2* double mutant was produced. Double mutants for *rpoH1* and the *rpoE3* and *rpoE5* genes were also generated, for the similarities in the sequences of these extracytoplasmic sigma factors could also imply similarities in their consensus binding sites. Real time RT-PCR was performed for the *rpoH 1rpoE2*, *rpoH1 rpoE3* and *rpoH1 rpoE5* double mutants. The cells were grown in Vincent minimal medium until an optical density of 0.8 was reached and total RNA was extracted for the analysis. The result of the real time RT-PCR can be seen in figure 18B. The expression of the rhizobactin operon gene *rhbA* was not reduced in the *rpoH1 rpoE2* double mutant. Actually, only for the *rpoH1 rpoE5* double mutant were the *rhbA* expression levels reduced in comparison to the *rpoH1* mutant. This indicates that the RpoE5 sigma factor could play a role in the upregulation of the rhizobactin biosynthesis operon, in the absence of RpoH1.

A

rhbA

gatctattccttctttatggcgtcctcgtccgacgaccagccgggaactgacttcaccatcctgagctgcgccga
gttggtcgtctatctgatcggcacctcgatcgcaggctacgttgccgaccggttcggctatgcgacgctcttctc
ctcggcgacggtcatttccgtgctcggcatcggattgtcagtctggatgctggaacggttgaaagcgcgtccttc
aagatcacgatgatttcaggtcgatcgactgaaaatcgcagaagcagagtgattgtaaagagttagaccgagATG

RpoH consensus cncttgaa – ccanat

RpoE2 consensus ggaacna – gcgttt – g

B

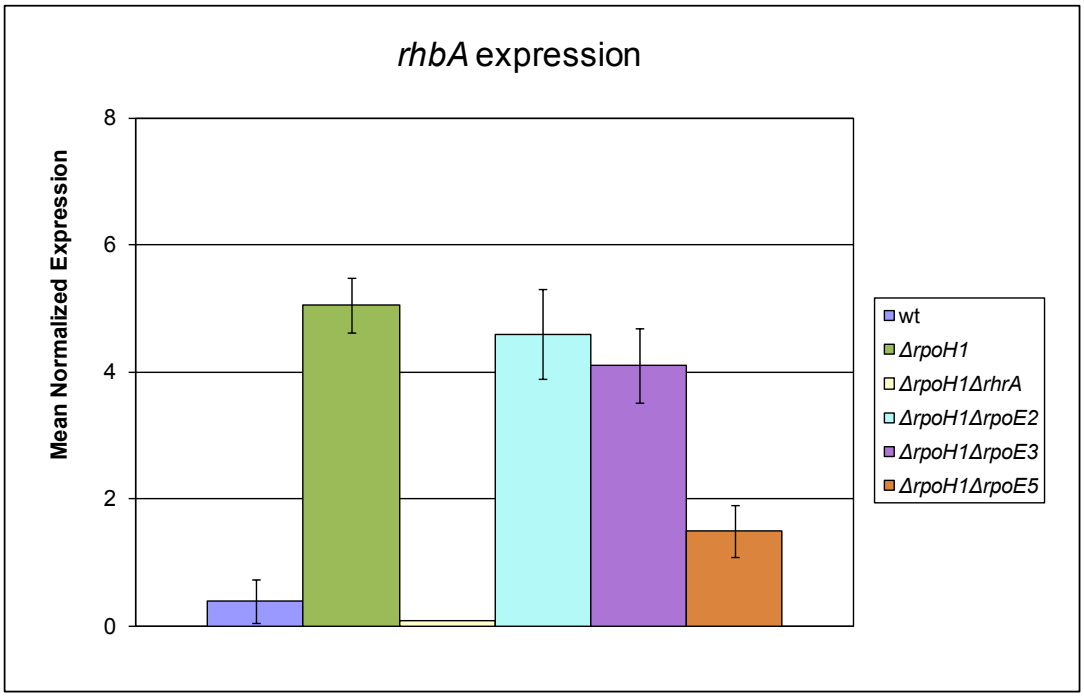


Figure 18. Indications for alternative sigma E regulation of the rhizobactin operon. A) Possible binding sites for RpoH1 (green) and RpoE2 (blue) were found in the putative promoter region of the *rhbA* gene, 96 and 90 nucleotides upstream the *rhbA* start codon, respectively. B) Reduction in rhbA expression levels measured by real time RT-PCR was observed for the *rpoH1 rpoE5* double mutant.

3. SIGMA FACTOR RPOH1 IS A REGULATOR OF PH STRESS RESPONSE IN *S. MELILOTI*

3.1 RPOH1 IS ESSENTIAL FOR GROWTH AT ACIDIC PH

To verify if the complementation of the *rpoH1* mutant phenotype at low pH could be achieved, a growth test was performed with *rpoH1* mutant cells bearing a plasmid that contains the *rpoH1* gene. For this purpose, the *rpoH1* gene was cloned into the plasmid pJN105. The recombinant vectors were first introduced into the *E. coli* strain S17-1 and then conjugated into *S. meliloti rpoH1* mutant cells. Restoration of the wild type growth phenotype was observed for the *rpoH1* mutant carrying the recombinant plasmid with the *rpoH1* gene, confirming therefore that the lack of growth at acidic pH values was indeed caused solely by the *rpoH1* mutation (Figure 19). The results indicate that the RpoH1 sigma factor is essential for growth at acidic pH.

3.2 CHARACTERIZATION OF *RPOH1* MUTANT AT DIFFERENT ACIDIC PH VALUES AND IN RESPONSE TO PH SHOCK

The *rpoH1* deletion mutant did not grow at all at pH 5.75. In order to better investigate regulation by RpoH1 at low pH, tests were done with the intent of finding the minimum pH value that permitted growth for this mutant. Growth tests were performed in Vincent minimal medium at pH values ranging from 5.7 to 7.2. The *rpoH1* mutant showed to be extremely sensitive to low pH, for the pH value 6.7 seems to be the minimum pH value that allows for *rpoH1* mutant cells to grow (Figure 20).

Since the *rpoH1* mutant is unable to grow at acidic pH, the RpoH1-dependent gene expression was investigated with a pH shift experiment. To this end, a growth test was performed in which *S. meliloti* wild type and *rpoH1* mutant were transferred from a medium at pH 7 to a medium at pH 5.75 after reaching an optical density of 0.8 at 580 nm (Figure 21). This test was useful to determine if the *rpoH1* mutant growth impairment was extended to sudden

acidic pH shift and also to test further for a role for *rpoH1* in pH shock response. *S. meliloti* wild type strain 1021 and the *rpoH1* mutant were grown under identical conditions at pH 7.0 until an optical density of 0.8 at 580 nanometers was reached. The cultures were then centrifuged and resuspended in fresh medium either at pH 5.75 or at pH 7.0 (control) (Figure 21). The samples continued to be measured for optical density after pH shock. In those experiments it was also observed that the *rpoH1* mutant displays a growth defect after pH shock in comparison to the wild type, suggesting once more the participation of the RpoH1 sigma factor in fighting pH stress (Figure 22). The viability of the mutant cells was tested 30 minutes after pH shock by observing their ability to form colonies in TY plates incubated at 30˚C overnight. The results indicated that the transfer to medium at acidic pH is not lethal to the *rpoH1* mutant.

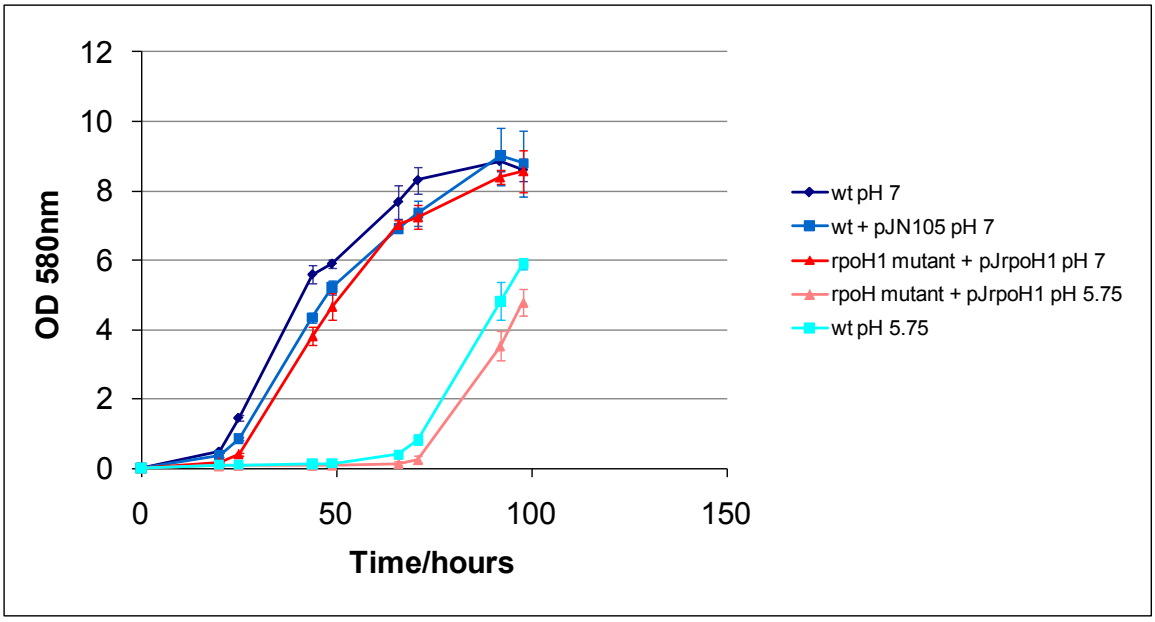


Figure 19. Complementation of the *rpoH1* mutation. Growth cultures of *S. meliloti* wild type strain and the *rpoH1* mutant carrying the plasmid pJ*rpoH1* containing the *rpoH1* gene were performed at both pH 7 and pH 5.75. Strains were grown in Vincent minimal medium and measured for optical density at 580 nm at different time points, for five days. The error bars indicate the standard deviation calculated from three independent cultures.

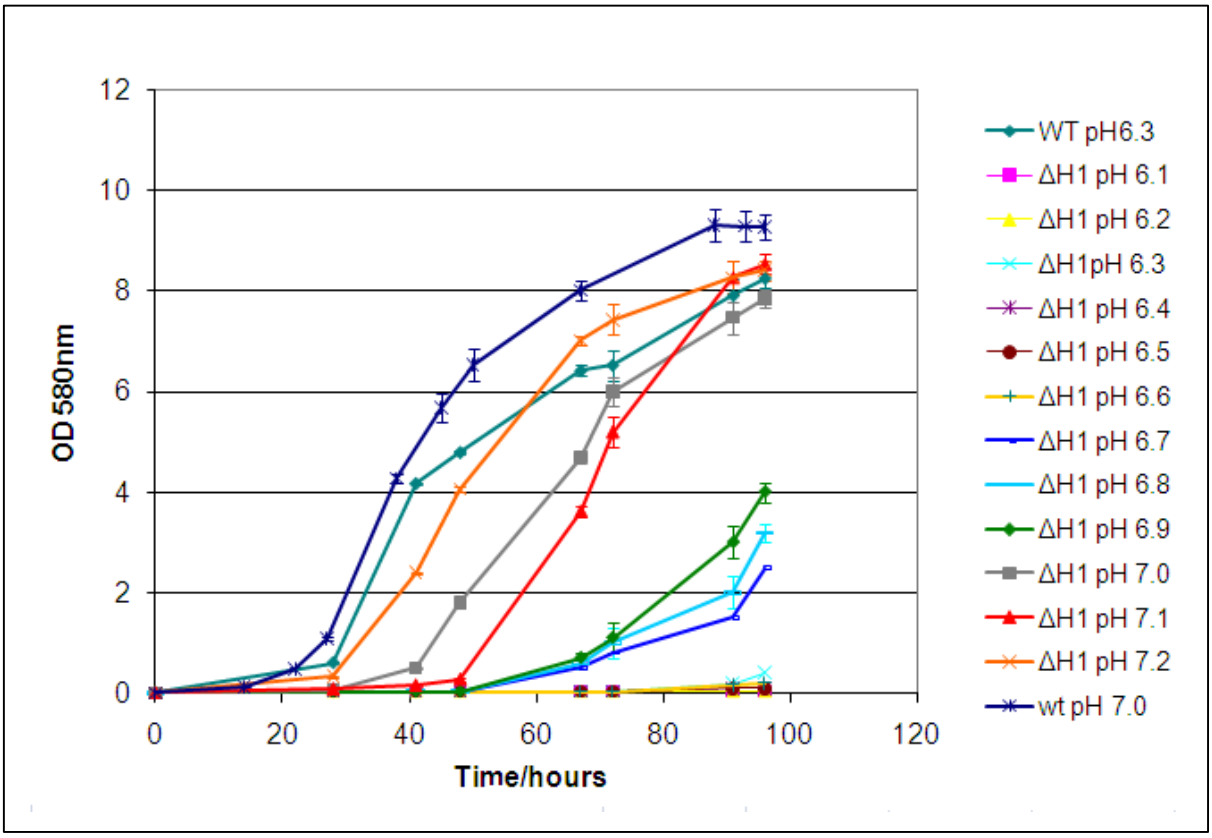


Figure 20. Growth cultures of the *S. meliloti rpoH1* deletion mutant at pH values ranging from 6.1 to 7.2. Samples were measured for optical density at 580 nm at different time points. No growth was observed below pH 6.7. Error bars indicate the standard deviation calculated from three independent cultures.

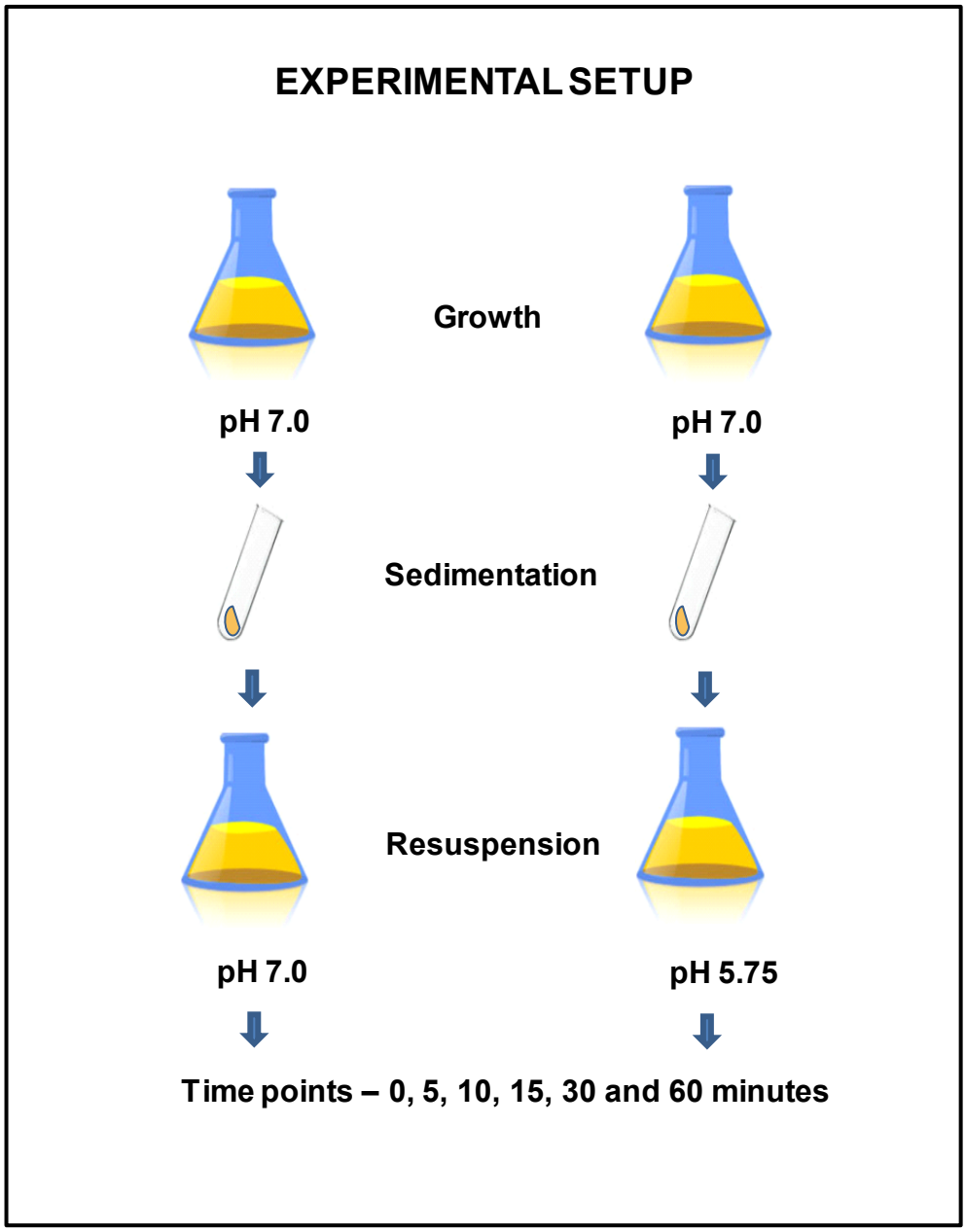


Figure 21. Experimental setup for pH shock experiment. Cells were grown in medium at pH 7 until reaching an optical density of 0.8 at 580 nm. The samples were then centrifuged and resuspended in either medium at pH 7 or pH 5.75. Cells were harvested at time points 0, 5, 10, 15, 20 and 60 minutes after pH shock.

3.3 GLOBAL TRANSCRIPTION PROFILING OF *S. MELILOTI* 1021 FOLLOWING PH SHOCK

3.3.1 Experimental setup and microarray analyses with *S. meliloti* wild type

Considering the striking phenotype for the *rpoH1* mutant at pH stress conditions, the identification of which genes involved in pH stress response of *S. meliloti* are under the control of RpoH1 was of high importance. Time-course global transcription profiling of *S. meliloti* wild type and *rpoH1* mutant were used as the main approach to fulfil that purpose. The first step was to identify *S. meliloti* wild type genes involved in pH stress, for later comparison to the *rpoH1* mutant. *S. meliloti* 1021 cells were grown in medium at pH 7.0 until reaching an optical density of 0.8 at 580 nm, and then transferred to medium at pH 5.75 or pH 7.0 (control). Samples were harvested at time points 0, 5, 10, 15, 30 and 60 minutes after the transfer (Figure 21). For each point of time, microarray hybridization analyses were performed comparing the cells shocked at pH 5.75 with control cells again transferred to medium at pH 7.0. The microarray images were analysed with the utilization of Imagene Software and EMMA [34, 36]. To determine differential gene expression, the $\log_2$ transforms of normalized model-based expression values of genes were compared. Log_2 ratio or fold change of gene expression was obtained for each gene at each time point against the time-matched control and the normalized model-based expression values of genes were compared. In order to identify genes that play a role in the cellular response to acidic pH, significant change in expression was determined in combination with a cut-off value of approximately threefold change. Only genes that showed a significant increase or decrease in the expression ratio of about threefold (M-value ≥ 1.4 or ≤ -1.4) between the two pH classes, for at least one of the six time points, were considered.

In the microarrays performed with *S. meliloti* wild type, a total of 210 nonredundant genes were selected whose expression was altered significantly at one or more time points. Those genes are listed in appendix 1. Overall, the observed response of the *S. meliloti* wild type following acid shift was in agreement with that described by Hellweg *et al*. [61]. Most transcriptional changes occurred within 20 minutes after pH shift and upregulation was dominant over downregulation at all time points. The response to acidic pH stress was characterized by the expression of gene sets associated with different cellular functions. Among the strongest upregulated genes (M-value ≥ 1.8) were *lpiA*, which codes for a low pH induced protein, the DegP1 protease encoding gene and *cah*, coding for a carbonic anhydrase. Among the groups of genes responding to the shift to acidic pH were those coding for chaperone proteins, which were upregulated, and genes involved in nitrogen uptake and

metabolism. There were also genes coding for exopolysaccharide I biosynthesis (EPSI) as well as flagellar and chemotaxis genes [12, 131]. While the EPSI genes were upregulated, the expression level of flagellar genes decreased in response to acidic pH.

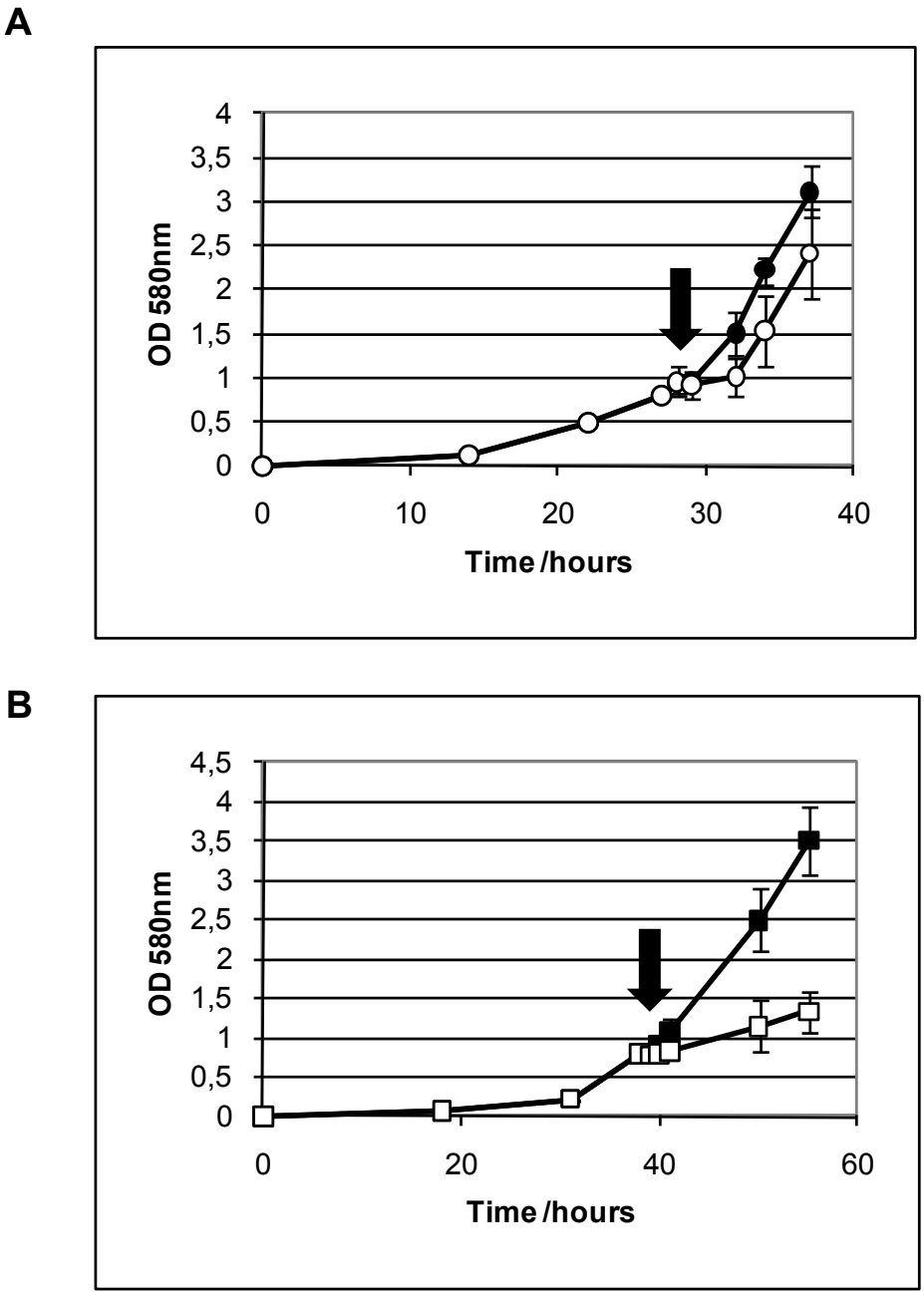


Figure 22. Growth curves of *S. meliloti* 1021 wild type strain and sigma factor *rpoH1* mutant after pH shock. *S. meliloti* 1021 wild type strain (A) and sigma factor *rpoH1* mutant (B) were grown in medium at pH 7.0 and transferred to medium at pH 5.75 (open signs) or at pH 7.0 (filled signs). The arrows indicate the moment of pH shift. Cell growth was measured every two hours after pH shift. The error bars indicate the standard deviation calculated from three independent cultures.

3.3.2 Time-course microarray data of *S. meliloti* wild type following an acidic pH shift were grouped in 6 K-means clusters

In order to identify patterns of gene expression from the data and to characterize the intricate dynamics of differential expressions from a temporal viewpoint, a clustering of genes according to their time-course transcription profiles was carried out. For the *S. meliloti* wild type, only those genes with a significantly altered expression (M-value $\geq$ 1.4 or $\leq$ -1.4) after pH shock were analysed and clustering of the time-course data ($\log_2$ ratio of gene expression) was performed using the Genesis software [138], which is suited for analysis of short time-series microarray data. The K-means clustering method is based on the mean values of similar expression data and it can be implemented to define a set of distinct and representative models of expression profiles. With K-means, each gene groups into the model profile to which its time series most closely matches, based on its Euclidian distance to the profiles. Clustering analysis was performed on the 210 genes that displayed significant differential expression at one or more time points in the wild type arrays. Genes with similar expression characteristics were grouped in the same cluster by the program. A total of 6 clusters were generated for the wild type microarray data, with distinct expression patterns over the 60-minute time-course. Clusters A to C represent the genes whose expression was upregulated and clusters D to F represent the genes whose expression was downregulated throughout the time that followed pH shift (Figure 23). Operons and genes involved in similar cellular functions were predominantly grouped in the same clusters. In the following, expression characteristics and relevant genes from the six wild type clusters will be further described.

<u>Cluster A</u>

Cluster A grouped genes with the strongest transcriptional induction after shift to low pH. It consists of 28 genes, including *nex18*, involved in the response to nutrient deprivation stress [28] and *lpiA*, involved in the formation of lysyl-phosphatidylglycerol, which is a low pH induced protein in *S. medicae* [118]. The exopolysaccharide biosynthesis (EPS) genes *exoV*,

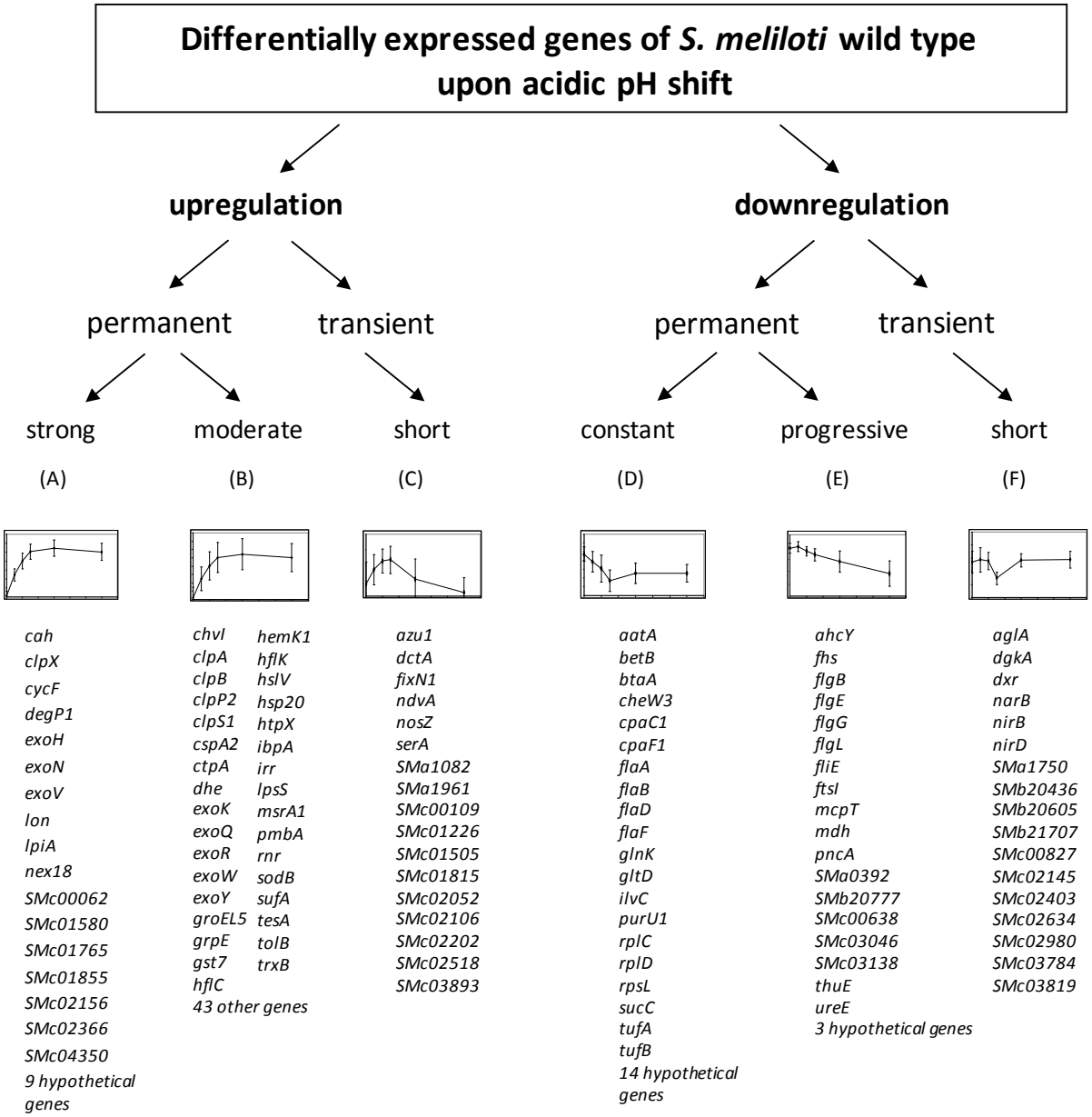


Figure 23. K-means clustering of *S. meliloti* 1021 genes differentially expressed after pH shift time-course. Six clusters (A-F) were characterized by their specific transcriptomic profiling over 60 minutes following acidic pH shift. Graphics illustrate the expression profile based on the mean values; the X-axis represents time, whereas the Y-axis represents the $\log_2$ ratio of gene expression. Tables below each graphic enlist genes distributed to the corresponding cluster.

exoH, *exoN*, and the gene for the Lon protease, a regulator of exopolysaccharide synthesis [139], also grouped in this cluster. The *lon* mutant in *S. meliloti* synthesizes EPSI, and also constitutively synthesizes EPSII, a galactoglucan which is the second major EPS known to be produced by *S. meliloti*, but typically is expressed only under conditions of phosphate limitation [139].

Cluster B

This cluster comprises genes that were gradually upregulated during the time-course and reached their maximum expression values at approximately 20 minutes after pH shift. Though not all, it includes most of the genes involved in EPS I biosynthesis. The upregulation of EPS biosynthesis genes upon sudden pH shift probably accounts for the mucoid phenotype in *S. meliloti* cells grown on plates at low pH and is in accordance to what has already been reported by Hellweg et al. [61]. Moreover, this cluster also includes a broad range of genes coding for heat shock proteins and chaperones involved in stress response, such as *ibpA*, *grpE*, *hslVU* and *groEL5* and the genes coding for the proteases HflCK, HtpX, FtsH, ClpAB, ClpP1 and ClpS.

Cluster C

Cluster C is composed of genes which were transiently induced after pH shift. It contains the dicarboxylate transport system DctA, which is essential for symbiosis in *S. meliloti* [158]. Also, the gene *rsiA1*, which plays the function of the anti-sigma factor for the extracytoplasmic function sigma factor RpoE2 [129], was transiently upregulated (Figure 23). RpoE2 activity is negatively controlled by two paralogous anti-sigma factors, RsiA1 (SMc01505) and RsiA2 (SMc04884), and RpoE2 activation by stress requires those two redundant paralogous. Activity of RsiA1, however, is negatively regulated by the anti-anti sigma factors RsiB1 and RsiB2 [10].

Cluster D

The genes in cluster D were gradually downregulated up to 30 minutes after pH shift, and maintained the peak of downregulation at 60 min. It comprises a number of genes related to

flagella biosynthesis and pillus assembly. It also includes the gene coding for GlnK, which is involved in generalized nitrogen stress response [159].

Cluster E

Cluster E is composed of genes whose expression decreased continuously for the whole duration of the time-series experiment. The expression was gradually downregulated as of 5 minutes after pH shift, followed by greater downregulation up to 60 min. Among the genes in this cluster were the flagellar genes *flgG flgL*, *flgB* and *fliE*. Whereas most flagellin genes clustered in cluster D, most of the *flg* and *fli* genes were grouped in cluster E.

Cluster F

Cluster F consists of genes which were transiently downregulated in their expression level after pH shift. It involves genes that play a role in nitrate assimilation, such as *nirB*, *nirD* and *narB* and the nitrate transporter *smb20436* (Figure 23). In *S. medicae*, *narB* expression is regulated by ActR, which is part of a two-component signal transduction system and is required for growth at low pH and microaerobic induction of the nitrogen fixation regulators [37].

3.4 Global transcription profiling analysis of *S. meliloti rpoH1* mutant cells following pH shock

In order to characterize the regulation of the *S. meliloti* response to pH stress, time-course microarray analyses were also performed to examine the differential expression of genes in *rpoH1* mutant cells following sudden pH shock. Cells were grown in medium at pH 7 until reaching an optical density of 0.8 at 580 nm, and then transferred to medium at pH 5.75. Then they were harvested at time points 0, 5, 10, 15, 30 and 60 minutes after pH shift. For each point of time, the microarray hybridization analyses were performed between *rpoH1* mutant cells shocked at pH 5.75 and control *rpoH1* mutant cells grown at and transferred to medium at pH 7. The experimental setup for the procedure with the *rpoH1* mutant was identical to that

of the wild type, allowing therefore for significant data comparison. Such adequate comparison is essential for modeling the differences in transcription between wild type and *rpoH1* mutant throughout the time-course that enable us accordingly to elucidate the role of RpoH1 during pH stress.

In the microarray analyses performed with the *S. meliloti rpoH1* mutant following acid shift, 132 of the 6,208 genes showed significant time-dependent variation in expression in at least one of the six time points. Those genes exhibited approximately threefold change in at least one time point throughout the 60 minute time-course. Among the genes most strongly induced were *nex18*, a gene that codes for a nutrient deprivation activated protein and the gene for the low pH induced protein *lpiA*. Both of these acid-induced genes display an extracellular stress response function [119]. Similarly to the wild type arrays, several genes of the flagellar regulon were repressed at low pH, whereas the genes of the EPS I biosynthesis were upregulated. Interestingly, in contrast to the *S. meliloti* wild type, some genes coding for nitrogen uptake and metabolism and several genes coding chaperone proteins, were not observed among the differentially expressed genes in the *rpoH1* mutant arrays. This was a strong indication that some genes for pH stress response in *S. meliloti* are under the regulation of the *rpoH1* sigma factor.

3.5 COMPARISON OF *RPOH1* MUTANT ARRAYS TO WILD TYPE TIME-COURSE GLOBAL TRANSCRIPTION PROFILES

3.5.1 K-means clustering performed with the expression data obtained from the *rpoH1*mutant microarray analyses

To elucidate the role of RpoH1 in transcription dynamics during pH stress response, the time-course transcriptomic analyses of the *rpoH1* mutant upon acidic pH shift were compared to those of the wild type. For a most effective comparative analysis, K-means clustering was performed for the 210 genes selected through the filtering of the wild type data, but this time the same type of clustering was carried out with the $\log_2$ expression data from the *rpoH1*

mutant arrays. This approach was applied because it enabled the identification of genes that, throughout the time-course, behaved in a similar fashion both in the *rpoH1* mutant arrays and in the wild type, and at the same time it also allowed for the identification of genes that displayed no differential expression in the *rpoH1* mutant arrays, even though they were differentially expressed, upon acidic pH shift, in the wild type. The dynamic gene expression profiles were also catalogued into six clusters for the *rpoH1* mutant, separating groups of genes with the highest possible similarity (Figure 24). Additionally, a program was created for plotting the profiles of each individual gene in the microarrays, both for the wild type and for the mutant in one same graphic (Appendix 8).

Clusters G, H and I

Clusters G and H comprise genes that were constantly upregulated over time, either with a very strong induction (M-value ≥ 2.5 for at least one time point) or a moderate one (M-value ≤ 2.5) (Figure 24). Among the strongly upregulated genes in cluster G were *nex18* and *lpiA*, the EPS biosynthesis genes *exoV*, *exoH*, *exoN* and the gene coding for the Cah carbonic anhydrase, which is also induced in response to phosphate starvation of *S. meliloti* [77]. Genes grouped in cluster H include many *exo* genes and the gene coding for a regulator of succynoglycan production *chvI* [150], as well as the gene encoding one of the homologs for the translocation protein TolB (Figure 24). A few transiently upregulated genes were listed in cluster I, such as the gene coding for SerA dehydrogenase, a transcriptional regulator that belongs to the ArsR family and the gene *azu1*, coding for pseudoazurin. The *azu* gene is not essential for nitrogen fixation in *R. leguminosarum* [114].

Clusters J and K

Those two clusters grouped genes that were downregulated throughout the time-course, with persistent and transient downregulation, respectively. Like in the wild type arrays, many flagellar genes were also downregulated in the mutant and grouped in cluster J. The phosphate transport system encoded in the *phoCDET* operon also grouped in this cluster. In *E. coli*, *phoB* is involved in the acid shock response [140]. Among the transiently downregulated genes in cluster K were genes involved in nitrogen metabolism, such as those coding for nitrite and nitrate reductases, *nirD*, *nirB* and *narB*, which play a role in the conversion of nitrate to ammonia.

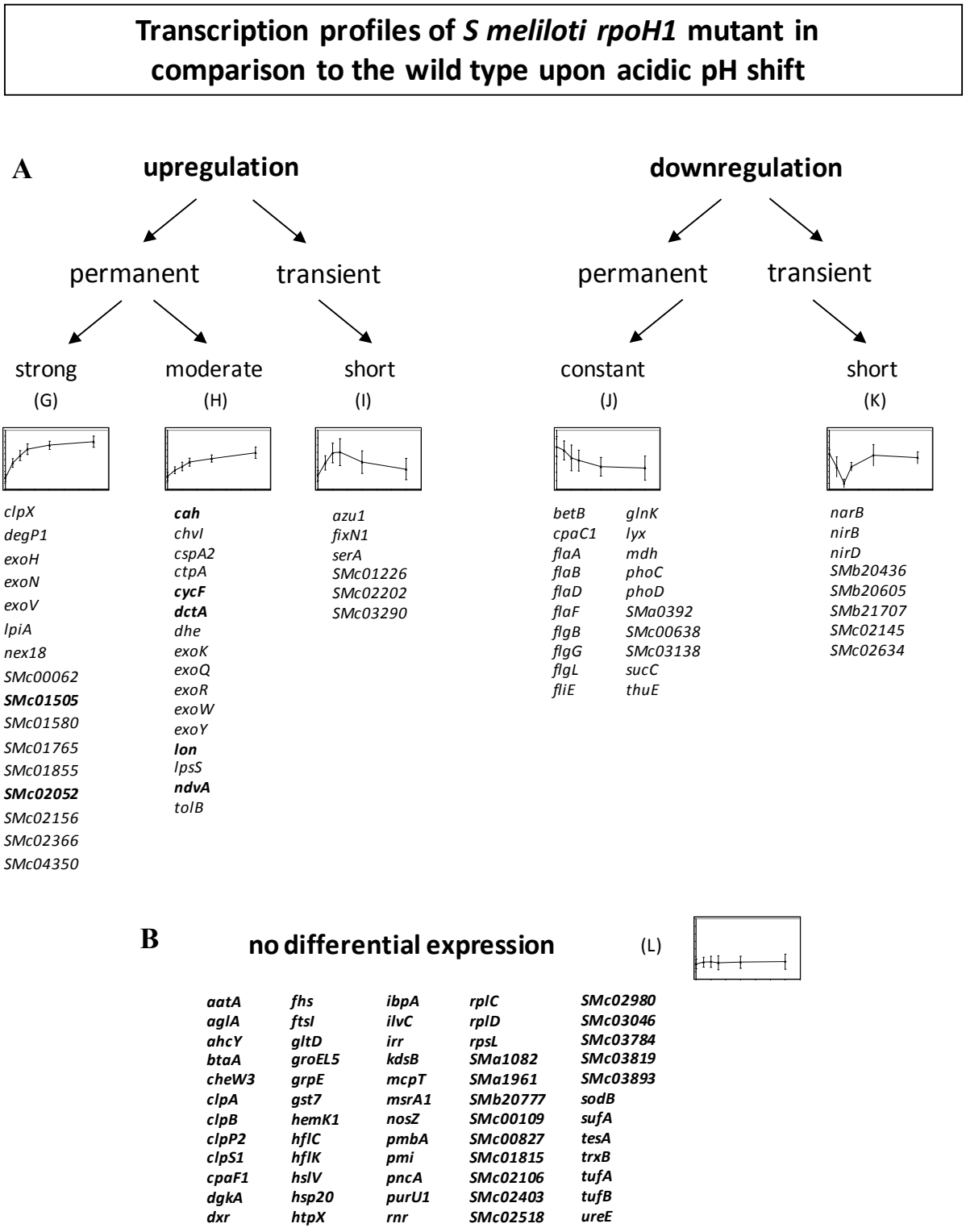


Figure 24. **Expression profiles of *S. meliloti rpoH1* mutant genes upon pH shift in comparison to wild type.** Uniquely classified groups (G-L) were obtained through K-means clustering of *rpoH1* mutant microarray data. Graphics illustrate the *(legend continues)*

expression profile based on the mean values; the X-axis represents time, whereas the Y-axis represents the $\log_2$ ratio of gene expression. Genes marked in bold present dissimilar expression profile in comparison to *S. meliloti* wild type clustering results.

Cluster L

Unlike the wild type, the clustering of the *rpoH1* mutant data yielded the observation of a large cluster of genes whose expression changed very little throughout the time-course. For the genes in cluster L, the M-values remained close to zero at all time points (Figure 24). Genes in cluster L include those coding for heat shock proteins and proteases, as well as the elongation factor *tufAB* operon and the gene coding for the putative chemotaxis protein *cheW3*. The complete lists of genes obtained from the clustering of the *rpoH1* mutant data can be seen in the appendix section 2.

3.5.2 Identification of *S. meliloti* genes that are regulated in an RpoH1-independent, RpoH1-dependent and complex manner following an acidic pH shift

Based on the cluster comparison between wild type and *rpoH1* mutant, the results allow for the dynamic distribution of the genes in three different classes: genes whose expression at low pH is independent of *rpoH1* expression, genes that display an expression dependent on *rpoH1* after pH shift, and genes that present a complex behaviour following pH shift. RpoH1-independent genes were designated as those distributed into similar expression profiles in both wild type and *rpoH1* mutant clustering analyses, that is, genes that were similarly up- or downregulated in both mutant and wild type arrays. Most genes from wild type cluster A (Figure 23) presented an RpoH1-independent expression, as they were also upregulated in the *rpoH1* mutant arrays and grouped at cluster G (Figure 24) in the *rpoH1* mutant clustering analysis, which was also clustering upregulated genes. The gene coding for the low pH induced protein LpiA also presented RpoH1-independent upregulation in the pH shift arrays, as did the EPS I biosynthesis genes *exoQ*, *exoW*, *exoV*, *exoH*, *exoK exoR*, *exoN,* and *exoY*, (Figure 25). The *exo* genes were also upregulated in both *rpoH1* mutant and wild type arrays. Similar expression profiles could also be observed for the genes coding for the carbonic anhydrase Cah and the cytochrome CycF protein. Almost all genes involved in motility and

flagellar biosynthesis, like the flagellar genes *flgB*, *fliE*, *flgG* and *flgL*, displayed similar expression profiles in both wild type and mutant arrays. The behavior of flagellar genes in the microarrays leads to the characterization of an RpoH1-independent downregulation of motility genes upon acid pH shift in *S. meliloti*. Flagellar genes *flaA*, *flaB*, *flaD*, and *flaF* were also downregulated in the mutant, showing therefore that the absence of RpoH1 probably did not interfere with the reduction of cell motility at low pH. Examples of RpoH1-independent expression, for genes *exoY* and *flgL* are shown in figure 25 A and B.

Genes classified as RpoH1-dependent did not present significant differential expression after pH shift in the *rpoH1* mutant arrays. On the other hand, these genes had displayed a circa threefold differential expression for at least one time point in the wild type arrays. They comprise as many as 101 genes of the *S. meliloti* genome whose transcription after pH shift seems to be dependent on *rpoH1* expression. A significant number of protein turnover and chaperone genes were upregulated in the wild type arrays, such as the ones coding for the heat shock proteins IbpA, GrpE and GroEL5 (Figure 25) as the ones coding for the Clp proteases, which are involved in the degradation of misfolded proteins [97]. However, absolutely no differential expression was observed for those genes in the *rpoH1* mutant arrays, characterizing thus an RpoH1-dependent expression of stress-response genes upon acid pH shift (Figure 25). Genes involved in translation, like *tufA* and *tufB, rplC rplD* and *rplS*, were downregulated only in the wild type arrays, characterizing a seemingly RpoH1-dependent inhibition of translational activity in *S. meliloti* cells under pH stress. Genes *cheW*3 and *mcpT* (Figure 25), coding for proteins involved in chemotaxis, were also downregulated only in the wild type arrays.

RpoH1 is also involved in the downregulation of specific transiently expressed genes. Interestingly, a few genes from wild type cluster C (Figure 23) were not grouped in cluster I (Figure 24) as transiently upregulated in the *rpoH1* mutant arrays. Those include the genes *dctA*, coding for a dicarboxylate transport protein, *ndvA*, coding for a beta glucan export protein, and the gene *rsiA*, which codes for an RpoE2 anti-sigma factor. These genes seem to have an RpoH1-independent upregulation, but an RpoH1-dependent downregulation as of 20 minutes following pH shift. In the wild type arrays, the expression of these genes is transient, but in the *rpoH1* mutant arrays these genes remained upregulated throughout the entire time period analysed (Figure 25).

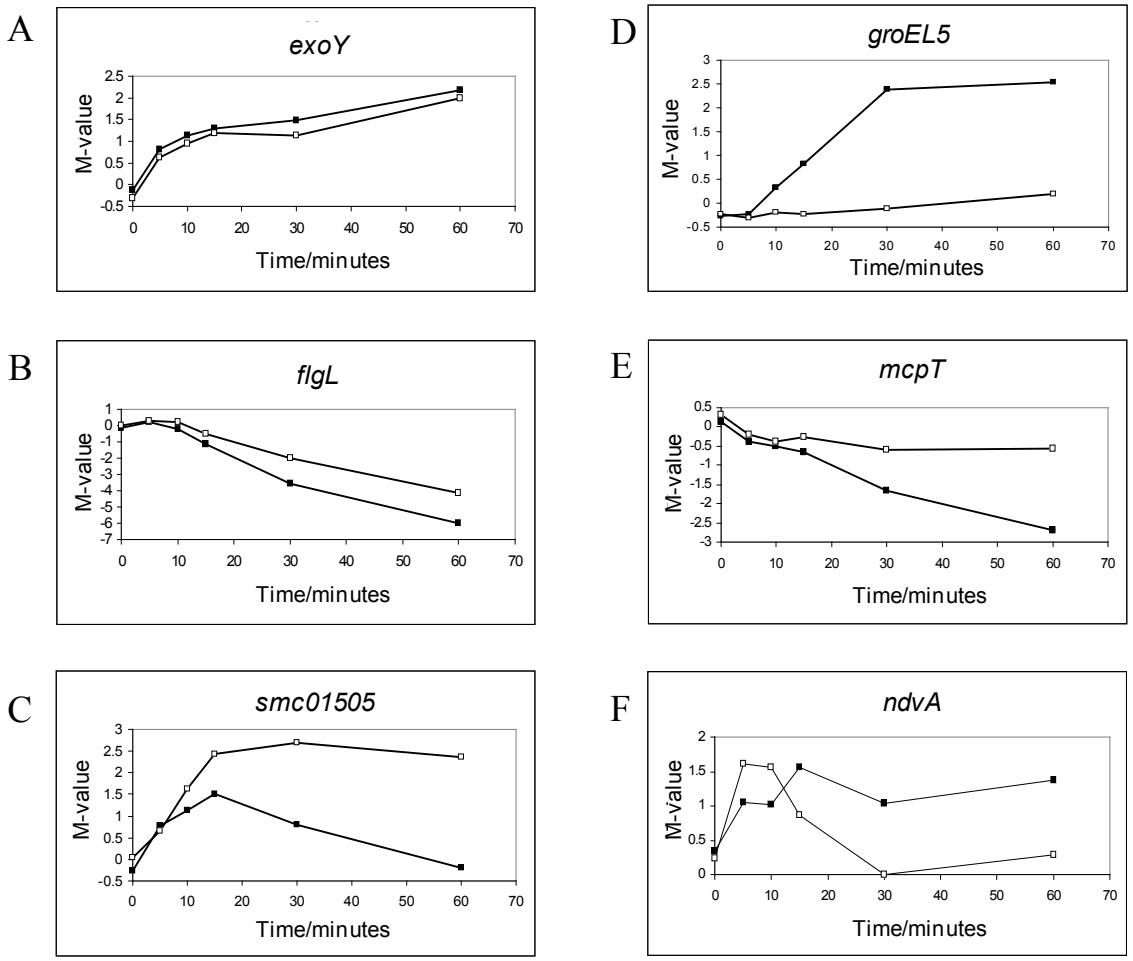


Figure 25. M-values of specific genes throughout the time-course following acidic pH shift in *S. meliloti* 1021 wild type strain (closed squares) and sigma factor *rpoH1* mutant (open squares). Graphics A and B exemplify RpoH1-independent up and downregulation, respectively, whereas graphics D and E show RpoH1-dependently regulated genes. C and F account for complex RpoH1-dependent downregulation in the later time points following acidic shift.

For those genes, the RpoH1 sigma factor is probably involved in the regulation, but is more likely to play a role in a more complex regulatory system that could include a secondary regulator, like for instance a repressor. The repressor could be under the control of RpoH1, whereas the transcription activation of the genes themselves seems to be RpoH1-independent.

3.5.3 Functional classification of genes regulated in an RpoH1-dependent manner

The genes that had distinct expression profiles in the *rpoH1* mutant arrays in comparison to the wild type, that is, the genes that presented an RpoH1-dependent differential expression, were also classified according to their COG categories. The COG classification distributes genes in orthologous groups on basis of functional predictions and patterns of sequence similarities [141]. The RpoH1-dependent genes were assigned to 18 functional categories. Among the known most representative classes were protein turnover and chaperones, followed by translation, transcription and by transport and metabolism of carbohydrates, nucleotides and amino acids (Figure 26). There is really a remarkable increase in the expression of chaperone proteins and heat shock genes in response to pH shock.

A total of 24 genes that presented an RpoH1-dependent upregulation following acid shift are known to be involved in heat shock and stress response. Among the proteases, the genes coding for HtpX, a membrane-bound and stress-controlled protease well characterized in *E. coli* [125], as well as those coding for ClpB and ClpP2, responsible for disassembling protein aggregates that accumulate in the cytoplasm under stress conditions [97], were expressed in dependence of RpoH1. The operon formed by the genes *hslUV*, which codes for an intrinsic ATP-dependent proteasome system for degradation of misfolded proteins in the cytoplasm, was also upregulated in an RpoH1-dependent fashion. Among the induced chaperones were also the gene *Smc00699*, coding for a heat shock DnaJ-like protein, as well as the gene coding for GrpE, which is part of the cellular chaperone machinery capable of repairing heat-induced protein damage [60]. Moreover, there was an RpoH1-dependent upregulation of the operon that codes for the only GroELS proteins specialized in stress response in *S. meliloti,* GroELS5 [97].

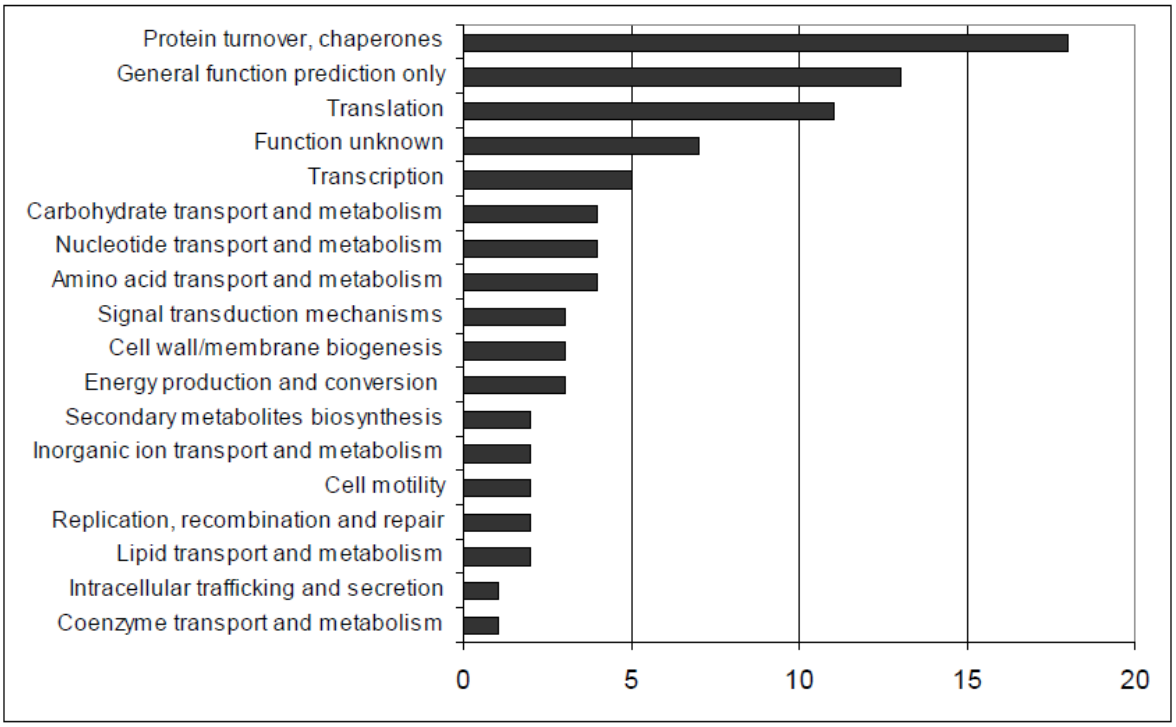


Figure 26. COG classification of RpoH1-dependent genes. COG classification of *S. meliloti* genes that are regulated in an RpoH1-dependent manner after pH shift. The x axis indicates how many genes belong to each category.

The gene coding for the small heat shock protein IbpA [103] was also upregulated. The group of proteins shown to be involved in the heat shock response under the transcriptional control of RpoH usually includes chaperones, proteases, and regulatory factors [76, 82]. The mutation in the *rpoH1* gene in *S. meliloti* and its characterization under pH stress revealed indeed a lack of activation of all major types of regulatory chaperones and key heat shock proteins usually activated in stress conditions. In this study, representatives of all of those groups seem to be involved in pH stress response. This definitely attests to the role of *rpoH1* in *S. meliloti* pH stress response. This role was clearly evidenced by the activation of acid-induced heat shock proteins in dependence of RpoH1 upon pH stress.

3.6 VALIDATION OF MICROARRAY RESULTS BY REAL TIME RT-PCR

The microarray results were confirmed using qRT-PCR analysis to observe the differential transcriptional expression some genes upon pH shock. M-values were obtained by qRT-PCR for six selected genes: *dctA*, *smc01505*, *grpE*, *lpiA, exoY* and *mcpT*. Comparison of expression data was performed between samples transferred to medium at pH 5.75 and control samples transferred to medium at pH 7, at time points 10 or 60 minutes after pH shift (Figure 27). Though the analysis of only two time points was not sufficient for the observation of the transient expression of *dctA,* the results obtained by real time RT-PCR were very similar to those acquired in the microarray.

3.7 BIOINFORMATIC ANALYSIS FOR IDENTIFICATION OF THE RPOH1 CONSENSUS SEQUENCE

Upstream regions of the genes that presented an RpoH1-dependent expression in the arrays were investigated with PatScan [35] for the recognition of the RpoH consensus binding site. The Web logo program was used to create a representation of the sequence, where lowercase letters are less highly conserved (Figure 28). Moreover, it was possible to more accurately define the RpoH1 consensus sequence for RpoH1-regulated genes, for until now this consensus sequence has been based on the promoter regions of only three RpoH-dependent genes [97]. The consensus sequence was located upstream from single ORFs, or from the first ORF in putative operons. This sequence was observed for many chaperone and heat shock genes that were identified for the first time as being under direct regulation of RpoH1. Interestingly, the same consensus sequence was also found upstream the *rpoH1* gene, indicating that this sigma factor is positively autoregulated. This finding strongly suggests that RpoH1 directly controls the transcription of the identified genes by binding the conserved sequences, which probably represent the -35 and -10 elements of their promoters (Figure 28). It is however likely that further RpoH1-regulated genes were not detected in our microarrays, either because their M-values were lower than the cutoff used for filtering the data, or because they are not part of the pH stress response in *S. meliloti*.

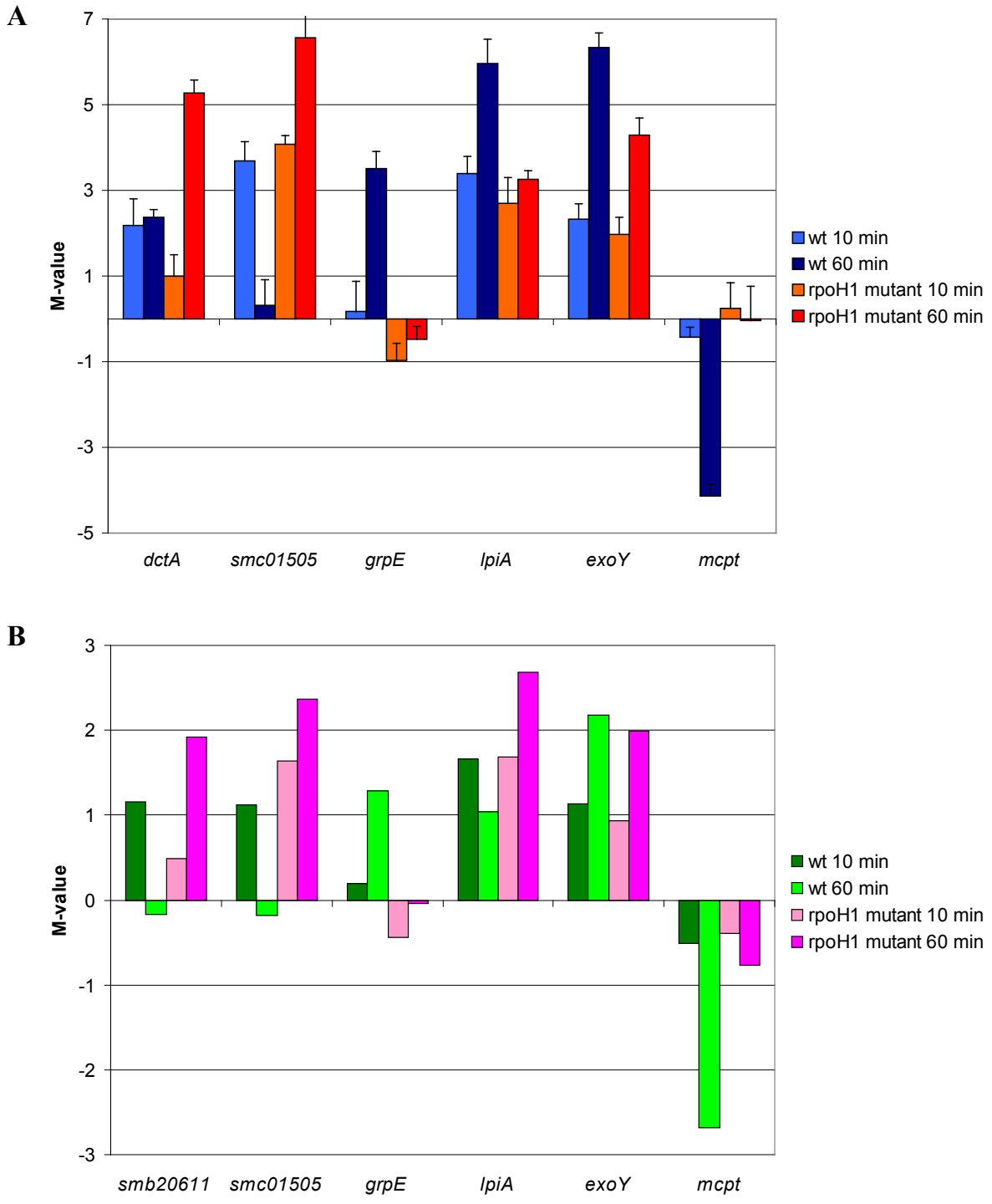


Figure 27. Validation of microarray data by quantitative RT-PCR. A) M-values obtained by qRT-PCR for six selected genes: *dctA*, *smc01505*, *grpE*, *lpiA, exoY* and *mcpT*. Comparison of expression data was performed between samples transferred to medium at pH 5.75 and control samples transferred to medium at pH 7, at 10 or 60 minutes after pH shift. Error bars indicate standard deviation calculated from three independent cultures. B) M-values obtained in the microarray analyses of the six selected genes in (A). Comparison of expression data between samples transferred to medium at pH 5.75 and control samples transferred to medium at pH 7, at time points 10 or 60 minutes after pH shift.

```
ibpA            cattcctCTTGAActcgtgcgcgggcattcCCAtgTcttctccg
ureE            ttcatcgCTTGAccctgacgcatgagcaggCCAtcTgcgccgcg
groEL5          ggggcctCTTGAAatccatttttagcccaaCCAgaTcatcccgc
msrA1           gacatgcCTTGAAaggtctgcggatgccggCCAttgcgcttgcc
clpB            tgccgctCTTGAATtcagaagtgcgctgccCCAtaTttcgcttc
trxB            gcgcgcCTTGAAagtgcgcgccgtgaataCCtaaagtaaggca
groES5          ggggcctCTTGAAatccatttttagcccaaCCAgaTcatcccgc
hslV            gctgcgaCTTGAATaggcctcaagcaa---CCAggctcattcga
clpP            ggcgaacCTTGAAaagtcgcgttcgatttCCAtaTcggaacga
grpE            atcgaccCTTGAATtttcgggc--------CCAaacatcgatat
grx             gcgtcgaCTTGACcgagtatgacg------CCActTatgcacct
glnII           gcggtgcCTTGAgccgaaaaaagagtctctCCAgcTgcccaact
sma1704         tgagagaCTTGAATttc-ttgccgt-cc--CCAagTagacgatg
sma2389         cCCAgtcCTTcAAatagatctcgacg----CCAtcTcgcgtgat
smc00699        gtcaattCTTGATgtggccctctgcaaaccCCAtaTgtgcgtcg
smc03802        tttctccCTTGAAatttcgaagggc-cgcaCCAttTtttggtcc
smc00949        cgcgcctCTTGATgttccatatgcgacaacCCAgaTaactgtct
smb21295        gctcgccCTTcAAgagcggCCActcgttttCCAggaagtcgagg
smc02576        aggtcagCTTGAgcccggtgg--ccgagggCCAcaTgccgatgc

rpoH            ggagccCCTTGAATttgcgtgtgtg-----gCAgaTtccgatct

Consensus               ncCTTGAA   (12-17)        CCAnaT
```

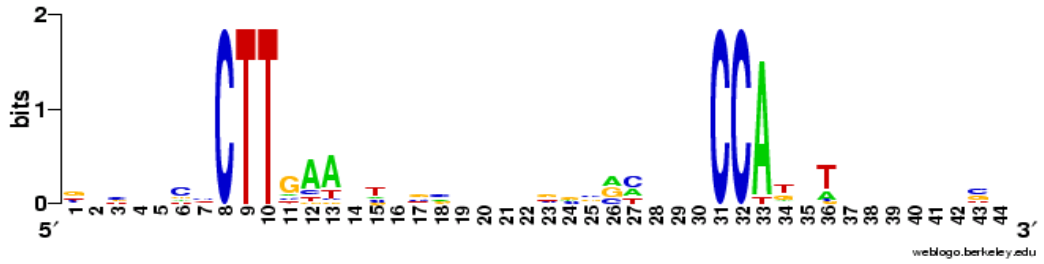


Figure 28. Consensus binding sequence for RpoH1-regulated genes. Genes are represented with the utilization of WebLogo (UC Berkeley). The overall height of each stack of letters indicates the sequence conservation at that position (measured in bits), whereas the height of symbols within the stack reflects the relative frequency of the corresponding nucleic acid at that position.

E. DISCUSSION

1. CHARACTERIZATION OF MULTIPLE ALTERNATIVE SIGMA FACTORS IN *S. MELILOTI* INDICATES OVERLAP IN THEIR FUNCTIONS

In the soil, *S. meliloti* deals with adverse environmental variations that could induce physiological stress responses. Alternative sigma factors play an important role in that they can respond with transcriptional activation to the presence of adverse conditions in their environment. Genomes of rhizobia are usually large (5.4–9.2 Mb) with one to seven replicons [84] and the presence of multiple sigma factors is a ubiquitous feature in those microrganisms. The genome of *S. meliloti* contains nine genes encoding putative ECF sigma factors and two genes encoding RpoH sigma factors [45]. Gene expression patterns are influenced by sigma factor availability and activity. Sigma factors that belong to the same family do not necessarily, but can have overlapping functions and regulate a similar set of genes under different stress conditions. It is also possible that sigma factors of the same family regulate different genes under different stress conditions. This seems to be the case for the different RpoH sigma factors of *B. japonicum* and *R etli*, for instance [87, 107]. Diverse signal transduction pathways could also exist for sigma factor activation, since there are usually different signaling pathways for physical and chemical stresses, as well as for nutrient stress.

The RpoE sigma factors of *S. meliloti* share similarities in their predicted protein sequence and this could account for similar promoter selectivity and possible overlapping functions (Figure 4). ECF sigma factors are responsible for dealing with environmental stress at the periplasmic space and within the bacterial cell wall. Oxidative and osmotic stresses are common effectors of ECF sigma factors in proteobacteria. It is the case for ECF factors in *C. crescentus* and *E. coli* [149]. In *C. crescentus*, the inactivation of an ECF sigma factor called sigma T provokes sensitivities to osmotic and oxidative stresses [4, 5], whereas sigma F is involved in the oxidative response. As for the RpoE sigma factors of *S. meliloti* analysed in this work, expression of sigma factors *rpoE5* and *rpoE1* was upregulated during oxidative stress. Expression of *rpoE2* was not observed for any of the conditions tested, but it has been described that RpoE2 is needed for H_2O_2 stress response in *S. meliloti* [41]. Induction of *rpoE5* expression was also observed under cold shock and heat shock, as well as under pH

stress. A deficiency in growth was observed for sigma factor mutants for *rpoE2* and *rpoE5* in heat shock and oxidative stress conditions. Though in a previous study the *S. meliloti rpoE2* mutant could not be associated with any phenotype [129], deficient growth was observed for the *rpoE2* mutant specially during heat and cold stress conditions. Though RpoE2 acts as a general stress response sigma factor in *S. meliloti* [129], RpoE5 also seems to be an important extracytoplasmic stress response sigma factor, for its expression was induced upon all the stress conditions tested (Figure 6). Also, growth was deficient for the *rpoE5* mutant in relation to the wild type in heat and cold shock, as well as pH stress. However, growth was not severely impaired for any of the ECF sigma factor mutants, in the tested stress conditions. It is unlikely that there are other mutations compensating for the absence of the sigma factor in the mutant strains tested, but it is probable that some of the RpoE sigma factors overlap in their functions and have similar consensus sequences for binding to the promoters. The lack of significantly severe phenotypes for the *rpoE* mutants tested in this work could then be due to the redundancy of extracytoplasmic sigma factors in the *S. meliloti* genome and due to the fact that those sigma factors can have similar functions. That is, there is strong possibility that some *rpoE* genes complement the absence of other *rpoEs* by taking part in the transcription of their targets. There are for instance large overlaps between the regulons of ECF sigma factors in *B. subtilis* [62, 88].

RpoH proteins from different rhizobia are very similar and have conserved domains, such as the RpoH box (Figure 5). A phylogenetic tree based on the similarities of RpoH protein sequences of different rhizobia is coherent with the evolutionary tree for those organisms. This is also an indication that this class of sigma factors is of ancient origin. In *R. etli*, *rpoH1* is involved mainly in oxidative and heat shock responses while *rpoH2* is involved in osmotic tolerance and oxidative stress. Both genes are also involved in the senescence of nodules in the symbiotic processes [87]. An *rpoH2* mutant was not studied in this work, but it has already been described in previous publications that it does not have a heat stress phenotype [97, 110]. In this work, a lack of growth was however observed for the *rpoH1* mutant at heat stress conditions, as previously described in the literature [97, 110]. Moreover, a novel phenotype for the *rpoH1* mutant was observed at pH stress conditions and the involvement of RpoH1 in pH stress response was further analysed. Its implications will be further discussed in section 3. With that in mind, it is coherent to discuss that the RpoH proteins probably regulate different sets of genes and are specialized for different stress responses in *S. meliloti*.

This diversity in RpoH and other alternative sigma factors could provide an increase in the bacterium's ability to adapt to a great diversity of inhospitable environmental conditions.

Studies with the *fecI* mutant were performed to a less extent and it was not possible to observe a peculiar growth phenotype for this mutant in the conditions tested. Although no nodulation experiments were performed in this work, it is likely that the many stress conditions potentially encountered by the bacteria in infection threads are signals for activation of alternative sigma factors and upregulation of their regulons. However, no extreme growth defect of the *rpoE* or *fecI* mutant strains was detected in the growth tests.

2. COMPLEX SIGMA FACTOR INVOLVEMENT IN THE REGULATION OF THE RHIZOBACTIN OPERON

The differential expression of very few genes in the microarrays between *rpoH1* mutant and *S. meliloti* 1021 at pH 7.0 most likely accounts for the absence of an inhospitable environmental condition or an effector that could activate the alternative *rpoH1* transcriptional response. Though the differential expression of the rhizobactin operon in *rpoH1* mutant cells was confirmed by CAS assays, it may simply reflect the need for iron uptake regulation at pH 7.0. Nevertheless, necessity for rhizobactin cannot be the reason for which the *rpoH1* mutant is defective in nitrogen fixation and undergoes senescence after nodulation [110], since rhizobactin is not expressed in the nodules [83].

Analyses of the promoter region of *rhbA*, the first gene of the rhizobactin biosynthesis operon, showed that there are possible binding sites for RpoH1 and RpoE2 about 90 nt upstream of the translation start codon. The expression of *rhbA* was not altered in the *rpoH1 rpoE2* double mutant, in comparison the *rpoH1* single mutant, but a reduction in *rhbA* expression was observed for the *rpoH1 rpoE5* double mutant. It is coherent to presume, then, that RpoE5 and RpoE2 have similar consensus binding sites, as it is a common feature observed in RpoE sigma factors [62, 88]. Figure 29 illustrates a possible model for the regulation of this operon. In hospitable conditions, the RNAP holoenzyme containing the RpoH1 sigma factor could be responsible for the transcription of the rhizobactin operon, with the indispensable regulation of the transcription activator RhrA. The expression of *rhrA* itself is repressed regulator RirA [24]. In the absence of RpoH1, RpoE5 could be responsible for promoting

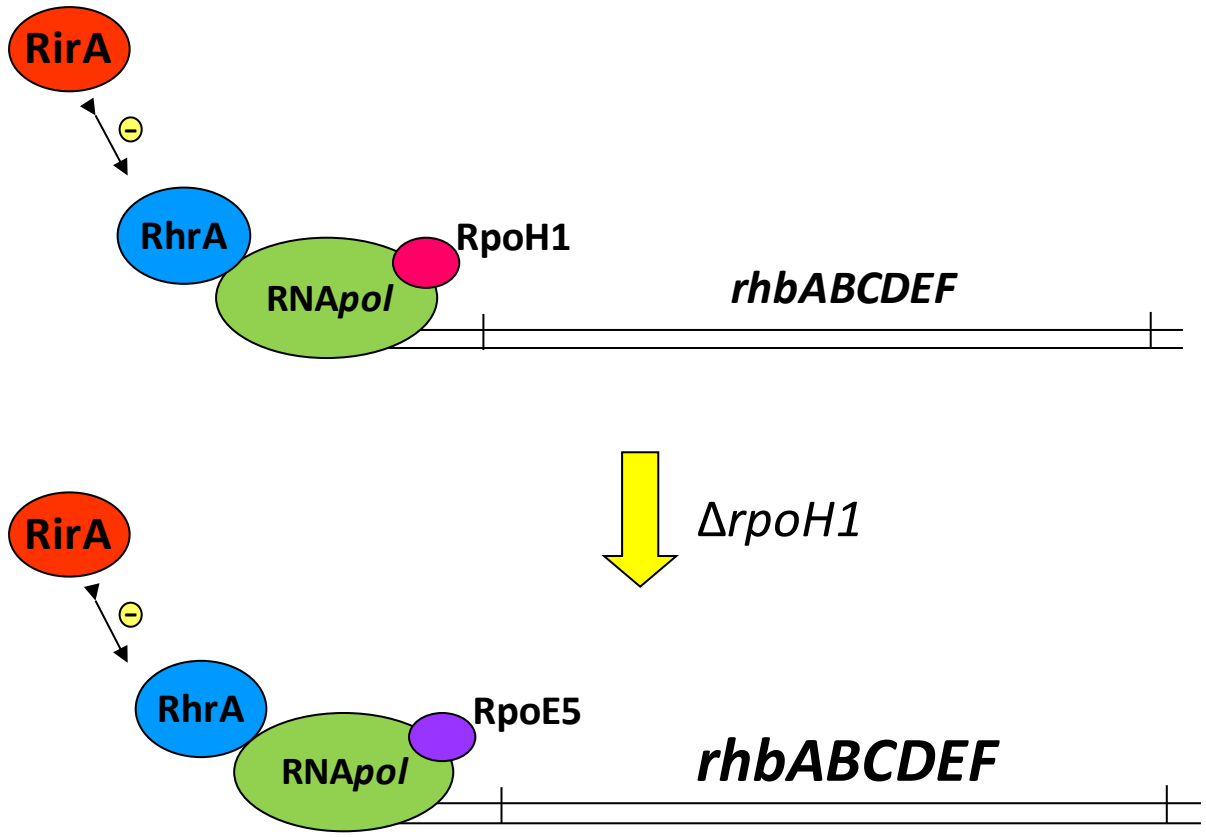


Figure 29. Possible model for the regulation of the rhizobactin operon. Under normal environmental conditions, the RNA polymerase holoenzyme containing the RpoH1 sigma factor could perform transcription of the rhizobactin biosynthesis operon, with the involvement of the transcription activator RhrA. The expression of *rhrA* itself is regulated by the global iron regulator RirA. In the absence of RpoH1, an alternative sigma factor is responsible for the transcription of the rhizobactin operon. Even though there is *rhrA* downregulation by RirA, the combination of the RNA polymerase with the alternative sigma factor could yield a higher expression of the rhizobactin operon. The difference in font size for the operon indicates significant increase of expression levels in the alternate condition.

transcription of the rhizobactin operon. Even though *rhrA* downregulation by RirA is more prominent, reflecting the increase in iron import, the combination of the RNA polymerase with the alternative sigma factor could yield more intense expression levels of the rhizobactin operon genes. However, it is important to point out that *rhbA* expression was not completely abolished, nor as reduced as in the wild type. Again, complementarity between the different alternative sigma factors could play a role in the regulation. There could be another binding site for another sigma factor that was not detected in our experiments. Or another model could be proposed in which absence of RpoH1 leads to the derepression of the *rhb* operon from a secondary unidentified regulator. Overall, it could be concluded that the upregulation of multiple genes related to rhizobactin synthesis in the microarray analyses of the *rpoH1* mutant at pH 7 might indicate a need for increased iron uptake regulation at neutral pH conditions.

3. THE *S. MELILOTI* SIGMA FACTOR RPOH1 IS A REGULATOR OF STRESS RESPONSE AT ACIDIC PH

This work suggests for the first time that RpoH1 efficiently regulates the expression of specific genes in response to pH stress in *S. meliloti.* The *S. meliloti rpoH1* mutant is unable to grow at pH 5.75 and a growth defect was also observed after pH shock experiments. The growth inhibition is probably caused as a consequence of the lowering of internal pH, which inhibits cell metabolism. Experiments by Mitsui et al. [97] confirmed the pH sensitivity phenotype in that an *rpoH1* mutant could not grow under pH stress. The microarray hybridization employed to investigate the time-course response of *S. meliloti* 1021 to a sudden acid shift revealed the induction of heat shock regulons and exopolysaccharide production and the repression of flagellar and chemotaxis genes. These results are in agreement with those described by Hellweg [61] and confirm the notion of an induced exopolysaccharide production and a hampered motility activity of *S. meliloti* upon pH shock. However, the microarray hybridization in the *rpoH1* mutant revealed novel information, in that many genes involved in heat shock and stress response that had been differentially expressed in the wild type were absent from the mutant arrays.

Clustering analyses and profile plotting enabled the classification of transcriptionally regulated *S. meliloti* genes in genes which were regulated in an RpoH1-independent, an RpoH1-dependent or in a complex manner upon pH stress. The genes included in this class of

RpoH1-independently regulated genes do not seem to have a specific stress response function. They comprise exopolysaccharide I biosynthesis genes, like *exoQ, exoP, exoN* and *exoY,* and also the group of genes involved in motility and flagellar biosynthesis, like the flagellar genes *flgA, flgL* and *mcpT* [131] [61]. So, the induction of exopolysaccharide production and the repression of motility is an RpoH1-independent process. The class of genes that responded in an RpoH1-dependent manner was composed of genes known to be involved in heat shock, such as *ibpA, grpE, clbP* and *groEL5,* as well as some genes involved in translation like *tufA* and *rplC*. Based on the functional classification of those genes, it can be inferred that a transcriptional response to pH takes place in which processes that are very energy-consuming are inhibited and the transcription of genes involved in chaperone mechanisms is upregulated. The RpoH1-dependent genes seem to be paramount for an appropriate cellular response in fighting pH stress.

Some genes were transiently upregulated only in the wild type arrays, whereas in the *rpoH1* mutant arrays those genes were constantly upregulated. This was the case for the genes *dctA, ndvA* and *rsiA1*. Those genes were classified as having a complex regulation. The lack of downregulation of those genes in the *rpoH1* mutant probably indicates that a second regulator is needed for transcription or for mRNA stability. This secondary regulation could involve one or more transcription repressors which are regulated by RpoH1 and control the activity of those genes. Interestingly, *rsiA1* codes for the RpoE2 anti-sigma factor. RpoE2 is known to be involved in general stress and in oxidative stress response in *S. meliloti* [41, 129]. In the time-course comparison, *rsiA1* was regulated differently from the wild type in the *rpoH1* mutant.

Though no differential expression of *rpoE2* was observed upon pH stress in our experiments, it is possible that extracytoplasmic sigma factors are also transcriptionally activated during pH stress, in addition to RpoH1, which is mostly responsible for the stress response in the intracellular compartment. In *E. coli*, for instance, pH stress response is under the regulation of more than one sigma factor [60]. Moreover, many stress genes are also regulated by transcriptional repressors and activators, a number of which were induced in our microarray experiments. Those constitute a secondary activation and are important for the response to specific intracellular cues for the precise coordination of transcription changes according to the physiological state of the cell.

4. CHARACTERIZATION OF RPOH1-DEPENDENT STRESS RESPONSE AND REGULATION OF RPOH1

4.1 Functional classification of RpoH1-dependently-regulated genes

Important genes involved in stress response are upregulated in the wild type arrays, but not in the *rpoH1* mutant arrays. The RpoH1-dependent genes were assigned to 18 functional COG categories. This suggests that there is a global effect on gene expression dependent on RpoH1 upon pH shock. Most of the genes encoded proteins that could be classified as chaperones, proteases, or proteins involved in adaptation to stress conditions. One major function of chaperones is to prevent both synthesized polypeptide chains and assembled subunits from aggregating into nonfunctional structures. It is for this reason that many chaperones are also heat shock proteins, because the tendency to aggregate increases as proteins are denatured by stress. Genes like *groEL5* and *clpB* have already been described as genes whose transcription is RpoH1-dependent in *S. meliloti* [16, 97], which corroborates with the results obtained in this work. The present results also identified a number of new targets not previously identified as RpoH-controlled genes in *S. meliloti*. They include many genes that encode the classical HSPs and other proteins involved in protein homeostasis. Additionally, genes that encode proteins that maintain genomic integrity by protecting DNA and RNA or function as effectors of transcription or translation were also identified. For instance, RpoH targets identified in the present analysis include genes encoding the chaperone machines GroEL-GroES and DnaK-DnaJ-GrpE and the proteases Lon and ClpB. The COG classification of RpoH1-dependent genes also confirmed that the class of protein turnover and chaperones was the most represented one. Moreover, quantitative real time RT-PCR of the sigma factor genes showed that there is indeed a differential regulation of these genes in response to unfavourable environmental conditions. The genetic circuits under the regulation of the RpoH1 sigma factors are likely to enable the cells to handle pH and even other environmental stresses.

The pH difference across the inner cell membrane contributes to cell energy in the form of proton potential or proton motive force. Low intracellular pH amplifies the dissipation of the proton potential that powers motility, ATP synthesis, and catabolite transport. For this reason, the cell's transmembrane proton potential is diminished by the maintenance of an inverted pH

gradient. In wild type arrays following pH shift, membrane-bound systems for proton transport were regulated by acid in a way that was very consistent with their relative degree of export or import of protons. An example is the downregulation of genes involved in proton import, such as ATP synthase genes. On the other hand, the *cyoABC* operon, which codes for a cytochrome oxidase proton pump, was upregulated, though their expression values were below the M-value threshold. Those genes associated with osmotic and oxidative stress were differentially expressed after acid shift in the wild type, but had however no differential expression in the *rpoH1* mutant arrays. A large part of functions and chemical events in Gram-negative bacteria take place in the outer membrane envelope and the periplasm, compartments essentially exposed to extracellular pH. Thus, several envelope and periplasmic components show pH-dependent expression. Genes included in this category are the ones coding for the proteases DegP1, ClpB and HtpX. Reduction of membrane lipid composition and overproduction of transmembrane proteins might indicate cellular strive to lower membrane permeability for protons. There is probably a connection of RpoH1 function to membrane integrity because the *rpoH1* mutant is hypersensitive to detergents and hydrophobic compounds [97], phenotypes that are associated with membrane integrity defects [15, 80].

4.2 RpoH1 response might not be unique to heat shock nor to pH stress

Often the consequences of a temperature shift rather than the temperature itself provide a signal that the cell recognizes and responds to. It is likely that the efficient regulation structure of RpoH sigma factors can be induced for adjustment not only to thermal adaptation, but also to other stresses that require rapid changes in the metabolism. The present results revealed patterns of pH response and clarified the overlap of pH stress with heat shock response. Since many of the proteins involved in heat shock response in bacteria are also induced by a variety of other environmental stress conditions, it can be concluded that such stress response can be activated by other stresses other than heat shock. RpoH1 has been described in *S. meliloti* as the heat shock response sigma factor [97, 110, 111]. The group of proteins shown to be involved in the heat shock response under the transcriptional control of RpoH1 includes chaperones, proteases, and regulatory factors. These groups of proteins are also involved in pH stress response, as seen by the microarray profiles in this work. Therefore, the pH stress

response in *S. meliloti* characterized here is likewise not specific for pH stress itself, but is likely to be a response to other types of environmental stress, including heat shock. Moreover, several stress responses are known to be linked with pH stress, including oxidative stress, heat shock, and envelope stress. Low pH usually accelerates acid consumption and proton export [42], and increases production of oxygen radicals, thus inducing a partial oxidative stress response. The RpoH sigma factor has already been implicated in the oxidative stress response in other rhizobia [29, 87].

4.3 Regulation of RpoH1 in *S. meliloti* 1021

A consensus binding sequence for RpoH1 was observed upstream of the *rpoH1* gene. This provides first indication that *rpoH1* is positively autoregulated from a σ^{32}-dependent promoter. However, that might not be the only promoter upstream the *rpoH1* gene. In *E.coli*, there are at least five promoters upstream of the *rpoH* gene, recognized by sigma 70, sigma E and sigma 54 [70]. In *R. etli*, there are two promoter sequences upstream of the *rpoH1* gene. P1 is a strong promoter that resembles the consensus boxes of the *R. etli* sigma 70 promoter and transcription from the P2 promoter is specific for heat shock stress [117]. In other α-proteobacteria, such as *Brucella melitensis*, *rpoE* sigma factors regulate the expression of *rpoH*. In *C. crescentus* the autoregulation of *rpoH* by a sigma 32 promoter also occurs [4, 29]. In *S. meliloti,* the presence of the RpoH1 consensus sequence indicates that RpoH1 controls its own transcription. The present analyses revealed that RpoH1 controls at least 100 genes, and about 20 of them contain the conserved motif in the upstream regulatory regions. This result may be explained by the possibility that RpoH1 controls other regulators, since *rpoE2* anti-sigma is regulated in a complex manner in the mutant, and therefore, more than one regulon may be affected. The RpoH consensus sequence identified here resembles the sequences of promoters recognized by other RpoH sigma factors, such as those of *E. coli* RpoH and *R etli* RpoH1.This study may have missed RpoH-dependent promoters that require the function of other transcription factors induced by other conditions. It is noticeably important that RpoH is required to regulate its own expression during environmental stress conditions such as low pH or heat shock

5. IMPLICATIONS ON THE ROLE OF RPOH1 IN SYMBIOSIS AND NITROGEN FIXATION

Rhizobia are usually exposed to oxidative and pH stresses during symbiosis, and stress response is required to withstand and to deal with eventual cellular damage. RpoH proteins are regulators that might indirectly connect the stress response with nitrogen fixation. The present work did not explore regulation within the nodule, another condition in which *rpoH1* is expressed [110]. Previous studies have shown that an *rpoH1* mutant is capable of eliciting the formation of nodules on alfalfa plants, but the *rpoH1* mutation causes early senescence of bacteroids during the endosymbiotic process [97, 110]. It can be speculated that the early senescence observed for *rpoH1* mutant nodules [97] is caused by an increased sensitivity to pH stress upon rhizosphere and plant acidification during nodulation. The endosymbiotic process would then be affected by the ability of rhizobial cells to protect themselves against environmental pH stress encountered within the host. In this case the role of RpoH1 during pH shift would be extremely important not only at free-living growth, as shown in this work, but also during symbiosis. Initiation, infection, and bacteroid development almost certainly require the synthesis of many proteins that are not present during free-living growth and multiple stress response genes might be required for symbiosis-specific functions. Although RpoH may direct expression of the classic heat shock genes, the requirement for RpoH1 during symbiosis may reflect the need for the expression of other genes, perhaps specific to rhizobia and nitrogen fixation. In fact, some of the genes in the microarray results for pH shock performed in this work overlap with transcriptomic studies that analysed *S. meliloti* under microoxic and symbiotic conditions [11]. Therefore, there is probably an additional function for the RpoH1-dependent genes identified in this work during symbiosis or there could be additional targets of RpoH1 that are necessary for symbiosis.

6. TIME-COURSE TRANSCRIPTIONAL PROFILING AND K-MEANS CLUSTERING ARE EFFICIENT APPROACHES FOR STUDYING GENES INVOLVED IN DYNAMIC PROCESSES SUCH AS STRESS RESPONSE

Time-course experiments of gene expression facilitate the understanding of the temporal structure of regulatory mechanisms and the identification of gene networks involved in stress response [30]. Here, the RpoH1-dependent pH stress response of *S. meliloti* was characterized

with the aid of transcriptomic studies and a time-course microarray hybridization was employed to investigate the response of *S. meliloti* to a sudden acid shift. The time-series microarrays, followed by clustering, enabled us to capture multiple expression profiles at discrete time points of a continuous, but very dynamic, cellular process. This was fundamental for extracting the basic patterns of gene expression inherent in the data. In addition to the recognition of individual genes with altered expressions, the proposed method for clustering of time-course data enabled us to identify gene clusters, each with a unique time-dependent expression pattern. K-means clustering provided the means to divide the robust microarray data in sets or groups of genes that had similar expression profile. This method proved to be very well suited for studying the RpoH1 regulation in the pH stress response of *S. meliloti*. A disadvantage of K-means clustering, however, is that the number of clusters is manually set by the user, so an optimal number of clusters cannot be automatically generated. Additionally, a program was designed for plotting the profiles of each individual gene, obtained from the microarray data, both for the wild type and for the *rpoH* mutant in one same graphic. This facilitated immensely the comparison of the transcription profiles of specific genes between wild type and *rpoH1* mutant (Appendix 7).

7. CONCLUSIONS AND OUTLOOK

The present study indicated that sigma factor RpoH1 plays an important role in fighting low pH stress in *S. meliloti*. This role was efficiently unravelled by time-course microarray studies, in which key players involved in stress response whose transcription is under regulation of RpoH1 were identified. A great advantage of the transcriptional profiling experiments was that it enabled the identification of genes and operons whose expression is RpoH-dependent. Clustering of time-course microarray data of *S. meliloti* wild type and *rpoH1* mutant allowed for the classification of three groups of genes that were transcriptionally regulated upon pH stress in an RpoH1-independent, in an RpoH1-dependent or in a complex manner. Among the genes that showed an RpoH1-dependent regulation, there were several coding for heat shock and chaperone proteins. RpoH1 is necessary for the dynamic response of *S. meliloti* to sudden pH shift and it accounts for critical changes in gene expression during pH stress response. These data suggest that *rpoH1* plays, directly or indirectly, a relevant role in survival under free-living conditions and possibly in the senescence of nodules as a master regulator of pH stress.

Time-course global gene expression analyses should be further employed to facilitate the temporal study of regulatory mechanisms and provide a more comprehensive framework for studying dynamic cellular processes such as stress response. Time-course microarray analyses could be performed, for instance, for testing the role of other sigma factor mutants in diverse stress conditions. The studies performed here, together with the observation in other bacteria, suggest that the pH stress response in rhizobia could overlap the response to other stimuli. Another valid approach would be to closely compare the RpoH pH stress regulon to the heat shock regulon to see if RpoH1 controls expression of genes that are induced exclusively in response to heat stress or to pH stress. This could be done, accordingly, for other sigma factors, as time-course stress response data is gathered for further sigma factor mutants and stress conditions. This would provide undoubtedly a more complete approach to understanding the intricate regulatory circuits under sigma regulation.

Characterization of alternative sigma factors and the ability of rhizobia to deal with environmental stress responses is of fundamental importance. Genetic studies on the regulatory events of *S. meliloti* cells undergoing environmental stress should continue to provide useful information for further understanding of the role of other alternative sigma factors in stress response. The findings in this work form a basis for subsequent analyses of regulation and function of the stress response in *S. meliloti*. The time-course study provides efficient methodology for hypothesis-driven investigations to dissect the roles of sigma factors and other key players in transcription regulation not only in pH stress conditions, but in general stress response and adaptation.

F. REFERENCES

1. Alexandre A, and Oliveira S. Most heat-tolerant rhizobia show high induction of major chaperone genes upon stress. *FEMS Microbiol Ecol* 75: 28-36, 2010.
2. Alloing G, Travers I, Sagot B, Le Rudulier D, and Dupont L. Proline betaine uptake in *Sinorhizobium meliloti*: Characterization of Prb, an opp-like ABC transporter regulated by both proline betaine and salinity stress. *J Bacteriol* 188: 6308-6317, 2006.
3. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, and Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389-3402, 1997.
4. Alvarez-Martinez CE, Lourenco RF, Baldini RL, Laub MT, and Gomes SL. The ECF sigma factor sigma(T) is involved in osmotic and oxidative stress responses in *Caulobacter crescentus*. *Mol Microbiol* 66: 1240-1255, 2007.
5. Anthony JR, Warczak KL, and Donohue TJ. A transcriptional response to singlet oxygen, a toxic byproduct of photosynthesis. *Proc Natl Acad Sci U S A* 102: 6502-6507, 2005.
6. Barne KA, Bown JA, Busby SJ, and Minchin SD. Region 2.5 of the *Escherichia coli* RNA polymerase sigma70 subunit is responsible for the recognition of the 'extended-10' motif at promoters. *EMBO J* 16: 4034-4040, 1997.
7. Barnett MJ, Fisher RF, Jones T, Komp C, Abola AP, Barloy-Hubler F, Bowser L, Capela D, Galibert F, Gouzy J, Gurjal M, Hong A, Huizar L, Hyman RW, Kahn D, Kahn ML, Kalman S, Keating DH, Palm C, Peck MC, Surzycki R, Wells DH, Yeh KC, Davis RW, Federspiel NA, and Long SR. Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti* pSymA megaplasmid. *Proc Natl Acad Sci U S A* 98: 9883-9888, 2001.
8. Barnett MJ, Toman CJ, Fisher RF, and Long SR. A dual-genome Symbiosis Chip for coordinate study of signal exchange and development in a prokaryote-host interaction. *Proc Natl Acad Sci U S A* 101: 16636-16641, 2004.
9. Barrios H, Valderrama B, and Morett E. Compilation and analysis of sigma(54)-dependent promoter sequences. *Nucleic Acids Res* 27: 4305-4313, 1999.
10. Bastiat B, Sauviac L, and Bruand C. Dual control of *Sinorhizobium meliloti* RpoE2 sigma factor activity by two PhyR-type two-component response regulators. *J Bacteriol* 192: 2255-2265, 2010.
11. Becker A, Bergès H, Krol E, Bruand C, Rüberg S, Capela D, Lauber E, Meilhoc E, Ampe F, Bruijn FJd, Fourment J, Francez-Charlot A, D. Kahn HK, Liebe C, Pühler A, Weidner S, and Batut J. Global changes in gene expression in *Sinorhizobium meliloti* 1021 under microoxic and symbiotic conditions. *Mol Plant-Microbe Interact* 17: 292-303, 2004.
12. Becker A, Ruberg S, Baumgarth B, Bertram-Drogatz PA, Quester I, and Pühler A. Regulation of succinoglycan and galactoglucan biosynthesis in *Sinorhizobium meliloti*. *J Mol Microbiol Biotechnol* 4: 187-190, 2002.
13. Bentley SD, Chater KF, Cerdeno-Tarraga AM, Challis GL, Thomson NR, James KD, Harris DE, Quail MA, Kieser H, Harper D, Bateman A, Brown S, Chandra G, Chen CW, Collins M, Cronin A, Fraser A, Goble A, Hidalgo J, Hornsby T, Howarth S,

Huang CH, Kieser T, Larke L, Murphy L, Oliver K, O'Neil S, Rabbinowitsch E, Rajandream MA, Rutherford K, Rutter S, Seeger K, Saunders D, Sharp S, Squares R, Squares S, Taylor K, Warren T, Wietzorrek A, Woodward J, Barrell BG, Parkhill J, and Hopwood DA. Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). *Nature* 417: 141-147, 2002.

14. Beringer JE. R factor transfer in *Rhizobium leguminosarum*. *J Gen Microbiol* 84: 188-198, 1974.

15. Bittner AN. Molecular genetic analysis of multiple *rpoH* and *groEL* genes in *Sinorhizobium meliloti PhD Thesis* 2008.

16. Bittner AN, and Oke V. Multiple groESL operons are not key targets of RpoH1 and RpoH2 in *Sinorhizobium meliloti*. *J Bacteriol* 188: 3507-3515, 2006.

17. Bolanos L, El-Hamdaoui A, and Bonilla I. Recovery of development and functionality of nodules and plant growth in salt-stressed Pisum sativum--*Rhizobium leguminosarum* symbiosis by boron and calcium. *J Plant Physiol* 160: 1493-1497, 2003.

18. Borukhov S, Lee J, and Laptenko O. Bacterial transcription elongation factors: new insights into molecular mechanism of action. *Mol Microbiol* 55: 1315-1324, 2005.

19. Braun V, Mahren S, and Ogierman M. Regulation of the FecI-type ECF sigma factor by transmembrane signalling. *Curr Opin Microbiol* 6: 173-180, 2003.

20. Brooks BE, and Buchanan SK. Signaling mechanisms for activation of extracytoplasmic function (ECF) sigma factors. *Biochim Biophys Acta* 1778: 1930-1945, 2008.

21. Cabiscol E, Tamarit J, and Ros J. Oxidative stress in bacteria and protein damage by reactive oxygen species. *Int Microbiol* 3: 3-8, 2000.

22. Camarero JA, Shekhtman A, Campbell EA, Chlenov M, Gruber TM, Bryant DA, Darst SA, Cowburn D, and Muir TW. Autoregulation of a bacterial sigma factor explored by using segmental isotopic labeling and NMR. *Proc Natl Acad Sci U S A* 99: 8536-8541, 2002.

23. Campbell EA, Westblade LF, and Darst SA. Regulation of bacterial RNA polymerase sigma factor activity: a structural perspective. *Curr Opin Microbiol* 11: 121-127, 2008.

24. Chao TC, Buhrmester J, Hansmeier N, Pühler A, and Weidner S. Role of the regulatory gene *rirA* in the transcriptional response of *Sinorhizobium meliloti* to iron limitation. *Appl Environ Microbiol* 71: 5969-5982, 2005.

25. Cleveland WS. Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association* 74: 829-836, 1979.

26. Cloutier J, Prevost D, Nadeau P, and Antoun H. Heat and cold shock protein synthesis in arctic and temperate strains of rhizobia. *Appl Environ Microbiol* 58: 2846-2853, 1992.

27. Dakora FD, Joseph CM, and Phillips DA. Alfalfa (Medicago sativa L.) Root Exudates Contain Isoflavonoids in the Presence of *Rhizobium meliloti*. *Plant Physiol* 101: 819-824, 1993.

28. Davey ME, and de Bruijn FJ. A homologue of the tryptophan-rich sensory protein TspO and FixL regulate a novel nutrient deprivation-induced *Sinorhizobium meliloti* locus. *Appl Environ Microbiol* 66: 5353-5359, 2000.

29. Delory M, Hallez R, Letesson J-J, and De Bolle X. An RpoH-like heat shock sigma factor is involved in stress response and virulence in *Brucella melitensis* 16M. *J Bacteriol* 188: 7707-7710, 2006.

30. DeRisi JL, Iyer VR, and Brown PO. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278: 680-686, 1997.

31. Diaz-Acosta A, Sandoval ML, Delgado-Olivares L, and Membrillo-Hernandez J. Effect of anaerobic and stationary phase growth conditions on the heat shock and oxidative stress responses in *Escherichia coli* K-12. *Arch Microbiol* 185: 429-438, 2006.

32. Diouf A, de Lajudie P, Neyra M, Kersters K, Gillis M, Martinez-Romero E, and Gueye M. Polyphasic characterization of rhizobia that nodulate *Phaseolus vulgaris* in West Africa (Senegal and Gambia). *Int J Syst Evol Microbiol* 50 Pt 1: 159-170, 2000.

33. Dombrecht B, Marchal K, Vanderleyden J, and Michiels J. Prediction and overview of the RpoN-regulon in closely related species of the Rhizobiales. *Genome Biol* 3: RESEARCH0076, 2002.

34. Dondrup M, Goesmann A, Bartels D, Kalinowski J, Krause L, Linke B, Rupp O, Sczyrba A, Pühler A, and Meyer F. EMMA: a platform for consistent storage and efficient analysis of microarray data. *J Biotechnol* 106: 135-146, 2003.

35. Dsouza M, Larsen N, and Overbeek R. Searching for patterns in genomic data. *Trends Genet* 13: 497-498, 1997.

36. EMMA server http://www.cebitec.uni-bielefeld.de/groups/brf/software/emma_info/.

37. Fenner BJ, Tiwari RP, Reeve WG, Dilworth MJ, and Glenn AR. *Sinorhizobium medicae* genes whose regulation involves the ActS and/or ActR signal transduction proteins. *FEMS Microbiol Lett* 236: 21-31, 2004.

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

39. Fisher RF, and Long SR. Rhizobium--plant signal exchange. *Nature* 357: 655-660, 1992.

40. Flechard M, Fontenelle C, Blanco C, Goude R, Ermel G, and Trautwetter A. RpoE2 of *Sinorhizobium meliloti* is necessary for trehalose synthesis and growth in hyperosmotic media. *Microbiology* 156: 1708-1718, 2010.

41. Flechard M, Fontenelle C, Trautwetter A, Ermel G, and Blanco C. *Sinorhizobium meliloti rpoE2* is necessary for H(2)O(2) stress resistance during the stationary growth phase. *FEMS Microbiol Lett* 290: 25-31, 2009.

42. Foster JW. *Escherichia coli* acid resistance: tales of an amateur acidophile. *Nat Rev Microbiol* 2: 898-907, 2004.

43. Foster PL. Stress-induced mutagenesis in bacteria. *Crit Rev Biochem Mol Biol* 42: 373-397, 2007.

44. Galibert F, Finan TM, Long SR, Pühler A, Abola P, Ampe F, Barloy-Hubler F, Barnett MJ, Becker A, Boistard P, Bothe G, Boutry M, Bowser L, Buhrmester J, Cadieu E, Capela D, Chain P, Cowie A, Davis RW, Dreano S, Federspiel NA, Fisher RF, Gloux S, Godrie T, Goffeau A, Golding B, Gouzy J, Gurjal M, Hernandez-Lucas

I, Hong A, Huizar L, Hyman RW, Jones T, Kahn D, Kahn ML, Kalman S, Keating DH, Kiss E, Komp C, Lelaure V, Masuy D, Palm C, Peck MC, Pohl TM, Portetelle D, Purnelle B, Ramsperger U, Surzycki R, Thebault P, Vandenbol M, Vorholter FJ, Weidner S, Wells DH, Wong K, Yeh KC, and Batut J. The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science* 293: 668-672, 2001.

45. Galibert F, Finan TM, Long SR, Pühler A, et al.. The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science* 293: 668-672, 2001.

46. Ghosh T, Bose D, and Zhang X. Mechanisms for activating bacterial RNA polymerase. *FEMS Microbiol Rev* 34: 611-627.

47. Gibson KE, Kobayashi H, and Walker GC. Molecular determinants of a symbiotic chronic infection. *Annu Rev Genet* 42: 413-441, 2008.

48. Giuliodori AM, Gualerzi CO, Soto S, Vila J, and Tavio MM. Review on bacterial stress topics. *Ann N Y Acad Sci* 1113: 95-104, 2007.

49. Gonzalez JE, York GM, and Walker GC. *Rhizobium meliloti* exopolysaccharides: synthesis and symbiotic function. *Gene* 179: 141-146, 1996.

50. Gonzalez V, Santamaria RI, Bustos P, Hernandez-Gonzalez I, Medrano-Soto A, Moreno-Hagelsieb G, Janga SC, Ramirez MA, Jimenez-Jacinto V, Collado-Vides J, and Davila G. The partitioned *Rhizobium etli* genome: genetic and metabolic redundancy in seven interacting replicons. *Proc Natl Acad Sci U S A* 103: 3834-3839, 2006.

51. Gottesman S. The small RNA regulators of *Escherichia coli*: roles and mechanisms*. *Annu Rev Microbiol* 58: 303-328, 2004.

52. Graham PH, Draeger, K.J., Ferrey, M.L., Conroy, M.J., Hammer, B.E., Martinez, E., Aarons SR, Quinto C. Acid pH tolerance in strains of *Rhizobium* and *Bradyrhizobium*, and initial studies on the basis for acid tolerance of *Rhizobium tropici* Umr1899. *Can J Microbiol* 40: 1994.

53. Grant SG, Jessee J, Bloom FR, and Hanahan D. Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proc Natl Acad Sci U S A* 87: 4645-4649, 1990.

54. Green HA, and Donohue TJ. Activity of *Rhodobacter sphaeroides* RpoHII, a second member of the heat shock sigma factor family. *J Bacteriol* 188: 5712-5721, 2006.

55. Gruber TM, and Gross CA. Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu Rev Microbiol* 57: 441-466, 2003.

56. Guisbert E, Yura T, Rhodius VA, and Gross CA. Convergence of molecular, modeling, and systems approaches for an understanding of the *Escherichia coli* heat shock response. *Microbiol Mol Biol Rev* 72: 545-554, 2008.

57. Gutteridge JM, and Halliwell B. Comments on review of Free Radicals in Biology and Medicine, second edition, by Barry Halliwell and John M. C. Gutteridge. *Free Radic Biol Med* 12: 93-95, 1992.

58. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser* 41: 95-98., 1999.

59. Harley CB, and Reynolds RP. Analysis of *E. coli* promoter sequences. *Nucleic Acids Res* 15: 2343-2361, 1987.

60. Harrison C. GrpE, a nucleotide exchange factor for DnaK. *Cell Stress Chaperones* 8: 218-224, 2003.

61. Hellweg C, Pühler A, and Weidner S. The time course of the transcriptomic response of *Sinorhizobium meliloti* 1021 following a shift to acidic pH. *BMC Microbiol* 9: 37-37, 2009.

62. Helmann JD. The extracytoplasmic function (ECF) sigma factors. *Adv Microb Physiol* 46: 47-110, 2002.

63. Hengge-Aronis R. Signal transduction and regulatory mechanisms involved in control of the sigma(S) (RpoS) subunit of RNA polymerase. *Microbiol Mol Biol Rev* 66: 373-395, table of contents, 2002.

64. Hengge R. The two-component network and the general stress sigma factor RpoS (sigma S) in *Escherichia coli*. *Adv Exp Med Biol* 631: 40-53, 2008.

65. Himmelreich R, Hilbert H, Plagens H, Pirkl E, Li BC, and Herrmann R. Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae. *Nucleic Acids Res* 24: 4420-4449, 1996.

66. Horn G, Hofweber R, Kremer W, and Kalbitzer HR. Structure and function of bacterial cold shock proteins. *Cell Mol Life Sci* 64: 1457-1470, 2007.

67. Horton RM. PCR-mediated recombination and mutagenesis. SOEing together tailor-made genes. *Mol Biotechnol* 3: 93-99, 1995.

68. Imamura S, Yoshihara S, Nakano S, Shiozaki N, Yamada A, Tanaka K, Takahashi H, Asayama M, and Shirai M. Purification, characterization, and gene expression of all sigma factors of RNA polymerase in a cyanobacterium. *J Mol Biol* 325: 857-872, 2003.

69. Ishihama A. Functional modulation of *Escherichia coli* RNA polymerase. *Annu Rev Microbiol* 54: 499-518, 2000.

70. Janaszak A, Majczak W, Nadratowska B, Szalewska-Palasz A, Konopa G, and Taylor A. A sigma54-dependent promoter in the regulatory region of the *Escherichia coli rpoH* gene. *Microbiology* 153: 111-123, 2007.

71. Kaneko T, Nakamura Y, Sato S, Asamizu E, Kato T, Sasamoto S, Watanabe A, Idesawa K, Ishikawa A, Kawashima K, Kimura T, Kishida Y, Kiyokawa C, Kohara M, Matsumoto M, Matsuno A, Mochizuki Y, Nakayama S, Nakazaki N, Shimpo S, Sugimoto M, Takeuchi C, Yamada M, and Tabata S. Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti DNA Res* 7: 331-338, 2000.

72. Kaneko T, Nakamura Y, Sato S, Minamisawa K, Uchiumi T, Sasamoto S, Watanabe A, Idesawa K, Iriguchi M, Kawashima K, Kohara M, Matsumoto M, Shimpo S, Tsuruoka H, Wada T, Yamada M, and Tabata S. Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110 (supplement). *DNA Res* 9: 225-256, 2002.

73. Kazmierczak MJ, Mithoe SC, Boor KJ, and Wiedmann M. *Listeria monocytogenes* sigma B regulates stress response and virulence functions. *J Bacteriol* 185: 5722-5734, 2003.

74. Keseler IM, and Kaiser D. sigma54, a vital protein for *Myxococcus xanthus*. *Proc Natl Acad Sci U S A* 94: 1979-1984, 1997.

75. Klinkert B, and Narberhaus F. Microbial thermosensors. *Cell Mol Life Sci* 66: 2661-2676, 2009.

76. Kogoma T, and Yura T. Sensitization of *Escherichia coli* cells to oxidative stress by deletion of the *rpoH* gene, which encodes the heat shock sigma factor. *J Bacteriol* 174: 630-632, 1992.

77. Krol E, and Becker A. Global transcriptional analysis of the phosphate starvation response in *Sinorhizobium meliloti* strains 1021 and 2011. *Mol Genet Genomics* 272: 1-17, 2004.

78. Kumari BDR. *Plant Proteomics*. A.P.H. Publishing Corporation, 2008.

79. Labes M, Rastogi V, Watson R, and Finan TM. Symbiotic nitrogen fixation by a nifA deletion mutant of *Rhizobium meliloti*: the role of an unusual ntrC allele. *J Bacteriol* 175: 2662-2673, 1993.

80. Lagares A, Caetano-Anolles G, Niehaus K, Lorenzen J, Ljunggren HD, Pühler A, and Favelukes G. A *Rhizobium meliloti* lipopolysaccharide mutant altered in competitiveness for nodulation of alfalfa. *J Bacteriol* 174: 5941-5952, 1992.

81. Lonetto M, Gribskov M, and Gross CA. The sigma 70 family: sequence conservation and evolutionary relationships. *J Bacteriol* 174: 3843-3849, 1992.

82. Lund PA. Multiple chaperonins in bacteria--why so many? *FEMS Microbiol Rev* 33: 785-800, 2009.

83. Lynch D, O'Brien J, Welch T, Clarke P, Cuiv PO, Crosa JH, and O'Connell M. Genetic organization of the region encoding regulation, biosynthesis, and transport of rhizobactin 1021, a siderophore produced by *Sinorhizobium meliloti*. *J Bacteriol* 183: 2576-2585, 2001.

84. MacLean AM, Finan TM, and Sadowsky MJ. Genomes of the symbiotic nitrogen-fixing bacteria of legumes. *Plant Physiol* 144: 615-622, 2007.

85. Marchler-Bauer A, Anderson JB, DeWeese-Scott C, Fedorova ND, Geer LY, He S, Hurwitz DI, Jackson JD, Jacobs AR, Lanczycki CJ, Liebert CA, Liu C, Madej T, Marchler GH, Mazumder R, Nikolskaya AN, Panchenko AR, Rao BS, Shoemaker BA, Simonyan V, Song JS, Thiessen PA, Vasudevan S, Wang Y, Yamashita RA, Yin JJ, and Bryant SH. CDD: a curated Entrez database of conserved domain alignments. *Nucleic Acids Res* 31: 383-387, 2003.

86. Margaret I, Becker A, Blom J, Bonilla I, Goesmann A, Gottfert M, Lloret J, Mittard-Runte V, Ruckert C, Ruiz-Sainz JE, Vinardell JM, and Weidner S. Symbiotic properties and first analyses of the genomic sequence of the fast growing model strain *Sinorhizobium fredii* HH103 nodulating soybean. *J Biotechnol*.

87. Martínez-Salazar JM, Sandoval-Calderon M, Guo X, Castillo-Ramirez S, Reyes A, Loza MG, Rivera J, Alvarado-Affantranger X, Sanchez F, Gonzalez V, Davila G, and Ramirez-Romero MA. The *Rhizobium etli* RpoH1 and RpoH2 sigma factors are involved in different stress responses. *Microbiology* 155: 386-397, 2009.

88. Mascher T, Hachmann AB, and Helmann JD. Regulatory overlap and functional redundancy among *Bacillus subtilis* extracytoplasmic function sigma factors. *J Bacteriol* 189: 6919-6927, 2007.

89. Masson-Boivin C, Giraud E, Perret X, and Batut J. Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes? *Trends Microbiol* 17: 458-466, 2009.

90. Meade HM, Long SR, Ruvkun GB, Brown SE, and Ausubel FM. Physical and genetic characterization of symbiotic and auxotrophic mutants of *Rhizobium meliloti* induced by transposon Tn5 mutagenesis. *J Bacteriol* 149: 114-122, 1982.

91. Mellor RB. Bacteroids in the *Rhizobium*-legume symbiosis inhabit a plant internal lytic compartment: implications for other microbial endosymbioses. *Journal of Experimental Botany* 831-839., 1989.

92. Mergaert P, Uchiumi T, Alunni B, Evanno G, Cheron A, Catrice O, Mausset AE, Barloy-Hubler F, Galibert F, Kondorosi A, and Kondorosi E. Eukaryotic control on bacterial cell cycle and differentiation in the Rhizobium-legume symbiosis. *Proc Natl Acad Sci U S A* 103: 5230-5235, 2006.

93. Merrick M, and Chambers S. The helix-turn-helix motif of sigma 54 is involved in recognition of the -13 promoter region. *J Bacteriol* 174: 7221-7226, 1992.

94. Meyer F, Goesmann A, McHardy AC, Bartels D, Bekel T, Clausen J, Kalinowski J, Linke B, Rupp O, Giegerich R, and Pühler A. GenDB--an open source genome annotation system for prokaryote genomes. *Nucl Acids Res* 31: 2187-2195, 2003.

95. Miksch G, and Dobrowolski P. Growth phase-dependent induction of stationary-phase promoters of *Escherichia coli* in different gram-negative bacteria. *J Bacteriol* 177: 5374-5378, 1995.

96. Missiakas D, and Raina S. The extracytoplasmic function sigma factors: role and regulation. *Mol Microbiol* 28: 1059-1066, 1998.

97. Mitsui H, Sato T, Sato Y, Ito N, and Minamisawa K. *Sinorhizobium meliloti* RpoH1 is required for effective nitrogen-fixing symbiosis with alfalfa. *Mol Genet Genomics* 271: 416-425, 2004.

98. Morett E, and Segovia L. The sigma 54 bacterial enhancer-binding protein family: mechanism of action and phylogenetic relationship of their functional domains. *J Bacteriol* 175: 6067-6074, 1993.

99. Morita MT, Kanemori M, Yanagi H, and Yura T. Dynamic interplay between antagonistic pathways controlling the sigma 32 level in *Escherichia coli*. *Proc Natl Acad Sci U S A* 97: 5860-5865, 2000.

100. Morita MT, Tanaka Y, Kodama TS, Kyogoku Y, Yanagi H, and Yura T. Translational induction of heat shock transcription factor sigma32: evidence for a built-in RNA thermosensor. *Genes Dev* 13: 655-665, 1999.

101. Mrabet M, Zribi K, Mhadhbi H, Djébali N, Mhamdi R, Aouani M, and Nakamura K. Salt tolerance of a *Sinorhizobium meliloti* strain isolated from dry lands: growth capacity and protein profile changes. *Annals of Microbiology* 61: 361-369.

102. Muller PY, Janovjak H, Miserez AR, and Dobbie Z. Processing of gene expression data generated by quantitative real-time RT-PCR. *Biotechniques* 32: 1372-1374, 1376, 1378-1379, 2002.

103. Münchbach M, Nocker A, and Narberhaus F. Multiple small heat shock proteins in rhizobia. *J Bacteriol* 181: 83-90, 1999.

104. Murakami KS, and Darst SA. Bacterial RNA polymerases: the wholo story. *Curr Opin Struct Biol* 13: 31-39, 2003.

105. Nakahigashi K, Yanagi H, and Yura T. Isolation and sequence analysis of *rpoH* genes encoding sigma 32 homologs from gram negative bacteria: conserved mRNA and protein segments for heat shock regulation. *Nucleic Acids Res* 23: 4383-4390, 1995.

106. Narberhaus F, and Balsiger S. Structure-function studies of *Escherichia coli* RpoH (sigma32) by in vitro linker insertion mutagenesis. *J Bacteriol* 185: 2731-2738, 2003.

107. Narberhaus F, Krummenacher P, Fischer HM, and Hennecke H. Three disparately regulated genes for sigma 32-like transcription factors in *Bradyrhizobium japonicum*. *Mol Microbiol* 24: 93-104, 1997.

108. Nies DH. Incidence and function of sigma factors in *Ralstonia metallidurans* and other bacteria. *Arch Microbiol* 181: 255-268, 2004.

109. O'Hara G W, Goss TJ, Dilworth MJ, and Glenn AR. Maintenance of Intracellular pH and Acid Tolerance in *Rhizobium meliloti*. *Appl Environ Microbiol* 55: 1870-1876, 1989.

110. Oke V, Rushing BG, Fisher EJ, Moghadam-Tabrizi M, and Long SR. Identification of the heat-shock sigma factor RpoH and a second RpoH-like protein in *Sinorhizobium meliloti*. *Microbiology* 147: 2399-2408, 2001.

111. Ono Y, Mitsui H, Sato T, and Minamisawa K. Two RpoH homologs responsible for the expression of heat shock protein genes in *Sinorhizobium meliloti*. *Mol Gen Genet* 264: 902-912, 2001.

112. Ott T, van Dongen JT, Gunther C, Krusell L, Desbrosses G, Vigeolas H, Bock V, Czechowski T, Geigenberger P, and Udvardi MK. Symbiotic leghemoglobins are crucial for nitrogen fixation in legume root nodules but not for general plant growth and development. *Curr Biol* 15: 531-535, 2005.

113. Paget MS, and Helmann JD. The sigma70 family of sigma factors. *Genome Biol* 4: 203, 2003.

114. Patschkowski T, Schluter A, and Priefer UB. *Rhizobium leguminosarum* bv. viciae contains a second fnr/fixK-like gene and an unusual fixL homologue. *Mol Microbiol* 21: 267-280, 1996.

115. Poole P, and Allaway D. Carbon and nitrogen metabolism in Rhizobium. *Adv Microb Physiol* 43: 117-163, 2000.

116. Raivio TL, and Silhavy TJ. Periplasmic stress and ECF sigma factors. *Annu Rev Microbiol* 55: 591-624, 2001.

117. Ramirez-Romero MA, Masulis I, Cevallos MA, Gonzalez V, and Davila G. The *Rhizobium etli* sigma70 (SigA) factor recognizes a lax consensus promoter. *Nucleic Acids Res* 34: 1470-1480, 2006.

118. Reeve WG, Brau L, Castelli J, Garau G, Sohlenkamp C, Geiger O, Dilworth MJ, Glenn AR, Howieson JG, and Tiwari RP. The *Sinorhizobium medicae* WSM419 *lpiA* gene is transcriptionally activated by FsrR and required to enhance survival in lethal acid conditions. *Microbiology* 152: 3049-3059, 2006.

119. Reeve WG, Tiwari RP, Guerreiro N, Stubbs J, Dilworth MJ, Glenn AR, Rolfe BG, Djordjevic MA, and Howieson JG. Probing for pH-regulated proteins in *Sinorhizobium medicae* using proteomic analysis. *J Mol Microbiol Biotechnol* 7: 140-147, 2004.

120. Reisenauer A, Mohr CD, and Shapiro L. Regulation of a heat shock sigma32 homolog in *Caulobacter crescentus*. *J Bacteriol* 178: 1919-1927, 1996.

121. Robson RL, and Postgate JR. Oxygen and hydrogen in biological nitrogen fixation. *Annu Rev Microbiol* 34: 183-207, 1980.

122. Ronson CW, Nixon BT, Albright LM, and Ausubel FM. *Rhizobium meliloti* ntrA (rpoN) gene is required for diverse metabolic functions. *J Bacteriol* 169: 2424-2431, 1987.

123. Rüberg S, Tian Z-X, Krol E, Linke B, Meyer F, Wang Y, Pühler A, Weidner S, and Becker A. Construction and validation of a *Sinorhizobium meliloti* whole genome DNA microarray: genome-wide profiling of osmoadaptive gene expression. *J Biotechnol* 106: 255-268, 2003.

124. Rushing BG, and Long SR. Cloning and characterization of the sigA gene encoding the major sigma subunit of *Rhizobium meliloti*. *J Bacteriol* 177: 6952-6957, 1995.

125. Sakoh M, Ito K, and Akiyama Y. Proteolytic activity of HtpX, a membrane-bound and stress-controlled protease from *Escherichia coli*. *J Biol Chem* 280: 33305-33310, 2005.

126. Sambrook J, Fritsch EF, and Maniatis T. *Molecular cloning: A laboratory manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press, 1989.

127. Sambrook J, Fritsch EF, and Maniatis T editors. *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y, 1989.

128. Santos R, Herouart D, Puppo A, and Touati D. Critical protective role of bacterial superoxide dismutase in rhizobium-legume symbiosis. *Mol Microbiol* 38: 750-759, 2000.

129. Sauviac L, Philippe H, Phok K, and Bruand C. An extracytoplasmic function sigma factor acts as a general stress response regulator in *Sinorhizobium meliloti*. *J Bacteriol* 189: 4204-4216, 2007.

130. Schäfer A, Tauch A, Jäger W, Kalinowski J, Thierbach G, and Pühler A. Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene* 145: 69-73, 1994.

131. Scharf B, and Schmitt R. Sensory transduction to the flagellar motor of *Sinorhizobium meliloti*. *J Mol Microbiol Biotechnol* 4: 183-186, 2002.

132. Schwyn B, and Neilands JB. Universal chemical assay for the detection and determination of siderophores. *Anal Biochem* 160: 47-56, 1987.

133. Sharma UK, and Chatterji D. Transcriptional switching in *Escherichia coli* during stress and starvation by modulation of sigma activity. *FEMS Microbiol Rev* 34: 646-657, 2010.

134. Simon LD, Randolph B, Irwin N, and Binkowski G. Stabilization of proteins by a bacteriophage T4 gene cloned in *Escherichia coli*. *Proc Natl Acad Sci U S A* 80: 2059-2062, 1983.

135. *Sinorhizobium meliloti* 1021 Sm14kOLI [http://www.cebitec.uni-bielefeld.de/transcriptomics/transcriptomics-facility/sm14koli.html].

136. Studer S, and Narberhaus F. Chaperone activity and homo- and hetero-oligomer formation of bacterial small heat shock proteins. *J Biol Chem* 275: 37212-37218, 2000.

137. Studholme DJ, and Buck M. The biology of enhancer-dependent transcriptional regulation in bacteria: insights from genome sequences. *FEMS Microbiol Lett* 186: 1-9, 2000.

138. Sturn A, Quackenbush J, and Trajanoski Z. Genesis: cluster analysis of microarray data. *Bioinformatics* 18: 207-208, 2002.

139. Summers ML, Botero LM, Busse SC, and McDermott TR. The *Sinorhizobium meliloti* lon protease is involved in regulating exopolysaccharide synthesis and is required for nodulation of alfalfa. *J Bacteriol* 182: 2551-2558, 2000.

140. Suziedeliene E, Suziedelis K, Garbenciute V, and Normark S. The acid-inducible *asr* gene in *Escherichia coli*: transcriptional control by the *phoBR* operon. *J Bacteriol* 181: 2084-2093, 1999.

141. Tatusov RL, Koonin EV, and Lipman DJ. A genomic perspective on protein families. *Science* 278: 631-637, 1997.

142. Thompson JD, Higgins DG, and Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22: 4673-4680, 1994.

143. Tiwari RP, Reeve WG, Dilworth MJ, and Glenn AR. Acid tolerance in *Rhizobium meliloti* strain WSM419 involves a two-component sensor-regulator system. *Microbiology* 142 (Pt 7): 1693-1704, 1996.

144. Tiwari RP, Reeve WG, Dilworth MJ, and Glenn AR. An essential role for *actA* in acid tolerance of *Rhizobium meliloti*. *Microbiology* 142 (Pt 3): 601-610, 1996.

145. Van de Velde W, Zehirov G, Szatmari A, Debreczeny M, Ishihara H, Kevei Z, Farkas A, Mikulass K, Nagy A, Tiricz H, Satiat-Jeunemaitre B, Alunni B, Bourge M, Kucho K, Abe M, Kereszt A, Maroti G, Uchiumi T, Kondorosi E, and Mergaert P. Plant peptides govern terminal differentiation of bacteria in symbiosis. *Science* 327: 1122-1126, 2010.

146. Vance CP. Symbiotic nitrogen fixation and phosphorus acquisition. Plant nutrition in a world of declining renewable resources. *Plant Physiol* 127: 390-397, 2001.

147. Vasse J, de Billy F, Camut S, and Truchet G. Correlation between ultrastructural differentiation of bacteroids and nitrogen fixation in alfalfa nodules. *J Bacteriol* 172: 4295-4306, 1990.

148. Vincent JM. A manual for the practical study of root nodule bacteria. In: *IBP Handbook No 15* Blackwell Scientific Publication, Oxford, 1970.

149. Wade JT, Roa DC, Grainger DC, Hurd D, Busby SJ, Struhl K, and Nudler E. Extensive functional overlap between sigma factors in *Escherichia coli*. *Nat Struct Mol Biol* 13: 806-814, 2006.

150. Wang C, Kemp J, Da Fonseca IO, Equi RC, Sheng X, Charles TC, and Sobral BW. *Sinorhizobium meliloti* 1021 loss-of-function deletion mutation in *chvI* and its phenotypic characteristics. *Mol Plant Microbe Interact* 23: 153-160, 2010.

151. Weir BS. The current taxonomy of rhizobia. *New Zealand rhizobia website* http://www.rhizobia.co.nz/taxonomy/rhizobia.html., 2011.

152. Wells DH, and Long SR. The *Sinorhizobium meliloti* stringent response affects multiple aspects of symbiosis. *Mol Microbiol* 43: 1115-1127, 2002.

153. Werner E, Holder AA, and Hoheisel JD. Growth and storage of YAC clones in Hogness Freezing Medium. *Nucl Acids Res* 25: 1467–1468, 1997.

154. Xu H, and Hoover TR. Transcriptional regulation at a distance in bacteria. *Curr Opin Microbiol* 4: 138-144, 2001.

155. Yang YH, Dudoit S, Luu P, Lin DM, Peng V, Ngai J, and Speed TP. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res* 30: e15, 2002.

156. Young JP, Crossman LC, Johnston AW, Thomson NR, Ghazoui ZF, Hull KH, Wexler M, Curson AR, Todd JD, Poole PS, Mauchline TH, East AK, Quail MA, Churcher C, Arrowsmith C, Cherevach I, Chillingworth T, Clarke K, Cronin A, Davis P, Fraser A, Hance Z, Hauser H, Jagels K, Moule S, Mungall K, Norbertczak H, Rabbinowitsch E, Sanders M, Simmonds M, Whitehead S, and Parkhill J. The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. *Genome Biol* 7: R34, 2006.

157. Yura T, and Nakahigashi K. Regulation of the heat-shock response. *Curr Opin Microbiol* 2: 153-158, 1999.

158. Yurgel S, Mortimer MW, Rogers KN, and Kahn ML. New substrates for the dicarboxylate transport system of *Sinorhizobium meliloti*. *J Bacteriol* 182: 4216-4221, 2000.

159. Yurgel SN, Rice J, Mulder M, and Kahn ML. GlnB/GlnK PII proteins and regulation of the *Sinorhizobium meliloti* Rm1021 nitrogen stress response and symbiotic function. *J Bacteriol* 192: 2473-2481.

160. Zahran HH. *Rhizobium*-legume symbiosis and nitrogen fixation under severe conditions and in an arid climate. *Microbiol Mol Biol Rev* 63: 968-989, 1999.

161. Zehr JP, Jenkins BD, Short SM, and Steward GF. Nitrogenase gene diversity and microbial community structure: a cross-system comparison. *Environ Microbiol* 5: 539-554, 2003.

162. Zhou YN, Kusukawa N, Erickson JW, Gross CA, and Yura T. Isolation and characterization of *Escherichia coli* mutants that lack the heat shock sigma factor sigma 32. *J Bacteriol* 170: 3640-3649, 1988.

1. SPREADSHEET OF *S. MELILOTI* WILD TYPE GENES WHICH WERE DIFFERENTIALLY EXPRESSED FOLLOWING ACIDIC pH SHIFT

Table 1. Wild type genes differentially expressed after acidic pH shift

Gene Name	Annotation	M-value at time point (minutes)					
		0	5	10	15	30	60
SMa0045	Cah Probable carbonic anhydrase	-0,25	1,98	3,19	3,64	2,75	2,40
SMa0172	Conserved Hypothetical protein	0,29	0,73	1,09	1,61	1,52	1,75
SMa0392	ABC transporter, periplasmic solute-binding protein	0,40	-0,33	-0,46	-0,49	-1,41	-1,43
SMa0473	Conserved Hypothetical protein	0,48	0,89	1,26	1,86	1,96	1,34
SMa0994	Hypothetical protein	0,19	0,24	0,55	0,75	1,28	2,25
SMa1077	Nex18 Symbiotically induced conserved protein	0,28	0,99	2,20	1,90	1,73	1,98
SMa1078	Conserved Hypothetical protein	0,64	0,67	1,76	2,10	1,95	2,41
SMa1079	TspO Tryptophan rich sensory protein homologue	0,13	0,45	1,50	2,08	2,08	1,50
SMa1082	Hypothetical protein	1,57	0,83	1,34	1,12	0,30	0,23
SMa1151	Conserved Hypothetical protein	0,98	0,93	2,06	2,58	2,15	2,07
SMa1182	NosZ N2O reductase	1,53	0,24	0,52	1,08	0,35	0,31
SMa1220	FixN1 Heme b / copper cytochrome c oxidase subunit	1,53	0,65	1,11	0,77	0,36	0,10
SMa1243	Azu1 pseudoazurin (blue copper protein)	1,48	0,45	0,65	1,11	0,33	-0,02
SMa1750	Hypothetical protein	0,33	0,20	-0,22	-0,26	0,08	1,53
SMa1898	Hypothetical protein	-0,31	-0,12	-0,22	-1,53	-0,15	-0,12
SMa1961	Putative Polyhydrixyalkanoate depolymerase	0,18	1,25	1,40	1,42	0,68	0,89
SMb20325	ThuE probable trehalosemaltose-binding protein	-0,02	-0,08	-0,11	0,02	-0,80	-1,75
SMb20359	Hypothetical protein	-0,03	1,79	1,81	1,88	2,11	1,83
SMb20436	Putative nitrate transporter protein	0,65	-0,59	-1,20	-0,83	-0,93	-1,41
SMb20486	Putative sugar ABC transporter permease protein	0,00	0,07	0,26	0,69	0,81	1,49
SMb20497	lyx Putative L-xylulose kinase protein	-0,16	-1,27	-0,65	-1,41	-0,95	-0,93
SMb20560	Conserved Hypothetical protein	-0,15	0,66	0,63	1,30	2,08	0,63
SMb20605	Putative ureashort-chain amide or branched-chain amino acid uptake ABC transporter	0,48	-2,68	-0,72	-0,08	-0,25	0,13
SMb20611	dctA C4-dicarboxylate transport protein	-0,15	0,52	1,16	1,53	1,95	-0,17
SMb20696	Hypothetical protein	0,07	0,76	1,08	1,51	1,79	1,07
SMb20707	cyaG2 Putative adenylate cyclase protein	0,13	0,62	1,02	1,64	2,50	1,90
SMb20777	TRm19 Putative transposase of insertion sequence ISRm19 protein	-0,65	0,05	-0,67	-0,16	-0,35	-1,43
SMb20944	exoQ Putative polysaccharide polymerase, similar to Wzy protein	0,37	0,52	1,02	1,25	1,27	1,78
SMb20946	exoY galactosyltransferase protein	-0,12	0,82	1,14	1,30	1,48	2,18
SMb20949	exoV Putative pyruvyltransferase protein	0,60	0,80	1,49	2,08	1,90	2,12
SMb20954	exoH succinyltransferase protein	0,30	1,05	1,91	2,36	2,68	3,17
SMb20955	exoK endo-beta-1,3-1,4-glycanase protein	0,09	0,52	1,23	1,53	1,70	2,06
SMb20956	exoL Putative glucosyltransferase protein	0,19	0,34	0,48	1,03	1,19	1,41
SMb20960	exoN UDPglucose pyrophosphorylase protein	-0,16	0,07	0,39	0,78	2,43	3,30
SMb20984	nirB Putative nitrite reductase [NAD(P)H], large subunit protein	0,50	-1,96	-1,28	-0,39	-0,45	-0,94
SMb20985	nirD Putative nitrite reductase [NAD(P)H], small subunit protein	0,58	-1,17	-1,47	-0,69	-0,24	-0,80
SMb20986	narB Putative nitrate reductase, large subunit protein	0,50	-0,05	-1,70	-0,48	-0,40	-0,89
SMb20988	Conserved Hypothetical protein	-0,76	-0,38	-1,82	-2,40	-0,49	-2,22
SMb21176	phoD phosphate uptake ABC transporter periplasmic solute-binding protein precursor	0,20	0,19	-1,33	-1,71	-1,76	-0,89
SMb21177	phoC phosphate uptake ABC transporter ATP-binding protein	0,22	-0,75	-1,65	-1,62	-1,59	-0,83
SMb21236	Putative ATPGTP-binding protein	0,16	0,21	0,43	0,82	1,72	1,09

SMb21259	Hypothetical exported protein precursor	0,11	0,00	0,33	0,77	1,63	1,11
SMb21295	Putative small heat shock protein, hsp20 family	0,02	0,74	0,93	1,08	1,97	1,92
SMb21440	Hypothetical protein	-0,24	0,13	0,50	1,11	2,24	2,96
SMb21491	Hypothetical exported protein	-0,08	0,22	0,36	0,55	1,45	2,17
SMb21516	Conserved Hypothetical protein	-0,09	1,13	1,66	1,84	1,55	2,14
SMb21566	groEL5 Putative heat shock protein groEL	-0,27	-0,24	0,32	0,82	2,38	2,54
SMb21690	exoW glucosyltransferase protein	0,00	0,61	1,15	1,30	1,33	1,62
SMb21707	Putative ureashort-chain amide or branched-chain amino acid uptake ABC transporter ATP-binding protein	0,02	0,29	-1,93	0,45	0,70	0,21
SMc00043	sodB Superoxide dismutase Fe protein	0,30	0,25	0,38	0,72	1,26	1,46
SMc00045	cycF Putative Cytochrome C signal peptide protein	0,01	1,23	2,06	2,17	2,04	2,27
SMc00062	Hypothetical protein	-0,01	0,96	1,67	1,77	2,14	1,95
SMc00063	Hypothetical transmembrane protein	-0,01	0,71	1,40	1,60	1,35	1,11
SMc00070	Conserved Hypothetical signal peptide protein	0,28	0,40	0,92	1,07	1,24	1,42
SMc00094	betB Betaine aldehyde dehydrogenase BADH oxireductese NAD protein	-0,01	-0,10	-0,77	-1,22	-1,45	-0,67
SMc00103	dhe Putative Alpha-halocarboxylic acid dehalogenase protein	0,25	1,11	1,39	1,40	1,41	1,62
SMc00109	Putative transcription regulator protein	-0,10	1,18	1,27	1,52	1,66	0,41
SMc00115	Conserved Hypothetical protein	-0,02	0,34	0,50	1,10	1,37	1,43
SMc00159	Hypothetical signal peptide protein	0,50	0,94	0,54	0,58	-0,38	-2,17
SMc00186	Putative ABC transporter ATP-binding transmembrane protein	0,69	0,84	1,73	2,42	2,15	1,19
SMc00276	Conserved Hypothetical protein	-1,05	0,17	-0,58	-0,94	-1,47	-0,44
SMc00283	Putative transcription regulator protein	-0,23	-0,84	-0,88	-1,24	-0,99	-1,47
SMc00301	sufA putative FeS assembly scaffold	0,15	0,35	0,63	1,01	1,51	1,40
SMc00302	Conserved Hypothetical protein	0,45	0,26	0,55	1,22	2,00	2,19
SMc00329	irr Putative Iron response regulator protein	0,25	0,62	0,79	1,15	1,41	0,94
SMc00341	Hypothetical transmembrane protein	0,20	0,66	0,97	1,12	2,24	1,82
SMc00346	Hypothetical transmembrane protein	-0,04	0,94	1,05	1,42	1,33	1,72
SMc00565	rplI Probable 50S ribosomal protein L9	0,20	0,12	-0,82	-2,12	-0,28	-0,09
SMc00591	Hypothetical/unknown signal peptide protein	-0,58	-0,03	-0,53	0,50	1,12	1,68
SMc00610	Conserved Hypothetical protein	0,75	0,88	1,26	1,38	1,32	1,49
SMc00611	lpiA Transmembrane protein	-0,05	1,11	1,66	2,31	2,40	1,04
SMc00638	Putative heat resistant agglutinin 1 protein	0,19	-0,61	-0,90	-0,91	-1,52	-1,76
SMc00641	serA Putative D-3-phosphoglycerate dehydrogenase protein	0,46	0,25	1,77	0,48	-0,43	-0,93
SMc00810	hypothetical protein	-0,52	-0,17	-1,43	-1,72	0,13	-0,27
SMc00827	Putative transport transmembrane protein	-1,44	-0,08	-0,32	-0,82	0,10	-0,61
SMc00897	pmbA Hypothetical PMBA protein	0,23	0,13	0,29	0,92	1,41	1,18
SMc00914	Putative oxidoreductase protein	0,28	0,37	0,69	1,04	1,42	1,35
SMc00949	Conserved Hypothetical protein	0,42	0,53	0,78	2,18	3,29	2,90
SMc01107	Conserved Hypothetical protein	0,09	0,59	0,85	1,45	1,92	1,33
SMc01142	grpE Probable heat shock protein	0,08	-0,10	0,19	0,46	1,43	1,29
SMc01224	trxB Probable Thioredoxin reductase protein	0,21	0,42	0,45	0,72	1,66	1,17
SMc01225	Putative transcription regulator protein	0,30	1,16	1,71	2,09	1,82	1,89
SMc01226	Putative transcription regulator protein	0,31	1,42	1,15	1,48	1,01	0,93
SMc01308	rplD Probable 50S ribosomal protein L4	0,21	-0,13	-0,01	-1,72	-0,78	-0,80
SMc01309	rplC Probable 50S ribosomal protein L3	0,28	-0,01	-0,41	-1,63	-0,71	-0,72
SMc01311	tufA Probable elongation factor TU protein	-0,06	-0,54	-0,93	-1,68	-1,06	-1,03
SMc01314	rpsL Probable 30S ribosomal protein S12	0,04	-0,84	-1,45	-1,40	-1,08	-0,90
SMc01326	tufB Probable elongation factor TU protein	-0,33	-0,62	-0,88	-1,77	-1,23	-1,18
SMc01341	Hypothetical/unknown protein	-0,31	0,65	0,94	1,26	1,75	2,01
SMc01365	rnr Putative Exoribonuclease II protein	0,06	0,42	0,74	1,31	1,02	1,43
SMc01428	cspA2 Probable cold shock transcription regulator protein	0,37	0,39	0,89	1,21	1,70	1,43
SMc01440	hflC Putative hydrolase serine protease transmembrane protein	0,20	-0,12	0,59	0,80	1,61	1,47
SMc01441	hflK Putative membrane bound Protease protein	0,54	0,57	0,61	0,30	1,77	1,21

SMc01505	Anti-sigma factor	-0,26	0,78	1,13	1,51	0,81	-0,18
SMc01518	Conserved Hypothetical protein	0,03	-0,51	-0,41	-1,22	-0,60	-1,44
SMc01556	Conserved Hypothetical protein	-0,09	0,21	0,43	0,77	1,26	1,66
SMc01578	aatA Aspartate aminotransferase A (transaminase)	-0,11	-0,27	-0,55	-1,68	-1,11	-0,93
SMc01580	Hypothetical transmembrane protein	0,15	0,98	2,14	2,64	3,29	3,46
SMc01719	mcpT Probable chemoreceptor (methyl accepting chemotaxis) transmembrane protein	0,13	-0,39	-0,51	-0,64	-1,65	-2,68
SMc01764	ate Putative arginyl-tRNA protein transferase	0,27	1,22	1,50	1,77	1,57	1,08
SMc01765	Hypothetical transmembrane protein	0,43	1,63	2,33	2,86	2,55	2,79
SMc01769	Hypothetical protein	0,01	1,09	1,62	1,94	1,60	1,57
SMc01774	Hypothetical transmembrane protein	-0,01	0,60	1,32	2,02	2,11	2,38
SMc01788	Hypothetical protein	0,00	0,31	1,27	1,64	1,44	1,27
SMc01813	Conserved Hypothetical protein	0,07	0,57	0,92	1,29	1,71	1,38
SMc01815	Putative oxidoreductase Iron-sulfur protein	-0,11	0,90	1,04	1,70	0,08	-0,17
SMc01832	ureE Putative Urease accessory protein	-0,74	0,61	-1,18	-0,20	0,31	-2,37
SMc01848	Conserved Hypothetical protein	-0,02	-0,97	-2,57	-1,46	-1,28	-1,37
SMc01855	Hypothetical transmembrane protein	0,88	1,13	2,75	3,58	4,09	5,11
SMc01860	ftsI Probable Penicillin-binding transmembrane protein	-0,09	0,56	-0,86	-0,74	-0,60	-1,58
SMc01904	clpX Probable ATP-dependent CLP Protease ATP-binding subunit protein	0,04	1,28	1,76	2,00	2,52	1,39
SMc01905	lon Probable ATP-dependent Protease LA protein	0,13	0,74	1,44	2,17	2,71	1,97
SMc01947	Conserved Hypothetical transmembrane protein	0,16	0,15	-0,16	-0,10	-0,59	-1,53
SMc02052	Conserved Hypothetical protein	-0,35	1,97	0,19	0,08	-0,35	-0,54
SMc02075	Conserved Hypothetical protein	0,18	0,52	0,75	1,59	2,10	1,35
SMc02078	exoR Exopolysaccharide biosynthesis regulatorY protein	0,04	0,30	0,05	0,19	1,16	1,58
SMc02106	Conserved Hypothetical protein	0,54	0,61	1,08	1,54	0,33	-0,07
SMc02109	clpA Probable ATP-dependent CLP Protease ATP-binding subunit protein	0,09	0,45	0,65	0,83	1,48	0,96
SMc02110	ATP-dependent Clp protease adapter protein clpS1	-0,08	0,55	0,64	1,30	1,76	1,21
SMc02139	Hypothetical protein	0,20	0,18	-0,30	-2,41	0,10	-0,17
SMc02145	Hypothetical signal peptide protein	0,19	-2,05	-1,50	-1,40	-1,25	-0,24
SMc02146	Putative Phosphate-binding periplasmic protein	0,30	-0,34	-1,06	-1,28	-1,66	-0,86
SMc02151	Hypothetical virulence associated protein homologue	0,07	-1,05	-0,38	-1,62	-0,40	-0,31
SMc02156	Conserved Hypothetical protein	0,50	1,95	2,62	3,41	3,24	3,17
SMc02187	Putative Integrase DNA protein	-0,16	0,64	1,14	1,42	1,23	0,79
SMc02202	Hypothetical protein	-0,23	0,59	1,41	1,52	1,25	0,68
SMc02275	pncA Probable Pyrazinamidase/nicotinamidase protein	-0,61	-0,82	0,14	-1,77	-1,57	-2,63
SMc02278	Hypothetical unknown transmembrane protein	-0,23	-0,02	0,12	0,33	1,55	1,36
SMc02284	Hypothetical signal peptide protein	-0,31	-0,05	-0,44	-1,40	-1,05	-0,45
SMc02365	degP1 Protease precursor protein	-1,54	3,13	4,26	4,93	5,63	4,08
SMc02366	Putative transcription regulator protein	-2,18	2,48	3,52	3,39	4,01	2,81
SMc02382	Conserved Hypothetical protein	-0,12	0,38	1,44	2,56	3,49	2,80
SMc02390	gst7 Putative Glutathione S-transferase protein	-0,15	0,14	0,36	0,98	1,95	1,63
SMc02392	Hypothetical protein	-0,31	-0,20	-0,55	-1,81	-0,35	-0,57
SMc02396	Probable Outer membrane protein	0,04	-0,44	-0,17	-1,49	-0,65	-0,39
SMc02400	Probable Outer membrane protein	-0,03	-0,38	-1,05	-2,43	-1,63	-1,31
SMc02403	Putative Murein transglycosylase protein	1,80	0,22	-0,51	-0,49	0,70	-0,35
SMc02433	clpB ATP-dependent Protease (heat shock protein)	-0,27	0,03	0,22	1,09	2,98	1,73
SMc02435	hemK1 Putative Methyltransferase protein	-0,18	0,30	0,04	1,59	1,82	1,35
SMc02443	grxC Probable Glutaredoxin 3 protein	-0,18	0,44	0,75	1,18	1,93	1,24
SMc02479	mdh Probable Malate dehydrogenase protein	-0,01	0,33	-0,50	-0,97	-1,74	-1,48
SMc02480	sucC Probable Succinyl-coA synthetase beta chain protein	0,45	0,27	-0,04	-1,36	-1,73	-0,87
SMc02491	Hypothetical protein	0,43	0,90	1,25	1,92	1,60	1,91
SMc02518	Putative ATP-binding ABC transporter protein	-0,15	0,62	1,47	0,82	0,23	-0,31
SMc02560	chvI transcriptionAL regulatorY protein	-0,55	0,45	1,25	1,60	1,96	1,78

SMc02575	hslV Probable heat shock protein	-0,28	-0,06	0,23	0,64	1,54	1,05
SMc02576	Hypothetical Acetyltransferase protein	-0,14	0,09	0,40	0,82	1,89	1,21
SMc02582	Conserved Hypothetical protein	0,13	-1,14	-2,19	-2,59	-1,70	-1,43
SMc02634	Hypothetical transmembrane protein	-0,48	-2,72	-3,77	-3,80	-3,28	-1,84
SMc02655	Hypothetical transmembrane protein	0,03	1,04	1,04	1,50	1,58	1,14
SMc02720	clpP2 CLP Protease Proteolytic subunit protein	0,15	0,44	0,60	1,03	1,54	1,05
SMc02728	fhs Probable Formate-tetrahydrofolate ligase protein	-0,60	-0,10	-0,09	-0,67	-0,27	-1,57
SMc02735	Hypothetical protein	0,05	0,38	0,52	1,19	1,65	1,03
SMc02755	ahcY Probable Adenosylhomocysteinase protein	0,15	0,29	-0,18	-0,73	-1,17	-1,82
SMc02820	cpaF1 Putative pilus assembly protein	0,07	-0,52	-1,13	-2,24	-1,39	-0,99
SMc02885	msrA1 Probable peptide Methionine sulfoxide reductase protein	0,20	-0,19	1,64	1,41	1,77	1,58
SMc02898	kdsB Probable 3-deoxy-manno-octulosonate cytidylyltransferase (CMP-KDO synthetase) protein	0,39	0,25	-0,28	-3,27	-2,36	-1,44
SMc02978	Conserved Hypothetical protein	0,13	0,12	0,58	0,66	1,70	1,19
SMc02980	Hypothetical protein	-0,16	-0,91	-2,63	-0,13	0,15	-0,55
SMc03000	Putative permease ABC transporter protein	-0,72	-0,28	-0,46	-1,59	-0,63	-0,71
SMc03016	Putative transcription regulator protein	-1,07	-1,49	-1,04	-1,60	-0,54	-0,58
SMc03027	flgB flagellar basal-body rod protein	0,18	0,13	-0,19	-0,80	-2,64	-3,57
SMc03029	fliE flagellar hook-basal body complex protein	-0,12	0,07	-0,10	-0,67	-1,77	-2,99
SMc03030	flgG flagellar basal-body rod protein	0,05	0,02	-0,17	-0,64	-1,56	-2,07
SMc03037	flaA flagellin A protein	0,11	-0,93	-0,93	-1,26	-0,83	-1,54
SMc03038	flaB flagellin B protein	0,16	-1,06	-1,16	-1,52	-1,28	-1,92
SMc03039	flaD Probable flagellin D protein	-0,04	-0,82	-1,27	-1,59	-1,46	-1,34
SMc03046	Putative transcription regulator protein	0,31	0,07	-0,15	-0,26	-0,64	-1,87
SMc03047	flgE flagellar hook protein	0,19	0,19	-0,29	-0,75	-1,28	-1,60
SMc03049	flgL Putative flagellar hook-associated protein	-0,14	0,20	-0,21	-1,17	-3,60	-5,98
SMc03050	flaF Putative flagellin synthesis regulator protein	-0,29	0,04	-0,42	-0,99	-1,63	-1,17
SMc03064	aglA Probable alpha-glucosidase protein	-1,11	-0,12	-2,75	-0,27	0,02	0,11
SMc03090	cheW3 Putative chemotaxis protein	-0,08	-0,26	-0,69	-0,83	-0,77	-1,41
SMc03105	dxr Probable 1-deoxy-D-xylulose 5-phosphate reductoisomerase protein	-0,03	0,20	-1,45	-1,08	0,12	-0,43
SMc03111	pmi Mannose-6-phosphate isomerase protein	-0,10	-0,74	0,52	-1,47	-0,50	-1,48
SMc03138	Putative Sugar kinase protein	-0,89	-0,35	-0,35	-1,19	-0,59	-1,92
SMc03151	Conserved Hypothetical protein	0,06	0,58	0,78	1,44	2,33	1,75
SMc03152	Hypothetical transmembrane protein	-0,20	0,03	0,31	0,76	1,75	1,38
SMc03205	purU1 Putative Formyltetrahydrofolate deformylase protein	-0,17	-0,61	-0,75	-1,45	-0,87	-0,83
SMc03245	Putative Amidase protein	-1,50	-0,96	-1,14	-2,09	-0,48	-0,46
SMc03290	Hypothetical protein	1,22	-2,13	-1,47	-3,94	0,92	-0,27
SMc03773	Conserved Hypothetical protein	-0,20	-0,97	-0,88	-1,13	-0,96	-1,44
SMc03780	Hypothetical protein	0,32	1,13	1,71	2,43	3,40	2,59
SMc03783	ctpA Putative Carboxy-terminal processing protease precursor signal peptide protein	0,27	1,00	1,33	1,82	2,22	1,60
SMc03784	Hypothetical transmembrane protein	-1,43	-0,85	-0,54	-0,15	0,20	0,02
SMc03788	dnaE2 Putative DNA Polymerase III Alpha chain protein	-1,40	-1,51	-0,10	-2,78	-0,12	-2,13
SMc03802	Conserved Hypothetical protein	-0,10	0,34	0,61	1,01	1,78	0,95
SMc03806	glnK Probable Nitrogen regulatorY protein PII 2	0,23	-1,12	-1,81	-1,52	-1,61	-2,74
SMc03819	Conserved Hypothetical protein	-0,85	-1,63	-0,65	-0,31	-0,69	-1,07
SMc03836	tesA Putative Acyl-coA Thioesterase I protein	0,38	0,37	1,24	1,97	2,31	1,82
SMc03838	Hypothetical transmembrane protein	0,10	0,31	0,58	1,05	1,62	1,32
SMc03893	Putative Amino-acid transport system permease ABC transporter protein	0,12	0,67	0,98	1,40	0,90	0,37
SMc03900	ndvA Beta-1-->2Glucan export ATP-binding protein	0,23	1,61	1,57	0,87	0,00	0,29
SMc03999	Hypothetical protein	0,17	1,06	1,33	1,66	1,97	2,42
SMc04009	Conserved Hypothetical protein	-0,12	-0,47	-1,14	-2,04	-1,79	-1,26
SMc04026	gltD Probable glutamate synthase small chain protein	-0,29	-0,33	-1,34	-1,09	-1,11	-1,44

SMc04040	ibpA Heat shock protein	-0,33	0,09	0,16	1,35	3,42	2,00
SMc04059	Hypothetical protein	-0,25	-0,21	-0,27	-0,83	-1,55	-1,87
SMc04091	htpX Putative Protease transmembrane protein	0,32	0,57	0,79	1,64	2,62	1,81
SMc04111	cpaC1 Putative pilus assembly transmembrane protein	-0,12	-0,18	-0,61	-1,01	-0,96	-1,52
SMc04128	Putative heavy metal transporting ATPase protein	-0,06	0,44	0,57	1,55	2,83	0,77
SMc04167	Putative HIistidine-rich transporter transmembrane protein	0,21	-0,08	-0,07	0,36	1,46	0,62
SMc04213	dgkA Diacylglycerol kinase protein	-0,11	-1,46	0,05	0,23	-0,17	-0,20
SMc04246	Hypothetical transmembrane signal peptide protein	-0,31	0,80	1,47	1,58	1,76	0,84
SMc04267	lpsS LPS sulfotransferase	-0,05	1,25	1,23	1,65	1,78	1,62
SMc04346	ilvC ketol-acid reductoisomerase protein	-0,35	-0,58	-0,72	-1,32	-1,45	-1,39
SMc04350	Putative multidrug efflux system transmembrane protein	0,05	2,93	3,64	4,82	4,08	2,76
SMc04459	ftsH Probable metalloprotease transmembrane protein	0,14	0,13	0,55	0,91	2,49	1,82
SMc04461	TolB protein precursor	-0,31	0,47	0,67	1,01	1,17	1,40
SMc04865	Hypothetical protein	-0,09	0,31	0,45	1,02	1,66	0,89

2. SPREADSHEET OF 210 GENES USED FOR ANALYSIS OF *RPOH1* MUTANT EXPRESSION PROFILING FOLLOWING ACIDIC pH SHIFT

Table 2. Spreadsheet of genes used for comparative analysis of *rpoH1* mutant to *S. meliloti* wild type following acidic pH shift

Gene Name	Annotation	M-value at time point (minutes)					
		0	5	10	15	30	60
SMa0045	Cah Probable carbonic anhydrase	-0,23	0,87	1,38	1,87	1,70	1,30
SMa0172	Conserved Hypothetical protein	-0,25	0,32	0,54	0,73	0,92	1,44
SMa0392	ABC transporter, periplasmic solute-binding protein	0,17	-0,34	-0,79	-0,69	-0,99	-0,91
SMa0473	Conserved Hypothetical protein	-0,01	0,22	0,54	0,66	0,51	0,70
SMa0994	Hypothetical protein	0,14	-0,10	-0,14	-0,07	-0,04	-0,17
SMa1077	Nex18 Symbiotically induced conserved protein	0,11	1,15	2,06	2,60	2,80	3,13
SMa1078	Conserved Hypothetical protein	0,31	0,48	0,39	1,27	1,32	1,44
SMa1079	TspO Tryptophan rich sensory protein homologue	-0,44	-0,07	0,08	0,54	0,39	0,58
SMa1082	Hypothetical protein	-0,91	0,03	0,59	0,37	-0,63	-0,72
SMa1151	Conserved Hypothetical protein	0,90	0,88	1,09	1,17	1,30	1,46
SMa1182	NosZ N2O reductase	-0,33	-0,11	0,45	0,52	-0,05	-0,50
SMa1220	FixN1 Heme b / copper cytochrome c oxidase subunit	0,16	1,30	1,33	0,96	-0,21	-0,33
SMa1243	Azu1 pseudoazurin (blue copper protein)	-0,62	0,65	1,14	1,05	0,30	-0,40
SMa1750	Hypothetical protein	-0,11	-0,43	-0,50	-0,78	-0,77	0,20
SMa1898	Hypothetical protein	-0,20	-0,20	-0,38	-0,28	-0,40	-0,41
SMa1961	Putative Polyhydrixyalkanoate depolymerase	-0,48	-0,10	0,27	0,10	-0,18	-0,43
SMb20325	ThuE probable trehalosemaltose-binding protein	0,48	0,45	-0,59	-0,21	-0,77	-0,94
SMb20359	Hypothetical protein	0,33	0,52	0,52	1,07	0,98	1,92
SMb20436	Putative nitrate transporter protein	0,36	-1,12	-2,00	-1,22	-0,51	-0,61
SMb20486	Putative sugar ABC transporter permease protein	-0,07	0,05	0,04	0,29	0,93	1,90
SMb20497	lyx Putative L-xylulose kinase protein	-0,40	-0,38	-0,97	-1,02	-1,55	-1,83
SMb20560	Conserved Hypothetical protein	-0,19	1,91	2,42	2,14	2,19	1,74
SMb20605	Putative ureashort-chain amide or branched-chain amino acid uptake ABC transporter	0,19	-4,39	-3,07	-0,43	2,05	0,67
SMb20611	dctA C4-dicarboxylate transport protein	-0,54	0,13	0,49	1,37	1,91	1,92
SMb20696	Hypothetical protein	-0,30	0,13	0,41	0,47	0,39	0,47
SMb20707	cyaG2 Putative adenylate cyclase protein	-0,31	0,22	0,11	0,26	0,02	0,01
SMb20777	TRm19 Putative transposase of insertion sequence ISRm19 protein	-0,11	-0,03	0,07	-0,01	-0,13	0,31
SMb20944	exoQ Putative polysaccharide polymerase, similar to Wzy protein	-0,19	0,33	0,57	0,72	0,82	1,54
SMb20946	exoY galactosyltransferase protein	-0,30	0,62	0,93	1,19	1,13	2,00
SMb20949	exoV Putative pyruvyltransferase protein	0,06	0,72	0,98	1,74	2,10	2,78
SMb20954	exoH succinyltransferase protein	-0,36	0,79	1,32	1,77	2,33	3,26
SMb20955	exoK endo-beta-1,3-1,4-glycanase protein	-0,29	0,27	0,09	0,47	0,89	1,48
SMb20956	exoL Putative glucosyltransferase protein	-0,33	-0,24	0,07	0,19	0,57	0,83
SMb20960	exoN UDPglucose pyrophosphorylase protein	0,14	0,21	0,30	1,23	2,55	3,93
SMb20984	nirB Putative nitrite reductase [NAD(P)H], large subunit protein	0,00	-2,58	-2,66	-1,19	-0,42	-0,73
SMb20985	nirD Putative nitrite reductase [NAD(P)H], small subunit protein	-0,15	-1,60	-2,95	-1,40	-0,53	-0,73
SMb20986	narB Putative nitrate reductase, large subunit protein	-0,15	-0,46	-3,17	-1,50	-0,21	-0,28
SMb20988	Conserved Hypothetical protein	-2,85	-1,95	-1,38	-0,77	-2,44	-1,67
SMb21176	phoD phosphate uptake ABC transporter periplasmic solute-binding protein precursor	0,03	-0,08	-0,86	-1,29	-1,84	-1,26
SMb21177	phoC phosphate uptake ABC transporter ATP-binding protein	0,25	-0,83	-1,54	-1,79	-1,66	-1,13

SMb21236	Putative ATPGTP-binding protein	-0,29	-0,16	0,03	0,11	-0,11	0,07
SMb21259	Hypothetical exported protein precursor	-0,07	-0,22	0,16	0,02	-0,40	-0,57
SMb21295	Putative small heat shock protein, hsp20 family	-0,71	-0,56	-0,49	-0,34	-0,18	-0,10
SMb21440	Hypothetical protein	-0,55	-0,52	-0,17	0,22	1,01	2,41
SMb21491	Hypothetical exported protein	0,09	0,24	0,09	0,15	1,11	2,65
SMb21516	Conserved Hypothetical protein	-0,52	0,53	0,97	1,23	1,05	1,39
SMb21566	groEL5 Putative heat shock protein groEL	-0,24	-0,30	-0,19	-0,23	-0,12	0,29
SMb21690	exoW glucosyltransferase protein	-0,26	0,48	0,62	1,29	1,59	2,19
SMb21707	Putative ureashort-chain amide or branched-chain amino acid uptake ABC transporter	-0,42	-0,28	-2,33	-1,08	1,71	-0,11
SMc00043	sodB Superoxide dismutase Fe protein	0,26	0,17	0,07	0,29	0,16	0,14
SMc00045	cycF Putative Cytochrome C signal peptide protein	-0,11	0,23	1,02	1,40	1,12	1,58
SMc00062	Hypothetical protein	-0,54	0,72	1,62	1,94	2,63	3,24
SMc00063	Hypothetical transmembrane protein	-0,46	0,17	0,47	1,43	1,81	2,16
SMc00070	Conserved Hypothetical signal peptide protein	0,35	0,53	-0,06	0,95	0,49	1,27
SMc00094	betB Betaine aldehyde dehydrogenase BADH oxireductese NAD protein	0,25	0,09	-0,72	-1,03	-1,56	-1,35
SMc00103	dhe Putative Alpha-halocarboxylic acid dehalogenase protein	0,17	1,35	1,10	1,70	1,61	2,46
SMc00109	Putative transcription regulator protein	0,00	0,21	0,34	0,45	0,61	0,98
SMc00115	Conserved Hypothetical protein	0,09	0,12	0,03	0,35	0,88	1,20
SMc00159	Hypothetical signal peptide protein	-0,19	-0,10	0,18	-0,26	-0,92	-1,86
SMc00186	Putative ABC transporter ATP-binding transmembrane protein	-0,01	0,81	0,62	0,58	1,41	1,36
SMc00276	Conserved Hypothetical protein	-0,90	-0,17	0,09	-0,20	-0,03	-0,14
SMc00283	Putative transcription regulator protein	0,20	-0,51	-0,42	-0,42	-0,39	-0,40
SMc00301	sufA putative FeS assembly scaffold	-0,07	0,14	0,36	0,00	0,01	0,26
SMc00302	Conserved Hypothetical protein	-0,46	0,03	-0,31	-0,80	-0,76	-0,17
SMc00329	irr Putative Iron response regulator protein	0,00	0,21	0,41	0,32	0,44	0,56
SMc00341	Hypothetical transmembrane protein	-0,13	0,10	0,18	0,24	0,27	0,27
SMc00346	Hypothetical transmembrane protein	0,23	1,13	1,28	1,56	1,97	1,97
SMc00565	rplI Probable 50S ribosomal protein L9	0,28	0,33	0,47	0,23	-0,33	0,12
SMc00591	Hypothetical/unknown signal peptide protein	0,17	-0,36	-0,41	0,22	0,18	0,24
SMc00610	Conserved Hypothetical protein	0,27	0,51	0,56	1,01	1,01	1,52
SMc00611	lpiA Transmembrane protein	-0,30	0,41	1,69	2,40	3,11	2,68
SMc00638	Putative heat resistant agglutinin 1 protein	0,27	-0,41	-0,72	-0,63	-0,86	-1,17
SMc00641	serA Putative D-3-phosphoglycerate dehydrogenase protein	0,26	1,25	2,62	3,04	1,61	0,36
SMc00810	hypothetical protein	0,27	0,12	-0,04	0,07	0,18	0,64
SMc00827	Putative transport transmembrane protein	-0,08	0,00	0,15	0,08	0,03	-0,05
SMc00897	pmbA Hypothetical PMBA protein	0,05	0,24	0,00	0,33	0,26	0,51
SMc00914	Putative oxidoreductase protein	-0,05	0,31	0,50	0,59	0,68	0,47
SMc00949	Conserved Hypothetical protein	0,31	-0,14	-0,54	-0,07	0,30	-0,05
SMc01107	Conserved Hypothetical protein	0,21	0,44	0,47	0,66	0,36	0,48
SMc01142	grpE Probable heat shock protein	0,10	-0,33	-0,44	-0,26	-0,14	-0,04
SMc01224	trxB Probable Thioredoxin reductase protein	0,30	0,15	-0,21	-0,28	-0,26	-0,09
SMc01225	Putative transcription regulator protein	-0,12	0,65	1,19	1,32	1,41	1,79
SMc01226	Putative transcription regulator protein	-0,06	0,56	0,90	0,86	1,05	0,91
SMc01308	rplD Probable 50S ribosomal protein L4	0,38	0,61	0,67	-0,29	-1,15	-0,19
SMc01309	rplC Probable 50S ribosomal protein L3	0,44	0,57	0,85	-0,44	-0,91	-0,18
SMc01311	tufA Probable elongation factor TU protein	0,08	0,49	-0,01	-0,73	-0,42	-0,48
SMc01314	rpsL Probable 30S ribosomal protein S12	0,12	0,45	-0,13	-1,06	0,12	-0,19
SMc01326	tufB Probable elongation factor TU protein	0,14	0,59	0,11	-0,69	-0,38	-0,35
SMc01341	Hypothetical/unknown protein	0,14	0,35	0,64	0,76	1,10	1,39
SMc01365	rnr Putative Exoribonuclease II protein	-0,39	0,11	0,56	0,42	0,63	1,10
SMc01428	cspA2 Probable cold shock transcription regulator protein	0,15	0,59	0,75	0,76	1,11	1,41
SMc01440	hflC Putative hydrolase serine protease	0,29	0,27	0,11	0,29	0,18	0,24

SMc01441	hflK Putative membrane bound Protease protein	0,28	0,36	0,25	0,34	0,22	0,28
SMc01505	Anti-sigma factor	0,05	0,66	1,63	2,44	2,69	2,37
SMc01518	Conserved Hypothetical protein	0,31	0,46	0,21	0,23	0,11	-0,04
SMc01556	Conserved Hypothetical protein	-0,28	0,15	0,50	0,47	0,99	1,89
SMc01578	aatA Aspartate aminotransferase A (transaminase) protein	0,20	-0,24	-0,39	-0,26	-0,23	-0,40
SMc01580	Hypothetical transmembrane protein	-0,01	1,28	1,75	2,46	2,73	3,75
SMc01719	mcpT Probable chemoreceptor (methyl accepting chemotaxis) transmembrane protein	0,31	-0,21	-0,40	-0,26	-0,68	-0,77
SMc01764	ate Putative arginyl-tRNA protein transferase	0,01	0,42	0,76	0,90	1,04	1,57
SMc01765	Hypothetical transmembrane protein	0,17	1,49	1,92	2,15	2,44	3,26
SMc01769	Hypothetical protein	-0,19	0,52	1,45	1,52	1,87	2,01
SMc01774	Hypothetical transmembrane protein	-0,27	0,09	0,33	1,07	1,91	2,69
SMc01788	Hypothetical protein	-0,26	0,25	0,53	1,03	0,86	0,87
SMc01813	Conserved Hypothetical protein	-0,10	0,18	0,33	0,33	0,15	0,19
SMc01815	Putative oxidoreductase Iron-sulfur protein	-0,15	-0,25	-0,04	0,65	0,40	0,07
SMc01832	ureE Putative Urease accessory protein	0,19	0,03	0,48	0,38	0,34	0,26
SMc01848	Conserved Hypothetical protein	0,03	-0,01	-0,03	-0,65	-0,09	-0,02
SMc01855	Hypothetical transmembrane protein	0,25	0,96	1,56	2,15	2,29	4,20
SMc01860	ftsI Probable Penicillin-binding transmembrane protein	-0,04	-0,21	-0,01	-0,02	-0,13	0,07
SMc01904	clpX Probable ATP-dependent CLP Protease ATP-binding subunit protein	0,27	1,29	1,40	2,09	2,47	2,20
SMc01905	lon Probable ATP-dependent Protease LA protein	0,17	0,70	1,23	1,28	1,15	1,52
SMc01947	Conserved Hypothetical transmembrane protein	0,15	0,45	0,22	0,18	0,26	-0,13
SMc02052	Conserved Hypothetical protein	0,31	2,38	2,51	1,55	1,71	1,51
SMc02075	Conserved Hypothetical protein	0,17	0,73	0,97	0,99	0,83	0,60
SMc02078	exoR Exopolysaccharide biosynthesis regulatorY protein	0,34	0,52	0,16	0,39	1,34	2,21
SMc02106	Conserved Hypothetical protein	0,01	0,00	0,00	0,14	0,13	0,16
SMc02109	clpA Probable ATP-dependent CLP Protease ATP-binding subunit protein	0,11	0,42	0,85	0,75	0,64	0,68
SMc02110	ATP-dependent Clp protease adapter protein clpS1	0,07	0,48	0,68	0,81	0,67	0,89
SMc02139	Hypothetical protein	0,62	0,46	0,30	0,19	0,20	0,13
SMc02145	Hypothetical signal peptide protein	0,46	-0,66	-1,38	-0,87	-0,75	-0,53
SMc02146	Putative Phosphate-binding periplasmic protein	0,46	0,44	-0,40	-0,80	-1,25	-1,10
SMc02151	Hypothetical virulence associated protein homologue	0,05	0,32	0,15	0,21	0,00	0,38
SMc02156	Conserved Hypothetical protein	0,05	1,53	2,60	2,89	3,66	3,49
SMc02187	Putative Integrase DNA protein	0,30	0,89	1,42	1,82	1,59	1,40
SMc02202	Hypothetical protein	-0,16	0,17	1,37	1,69	1,41	1,25
SMc02275	pncA Probable Pyrazinamidase/nicotinamidase protein	-0,01	0,28	-0,04	-0,02	0,11	1,57
SMc02278	Hypothetical unknown transmembrane protein	-0,76	0,10	0,09	0,69	1,33	2,44
SMc02284	Hypothetical signal peptide protein	1,12	0,81	0,65	0,22	0,00	0,09
SMc02365	degP1 Protease precursor protein	-0,47	2,96	3,22	4,64	5,84	5,90
SMc02366	Putative transcription regulator protein	-1,32	1,64	2,59	3,39	3,26	3,87
SMc02382	Conserved Hypothetical protein	-0,25	-0,19	0,47	0,62	1,23	1,20
SMc02390	gst7 Putative Glutathione S-transferase protein	-0,17	-0,01	0,27	-0,04	0,28	0,14
SMc02392	Hypothetical protein	-0,40	-0,46	-0,46	-0,47	-0,31	-0,34
SMc02396	Probable Outer membrane protein	-0,13	0,07	-0,20	-0,77	-0,83	-0,68
SMc02400	Probable Outer membrane protein	-0,03	-0,21	-0,93	-1,95	-1,89	-1,38
SMc02403	Putative Murein transglycosylase protein	-1,27	0,07	0,52	-0,46	0,59	1,37
SMc02433	clpB ATP-dependent Protease (heat shock protein)	0,09	-0,10	0,01	0,04	0,35	0,84
SMc02435	hemK1 Putative Methyltransferase protein	0,06	0,01	0,32	0,26	0,60	0,56
SMc02443	grxC Probable Glutaredoxin 3 protein	-0,23	0,09	0,27	0,24	0,26	0,64
SMc02479	mdh Probable Malate dehydrogenase protein	0,41	0,79	-0,06	-0,48	-0,96	-0,64
SMc02480	sucC Probable Succinyl-coA synthetase beta chain protein	0,39	0,85	-0,36	-0,33	-1,05	-0,63
SMc02491	Hypothetical protein	0,14	0,31	0,84	1,02	1,26	1,76
SMc02518	Putative ATP-binding ABC transporter protein	0,06	0,65	0,30	0,75	0,11	0,14

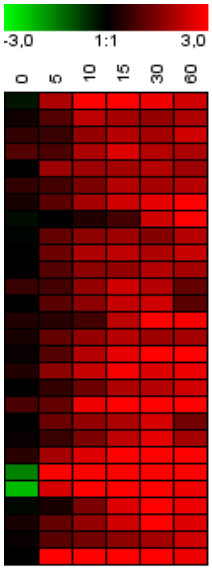
SMc02560	chvI transcriptionAL regulatorY protein	-0,03	0,85	1,14	1,67	1,90	2,91
SMc02575	hslV Probable heat shock protein	-0,18	-0,42	-0,01	-0,34	-0,02	-0,09
SMc02576	Hypothetical Acetyltransferase protein	0,18	-0,04	0,10	0,03	0,05	0,07
SMc02582	Conserved Hypothetical protein	0,34	-0,43	-1,78	-1,89	-1,15	-1,23
SMc02634	Hypothetical transmembrane protein	-0,09	-0,88	-3,16	-2,56	-2,64	-2,23
SMc02655	Hypothetical transmembrane protein	0,08	1,19	1,64	1,79	2,23	2,34
SMc02720	clpP2 CLP Protease Proteolytic subunit protein	0,06	0,12	0,27	0,20	0,21	0,21
SMc02728	fhs Probable Formate-tetrahydrofolate ligase protein	-0,32	-0,14	0,12	-0,22	0,00	-0,43
SMc02735	Hypothetical protein	-0,06	-0,06	0,29	0,02	-0,08	-0,19
SMc02755	ahcY Probable Adenosylhomocysteinase protein	0,51	0,58	0,68	0,14	-0,03	-1,00
SMc02820	cpaF1 Putative pilus assembly protein	0,17	-0,12	-0,79	-0,08	-0,60	-0,34
SMc02885	msrA1 Probable peptide Methionine sulfoxide reductase protein	0,30	0,09	0,59	0,20	0,38	0,24
SMc02898	kdsB Probable 3-deoxy-manno-octulosonate cytidylyltransferase (CMP-KDO synthetase) protein	0,22	0,29	0,14	0,08	0,09	0,25
SMc02978	Conserved Hypothetical protein	0,31	0,51	0,16	0,49	0,22	0,56
SMc02980	Hypothetical protein	0,41	0,37	0,45	0,33	-0,07	0,43
SMc03000	Putative permease ABC transporter protein	-0,17	-0,21	-0,08	-0,15	0,00	-0,28
SMc03016	Putative transcription regulator protein	-0,13	-0,31	-0,28	-0,35	0,16	-0,18
SMc03027	flgB flagellar basal-body rod protein	0,19	0,18	0,20	-0,35	-1,40	-1,97
SMc03029	fliE flagellar hook-basal body complex protein	0,00	0,30	0,68	0,05	-0,92	-1,60
SMc03030	flgG flagellar basal-body rod protein	0,23	0,41	0,33	0,01	-1,05	-1,95
SMc03037	flaA flagellin A protein	0,23	-0,52	-0,68	-0,48	-0,58	-0,71
SMc03038	flaB flagellin B protein	0,24	-0,60	-0,76	-0,61	-0,90	-1,21
SMc03039	flaD Probable flagellin D protein	0,21	-0,52	-0,98	-0,76	-0,94	-1,29
SMc03046	Putative transcription regulator protein	0,31	-0,22	-0,26	-0,51	-0,14	-0,50
SMc03047	flgE flagellar hook protein	0,01	-0,21	0,13	-0,33	-0,43	-0,82
SMc03049	flgL Putative flagellar hook-associated protein	0,02	0,26	0,21	-0,47	-1,98	-4,11
SMc03050	flaF Putative flagellin synthesis regulator protein	0,04	0,26	0,10	-0,58	-1,11	-0,62
SMc03064	aglA Probable alpha-glucosidase protein	-0,14	-0,06	-0,18	-0,60	-0,38	-0,58
SMc03090	cheW3 Putative chemotaxis protein	0,26	-0,04	-0,30	-0,11	-0,10	-0,28
SMc03105	dxr Probable 1-deoxy-D-xylulose 5-phosphate reductoisomerase protein	-0,24	0,04	-0,24	-0,13	-0,05	0,31
SMc03111	pmi Mannose-6-phosphate isomerase protein	0,30	0,57	0,45	0,28	0,33	0,35
SMc03138	Putative Sugar kinase protein	-0,75	-0,56	-0,42	-1,41	-1,08	-0,25
SMc03151	Conserved Hypothetical protein	0,11	0,36	0,09	0,18	0,22	0,23
SMc03152	Hypothetical transmembrane protein	0,12	0,20	0,01	0,20	0,36	0,45
SMc03205	purU1 Putative Formyltetrahydrofolate deformylase protein	0,06	-0,09	-0,18	-0,19	-0,12	-0,18
SMc03245	Putative Amidase protein	0,46	0,12	0,17	0,47	0,54	0,19
SMc03290	Hypothetical protein	0,79	0,58	1,45	0,58	-0,43	-0,93
SMc03773	Conserved Hypothetical protein	0,23	0,44	0,04	-0,36	0,61	0,80
SMc03780	Hypothetical protein	-0,35	0,04	0,63	1,04	1,03	0,90
SMc03783	ctpA Putative Carboxy-terminal processing protease precursor signal peptide protein	-0,04	0,79	0,97	1,56	1,60	1,78
SMc03784	Hypothetical transmembrane protein	-0,33	-0,10	0,01	0,04	0,19	0,14
SMc03788	dnaE2 Putative DNA Polymerase III Alpha chain protein	0,40	0,77	-0,20	-1,68	1,05	-0,26
SMc03802	Conserved Hypothetical protein	-0,13	0,13	0,21	0,20	0,59	0,58
SMc03806	glnK Probable Nitrogen regulatorY protein PII 2	0,26	-0,80	-1,27	-0,66	-0,09	-0,59
SMc03819	Conserved Hypothetical protein	-0,47	-0,27	0,01	-0,07	-0,08	-0,21
SMc03836	tesA Putative Acyl-coA Thioesterase I protein	0,03	0,09	0,23	0,42	0,79	0,87
SMc03838	Hypothetical transmembrane protein	0,36	0,22	0,27	0,19	0,21	0,19
SMc03893	Putative Amino-acid transport system permease ABC transporter protein	0,21	0,50	0,42	0,50	0,60	0,65
SMc03900	ndvA Beta-1-->2Glucan export ATP-binding protein	0,34	1,04	1,02	1,56	0,90	1,38
SMc03999	Hypothetical protein	-0,19	0,19	0,94	1,68	1,40	2,32
SMc04009	Conserved Hypothetical protein	0,06	-0,34	-1,01	-1,29	-1,20	-1,12

SMc04026	gltD Probable glutamate synthase small chain protein	0,11	0,31	0,05	-0,20	-0,18	-0,56
SMc04040	ibpA Heat shock protein	-0,04	-0,26	-0,22	-0,19	0,01	0,21
SMc04059	Hypothetical protein	0,00	0,06	0,09	-0,17	-0,74	-1,02
SMc04091	htpX Putative Protease transmembrane protein	0,13	0,10	0,49	0,59	0,75	0,38
SMc04111	cpaC1 Putative pilus assembly transmembrane protein	-0,10	-0,28	-0,35	-0,34	-0,51	-0,91
SMc04128	Putative heavy metal transporting ATPase protein	0,17	0,61	1,38	2,64	3,03	3,37
SMc04167	Putative HIistidine-rich transporter transmembrane protein	0,31	-0,06	0,16	0,64	1,83	2,24
SMc04213	dgkA Diacylglycerol kinase protein	-0,11	0,09	-0,12	0,04	0,31	0,72
SMc04246	Hypothetical transmembrane signal peptide protein	-0,29	-0,05	0,62	0,73	1,38	1,44
SMc04267	lpsS LPS sulfotransferase	0,19	0,90	1,14	1,38	1,70	2,70
SMc04346	ilvC ketol-acid reductoisomerase protein	0,20	0,12	-0,24	-0,33	-0,38	-0,55
SMc04350	Putative multidrug efflux system transmembrane protein	0,01	2,14	2,97	3,62	4,00	3,62
SMc04459	ftsH Probable metalloprotease transmembrane protein	0,22	0,55	0,84	1,16	1,83	2,13
SMc04461	TolB protein precursor	0,11	0,66	0,76	1,13	1,81	2,14
SMc04865	Hypothetical protein	-0,13	0,18	0,05	0,38	0,14	0,41

3. HEAT MAPS OF *S. MELILOTI 1021* CLUSTERS A TO F

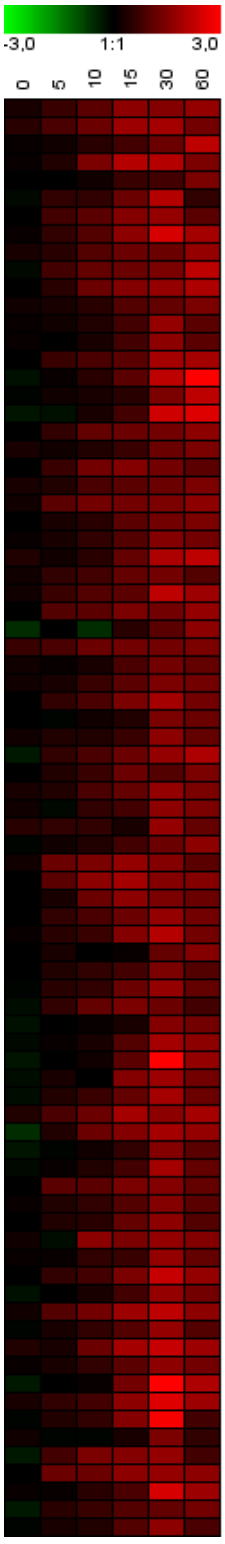
The transcriptional data obtained by microarray analysis of the *S. meliloti* 1021 pH shock experiment were grouped into six K-means clusters (A-F). Each column of the heat map represents one time point of the time-course experiment, after shift from pH 7.0 to pH 5.75, in the following order: 0, 5, 10, 15, 30 and 60 minutes. The color intensity on the heat map correlates to the intensity (log ratio) of the expression of each gene at the specified time point, with red representing overexpression and green indicating reduced expression.

CLUSTER A



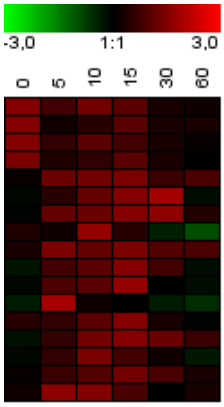
SMa0045	Cah Probable carbonic anhydrase
SMa1077	Nex18 Symbiotically induced conserved protein
SMa1078	Conserved Hypothetical protein
SMa1151	Conserved Hypothetical protein
SMb20359	Hypothetical protein
SMb20949	exoV Putative pyruvyltransferase protein
SMb20954	exoH succinyltransferase protein
SMb20960	exoN UDPglucose pyrophosphorylase protein
SMb21516	Conserved Hypothetical protein
SMc00045	cycF Putative Cytochrome C signal peptide protein
SMc00062	Hypothetical protein
SMc00186	Putative ABC transporter ATP-binding transmembrane protein
SMc00611	IpiA Transmembrane protein
SMc00949	Conserved Hypothetical protein
SMc01225	Putative transcription regulator protein
SMc01580	Hypothetical transmembrane protein
SMc01765	Hypothetical transmembrane protein
SMc01774	Hypothetical transmembrane protein
SMc01855	Hypothetical transmembrane protein
SMc01904	clpX Probable ATP-dependent CLP Protease ATP-binding subunit protein
SMc01905	lon Probable ATP-dependent Protease LA protein
SMc02156	Conserved Hypothetical protein
SMc02365	degP1 Protease precursor protein
SMc02366	Putative transcription regulator protein
SMc02382	Conserved Hypothetical protein
SMc03780	Hypothetical protein
SMc03999	Hypothetical protein
SMc04350	Putative multidrug efflux system transmembrane protein

CLUSTER B



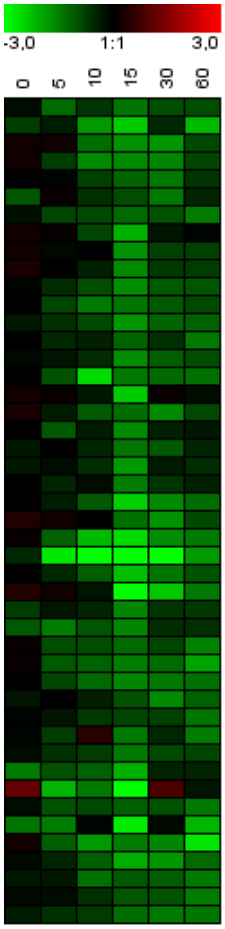
SMa0172	Conserved Hypothetical protein
SMa0473	Conserved Hypothetical protein
SMa0994	Hypothetical protein
SMa1079	TspO Tryptophan rich sensory protein homologue
SMb20486	Putative sugar ABC transporter permease protein
SMb20560	Conserved Hypothetical protein
SMb20696	Hypothetical protein
SMb20707	cyaG2 Putative adenylate cyclase protein
SMb20944	exoQ Putative polysaccharide polymerase, similar to Wzy protein
SMb20946	exoY galactosyltransferase protein
SMb20955	exoK endo-beta-1,3-1,4-glycanase protein
SMb20956	exoL Putative glucosyltransferase protein
SMb21236	Putative ATPGTP-binding protein
SMb21259	Hypothetical exported protein precursor
SMb21295	Putative small heat shock protein, hsp20 family
SMb21440	Hypothetical protein
SMb21491	Hypothetical exported protein
SMb21566	groEL5 Putative heat shock protein groEL
SMb21690	exoW glucosyltransferase protein
SMc00043	sodB Superoxide dismutase Fe protein
SMc00063	Hypothetical transmembrane protein
SMc00070	Conserved Hypothetical signal peptide protein
SMc00103	dhe Putative Alpha-halocarboxylic acid dehalogenase protein
SMc00115	Conserved Hypothetical protein
SMc00301	Conserved Hypothetical protein
SMc00302	Conserved Hypothetical protein
SMc00329	irr Putative Iron response regulator protein
SMc00341	Hypothetical transmembrane protein
SMc00346	Hypothetical transmembrane protein
SMc00591	Hypothetical/unknown signal peptide protein
SMc00610	Conserved Hypothetical protein
SMc00897	pmbA Hypothetical PMBA protein
SMc00914	Putative oxidoreductase protein
SMc01107	Conserved Hypothetical protein
SMc01142	grpE Probable heat shock protein
SMc01224	trxB Probable Thioredoxin reductase protein
SMc01341	Hypothetical/unknown protein
SMc01365	rnr Putative Exoribonuclease II protein
SMc01428	cspA2 Probable cold shock transcription regulator protein
SMc01440	hflC Putative hydrolase serine protease transmembrane protein
SMc01441	hflK Putative membrane bound Protease protein
SMc01556	Conserved Hypothetical protein
SMc01764	Hypothetical protein
SMc01769	Hypothetical protein
SMc01788	Hypothetical protein
SMc01813	Conserved Hypothetical protein
SMc02075	Conserved Hypothetical protein
SMc02078	exoR Exopolysaccharide biosynthesis regulatorY protein
SMc02109	clpA Probable ATP-dependent CLP Protease ATP-binding subunit protein
SMc02110	Conserved Hypothetical protein
SMc02187	Putative Integrase DNA protein
SMc02278	Hypothetical unknown transmembrane protein
SMc02390	gst7 Putative Glutathione S-transferase protein
SMc02433	clpB Probable ATP-dependent Protease (heat shock protein)
SMc02435	hemK1 Putative Methyltransferase protein
SMc02443	grxC Probable Glutaredoxin 3 protein
SMc02491	Hypothetical protein
SMc02560	chvI transcriptionAL regulatorY protein
SMc02575	hslV Probable heat shock protein
SMc02576	Hypothetical Acetyltransferase protein
SMc02655	Hypothetical transmembrane protein
SMc02720	clpP2 CLP Protease Proteolytic subunit protein
SMc02735	Hypothetical protein
SMc02885	msrA1 Probable peptide Methionine sulfoxide reductase protein
SMc02978	Conserved Hypothetical protein
SMc03151	Conserved Hypothetical protein
SMc03152	Hypothetical transmembrane protein
SMc03783	ctpA Putative Carboxy-terminal processing protease precursor signal peptide protein
SMc03802	Conserved Hypothetical protein
SMc03836	tesA Putative Acyl-coA Thioesterase I protein
SMc03838	Hypothetical transmembrane protein
SMc04040	ibpA Heat shock protein
SMc04091	htpX Putative Protease transmembrane protein
SMc04128	Putative heavy metal transporting ATPase protein
SMc04167	Putative Hlistidine-rich transporter transmembrane protein
SMc04246	Hypothetical transmembrane signal peptide protein
SMc04267	Hypothetical protein
SMc04459	ftsH Probable metalloprotease transmembrane protein
SMc04461	TolB protein precursor
SMc04865	Hypothetical protein

CLUSTER C



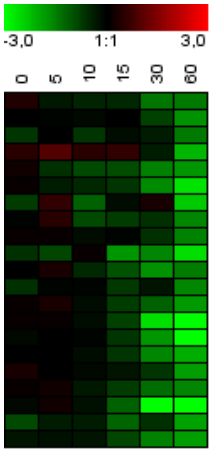
SMa1082	Hypothetical protein
SMa1182	NosZ N2O reductase
SMa1220	FixN1 Heme b / copper cytochrome c oxidase subunit
SMa1243	Azu1 pseudoazurin (blue copper protein)
SMa1961	Putative Polyhydrixyalkanoate depolymerase
SMb20611	dctA C4-dicarboxylate transport protein
SMc00109	Putative transcription regulator protein
SMc00641	serA Putative D-3-phosphoglycerate dehydrogenase protein
SMc01226	Putative transcription regulator protein
SMc01505	Hypothetical protein
SMc01815	Putative oxidoreductase Iron-sulfur protein
SMc02052	Conserved Hypothetical protein
SMc02106	Conserved Hypothetical protein
SMc02202	Hypothetical protein
SMc02518	Putative ATP-binding ABC transporter protein
SMc03893	Putative Amino-acid transport system permease ABC transporter protein
SMc03900	ndvA Beta-1-->2Glucan export ATP-binding protein

CLUSTER D



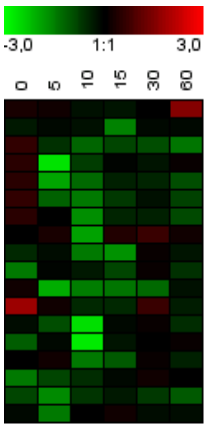
SMb20497	lyx Putative L-xylulose kinase protein
SMb20988	Conserved Hypothetical protein
SMb21176	phoD phosphate uptake ABC transporter periplasmic solute-binding protein precursor
SMb21177	phoC phosphate uptake ABC transporter ATP-binding protein
SMc00094	betB Betaine aldehyde dehydrogenase BADH oxireductese NAD protein
SMc00276	Conserved Hypothetical protein
SMc00283	Putative transcription regulator protein
SMc00565	rplI Probable 50S ribosomal protein L9
SMc01308	rplD Probable 50S ribosomal protein L4
SMc01309	rplC Probable 50S ribosomal protein L3
SMc01311	tufA Probable elongation factor TU protein
SMc01314	rpsL Probable 30S ribosomal protein S12
SMc01326	tufB Probable elongation factor TU protein
SMc01518	Conserved Hypothetical protein
SMc01578	aatA Aspartate aminotransferase A (transaminase) protein
SMc01848	Conserved Hypothetical protein
SMc02139	Hypothetical protein
SMc02146	Putative Phosphate-binding periplasmic protein
SMc02151	Hypothetical virulence associated protein homologue
SMc02284	Hypothetical signal peptide protein
SMc02392	Hypothetical protein
SMc02396	Probable Outer membrane protein
SMc02400	Probable Outer membrane protein
SMc02480	sucC Probable Succinyl-coA synthetase beta chain protein
SMc02582	Conserved Hypothetical protein
SMc02634	Hypothetical transmembrane protein
SMc02820	cpaF1 Putative pilus assembly protein
SMc02898	kdsB Probable 3-deoxy-manno-octulosonate cytidylyltransferase (CMP-KDO synthetase)
SMc03000	Putative permease ABC transporter protein
SMc03016	Putative transcription regulator protein
SMc03037	flaA flagellin A protein
SMc03038	flaB flagellin B protein
SMc03039	flaD Probable flagellin D protein
SMc03050	flaF Putative flagellin synthesis regulator protein
SMc03090	cheW3 Putative chemotaxis protein
SMc03111	pmi Mannose-6-phosphate isomerase protein
SMc03205	purU1 Putative Formyltetrahydrofolate deformylase protein
SMc03245	Putative Amidase protein
SMc03290	Hypothetical protein
SMc03773	Conserved Hypothetical protein
SMc03788	dnaE2 Putative DNA Polymerase III Alpha chain protein
SMc03806	glnK Probable Nitrogen regulatorY protein PII 2
SMc04009	Conserved Hypothetical protein
SMc04026	gltD Probable glutamate synthase small chain protein
SMc04111	cpaC1 Putative pilus assembly transmembrane protein
SMc04346	ilvC ketol-acid reductoisomerase protein

CLUSTER E



SMa0392	ABC transporter, periplasmic solute-binding protein
SMb20325	ThuE probable trehalosemaltose-binding protein
SMb20777	TRm19 Putative transposase of insertion sequence ISRm19 protein
SMc00159	Hypothetical signal peptide protein
SMc00638	Putative heat resistant agglutinin 1 protein
SMc01719	mcpT Probable chemoreceptor (methyl accepting chemotaxis) transmembrane protein
SMc01832	ureE Putative Urease accessory protein
SMc01860	ftsI Probable Penicillin-binding transmembrane protein
SMc01947	Conserved Hypothetical transmembrane protein
SMc02275	pncA Probable Pyrazinamidase/nicotinamidase protein
SMc02479	mdh Probable Malate dehydrogenase protein
SMc02728	fhs Probable Formate-tetrahydrofolate ligase protein
SMc02755	ahcY Probable Adenosylhomocysteinase protein
SMc03027	flgB flagellar basal-body rod protein
SMc03029	fliE flagellar hook-basal body complex protein
SMc03030	flgG flagellar basal-body rod protein
SMc03046	Putative transcription regulator protein
SMc03047	flgE flagellar hook protein
SMc03049	flgL Putative flagellar hook-associated protein
SMc03138	Putative Sugar kinase protein
SMc04059	Hypothetical protein

CLUSTER F

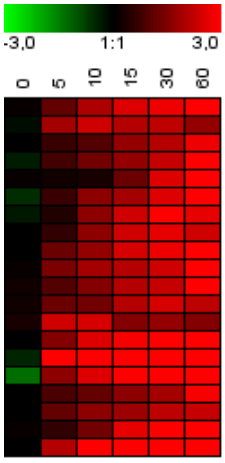


SMa1750	Hypothetical protein
SMa1898	Hypothetical protein
SMb20436	Putative nitrate transporter protein
SMb20605	Putative ureashort-chain amide or branched-chain amino acid uptake ABC transporter
SMb20984	nirB Putative nitrite reductase [NAD(P)H], large subunit protein
SMb20985	nirD Putative nitrite reductase [NAD(P)H], small subunit protein
SMb20986	narB Putative nitrate reductase, large subunit protein
SMb21707	Putative ureashort-chain amide or branched-chain amino acid uptake ABC transporter
SMc00810	hypothetical protein
SMc00827	Putative transport transmembrane protein
SMc02145	Hypothetical signal peptide protein
SMc02403	Putative Murein transglycosylase protein
SMc02980	Hypothetical protein
SMc03064	aglA Probable alpha-glucosidase protein
SMc03105	dxr Probable 1-deoxy-D-xylulose 5-phosphate reductoisomerase protein
SMc03784	Hypothetical transmembrane protein
SMc03819	Conserved Hypothetical protein
SMc04213	dgkA Diacylglycerol kinase protein

4. HEAT MAPS OF *S. MELILOTI RPOH1* MUTANT CLUSTERS G TO L

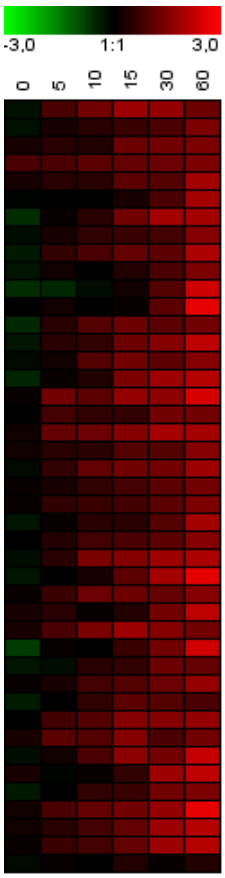
The transcriptional data obtained by microarray analysis of the *S. meliloti rpoH1* mutant following acidic pH shift was analyzed taking into consideration the 210 genes that were also analysed in the wild type experiments. The *rpoH1* mutant microarray data were also grouped into six K-means clusters (G-L). Each column of the heat map represents one time point after shift from pH 7.0 to pH 5.75 of the time-course experiment, in the following order: 0, 5, 10, 15, 30 and 60 minutes. The color intensity on the heat map correlates to the intensity (log ratio) of the expression, with red representing overexpression and green indicating reduced expression.

CLUSTER G



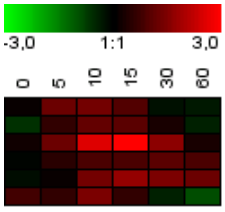
SMa1077	Nex18 Symbiotically induced conserved protein
SMb20560	Conserved Hypothetical protein
SMb20949	exoV Putative pyruvyltransferase protein
SMb20954	exoH succinyltransferase protein
SMb20960	exoN UDPglucose pyrophosphorylase protein
SMc00062	Hypothetical protein
SMc00611	IpiA Transmembrane protein
SMc01505	Hypothetical protein
SMc01580	Hypothetical transmembrane protein
SMc01765	Hypothetical transmembrane protein
SMc01855	Hypothetical transmembrane protein
SMc01904	clpX Probable ATP-dependent CLP Protease ATP-binding subunit protein
SMc02052	Conserved Hypothetical protein
SMc02156	Conserved Hypothetical protein
SMc02365	degP1 Protease precursor protein
SMc02366	Putative transcription regulator protein
SMc02560	chvI transcriptionAL regulatorY protein
SMc02655	Hypothetical transmembrane protein
SMc04128	Putative heavy metal transporting ATPase protein
SMc04350	Putative multidrug efflux system transmembrane protein

CLUSTER H



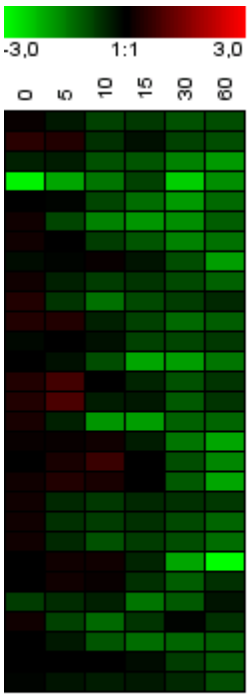
SMa0045	Cah Probable carbonic anhydrase
SMa0172	Conserved Hypothetical protein
SMa1078	Conserved Hypothetical protein
SMa1151	Conserved Hypothetical protein
SMb20359	Hypothetical protein
SMb20486	Putative sugar ABC transporter permease protein
SMb20611	dctA C4-dicarboxylate transport protein
SMb20944	exoQ Putative polysaccharide polymerase, similar to Wzy protein
SMb20946	exoY galactosyltransferase protein
SMb20955	exoK endo-beta-1,3-1,4-glycanase protein
SMb21440	Hypothetical protein
SMb21491	Hypothetical exported protein
SMb21516	Conserved Hypothetical protein
SMb21690	exoW glucosyltransferase protein
SMc00045	cycF Putative Cytochrome C signal peptide protein
SMc00063	Hypothetical transmembrane protein
SMc00103	dhe Putative Alpha-halocarboxylic acid dehalogenase protein
SMc00186	Putative ABC transporter ATP-binding transmembrane protein
SMc00346	Hypothetical transmembrane protein
SMc00610	Conserved Hypothetical protein
SMc01225	Putative transcription regulator protein
SMc01341	Hypothetical/unknown protein
SMc01428	cspA2 Probable cold shock transcription regulator protein
SMc01556	Conserved Hypothetical protein
SMc01764	ate Putative arginyl-tRNA protein transferase
SMc01769	Hypothetical protein
SMc01774	Hypothetical transmembrane protein
SMc01905	lon Probable ATP-dependent Protease LA protein
SMc02078	exoR Exopolysaccharide biosynthesis regulatorY protein
SMc02187	Putative Integrase DNA protein
SMc02278	Hypothetical unknown transmembrane protein
SMc02382	Conserved Hypothetical protein
SMc02491	Hypothetical protein
SMc03780	Hypothetical protein
SMc03783	ctpA Putative Carboxy-terminal processing protease precursor signal peptide
SMc03900	ndvA Beta-1-->2Glucan export ATP-binding protein
SMc03999	Hypothetical protein
SMc04167	Putative Hlistidine-rich transporter transmembrane protein
SMc04246	Hypothetical transmembrane signal peptide protein
SMc04267	lpsS LPS sulfotransferase
SMc04459	ftsH Probable metalloprotease transmembrane protein
SMc04461	TolB protein precursor
SMc04865	Hypothetical protein

CLUSTER I



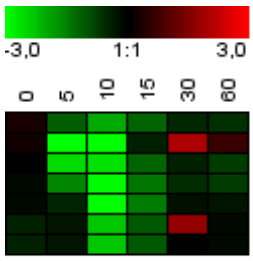
SMa1220	FixN1 Heme b / copper cytochrome c oxidase subunit
SMa1243	Azu1 pseudoazurin (blue copper protein)
SMc00641	serA Putative D-3-phosphoglycerate dehydrogenase protein
SMc01226	Putative transcription regulator protein
SMc02202	Hypothetical protein
SMc03290	Hypothetical protein

CLUSTER J



SMa0392	ABC transporter, periplasmic solute-binding protein
SMb20325	ThuE probable trehalosemaltose-binding protein
SMb20497	lyx Putative L-xylulose kinase protein
SMb20988	Conserved Hypothetical protein
SMb21176	phoD phosphate uptake ABC transporter periplasmic solute-binding protein
SMb21177	phoC phosphate uptake ABC transporter ATP-binding protein
SMc00094	betB Betaine aldehyde dehydrogenase BADH oxireductese NAD protein
SMc00159	Hypothetical signal peptide protein
SMc00638	Putative heat resistant agglutinin 1 protein
SMc02145	Hypothetical signal peptide protein
SMc02146	Putative Phosphate-binding periplasmic protein
SMc02396	Probable Outer membrane protein
SMc02400	Probable Outer membrane protein
SMc02479	mdh Probable Malate dehydrogenase protein
SMc02480	sucC Probable Succinyl-coA synthetase beta chain protein
SMc02582	Conserved Hypothetical protein
SMc03027	flgB flagellar basal-body rod protein
SMc03029	fliE flagellar hook-basal body complex protein
SMc03030	flgG flagellar basal-body rod protein
SMc03037	flaA flagellin A protein
SMc03038	flaB flagellin B protein
SMc03039	flaD Probable flagellin D protein
SMc03049	flgL Putative flagellar hook-associated protein
SMc03050	flaF Putative flagellin synthesis regulator protein
SMc03138	Putative Sugar kinase protein
SMc03806	glnK Probable Nitrogen regulatorY protein PII 2
SMc04009	Conserved Hypothetical protein
SMc04059	Hypothetical protein
SMc04111	cpaC1 Putative pilus assembly transmembrane protein

CLUSTER K



SMb20436	Putative nitrate transporter protein
SMb20605	Putative ureashort-chain amide or branched-chain amino acid uptake ABC transporter
SMb20984	nirB Putative nitrite reductase [NAD(P)H], large subunit protein
SMb20985	nirD Putative nitrite reductase [NAD(P)H], small subunit protein
SMb20986	narB Putative nitrate reductase, large subunit protein
SMb21707	Putative ureashort-chain amide or branched-chain amino acid uptake ABC transporter
SMc02634	Hypothetical transmembrane protein

CLUSTER L

SMa0473	Conserved Hypothetical protein
SMa0994	Hypothetical protein
SMa1079	TspO Tryptophan rich sensory protein homologue
SMa1082	Hypothetical protein
SMa1182	NosZ N2O reductase
SMa1750	Hypothetical protein
SMa1898	Hypothetical protein
SMa1961	Putative Polyhydrixyalkanoate depolymerase
SMb20696	Hypothetical protein
SMb20707	cyaG2 Putative adenylate cyclase protein
SMb20777	TRm19 Putative transposase of insertion sequence ISRm19 protein
SMb20956	exoL Putative glucosyltransferase protein
SMb21236	Putative ATPGTP-binding protein
SMb21259	Hypothetical exported protein precursor
SMb21295	Putative small heat shock protein, hsp20 family
SMb21566	groEL5 Putative heat shock protein groEL
SMc00043	sodB Superoxide dismutase Fe protein
SMc00070	Conserved Hypothetical signal peptide protein
SMc00115	Conserved Hypothetical protein
SMc00276	Conserved Hypothetical protein
SMc00283	Putative transcription regulator protein
SMc00301	sufA Putative FeS assembly scaffold
SMc00302	Conserved Hypothetical protein
SMc00329	irr Putative Iron response regulator protein
SMc00341	Hypothetical transmembrane protein
SMc00565	rplI Probable 50S ribosomal protein L9
SMc00591	Hypothetical/unknown signal peptide protein
SMc00810	hypothetical protein
SMc00827	Putative transport transmembrane protein
SMc00897	pmbA Hypothetical PMBA protein
SMc00914	Putative oxidoreductase protein
SMc00949	Conserved Hypothetical protein
SMc01107	Conserved Hypothetical protein
SMc01142	grpE Probable heat shock protein
SMc01224	trxB Probable Thioredoxin reductase protein
SMc01309	rplC Probable 50S ribosomal protein L3
SMc01311	tufA Probable elongation factor TU protein
SMc01326	tufB Probable elongation factor TU protein
SMc01365	rnr Putative Exoribonuclease II protein
SMc01440	hflC Putative hydrolase serine protease transmembrane protein
SMc01441	hflK Putative membrane bound Protease protein
SMc01518	Conserved Hypothetical protein
SMc01578	aatA Aspartate aminotransferase A (transaminase) protein
SMc01719	mcpT chemoreceptor transmembrane protein
SMc01788	Hypothetical protein
SMc01813	Conserved Hypothetical protein
SMc01815	Putative oxidoreductase Iron-sulfur protein
SMc01832	ureE Putative Urease accessory protein
SMc01848	Conserved Hypothetical protein
SMc01860	ftsI Probable Penicillin-binding transmembrane protein
SMc01947	Conserved Hypothetical transmembrane protein
SMc02075	Conserved Hypothetical protein
SMc02106	Conserved Hypothetical protein
SMc02109	clpA Probable ATP-dependent CLP Protease ATP-binding subunit protein
SMc02110	ATP-dependent Clp protease adapter protein clpS1
SMc02139	Hypothetical protein
SMc02151	Hypothetical virulence associated protein homologue
SMc02275	pncA Probable Pyrazinamidase/nicotinamidase protein
SMc02284	Hypothetical signal peptide protein
SMc02390	gst7 Putative Glutathione S-transferase protein
SMc02392	Hypothetical protein
SMc02403	Putative Murein transglycosylase protein
SMc02433	clpB Probable ATP-dependent Protease (heat shock protein)
SMc02435	hemK1 Putative Methyltransferase protein
SMc02443	grxC Probable Glutaredoxin 3 protein
SMc02518	Putative ATP-binding ABC transporter protein
SMc02575	hslV Probable heat shock protein
SMc02576	Hypothetical Acetyltransferase protein
SMc02720	clpP2 CLP Protease Proteolytic subunit protein
SMc02728	fhs Probable Formate-tetrahydrofolate ligase protein
SMc02735	Hypothetical protein
SMc02755	ahcY Probable Adenosylhomocysteinase protein
SMc02820	cpaF1 Putative pilus assembly protein
SMc02885	msrA1 Probable peptide Methionine sulfoxide reductase protein
SMc02898	kdsB Probable 3-deoxy-manno-octulosonate cytidylyltransferase (CMP-KDO synthetase) p
SMc02978	Conserved Hypothetical protein
SMc02980	Hypothetical protein
SMc03000	Putative permease ABC transporter protein
SMc03016	Putative transcription regulator protein
SMc03046	Putative transcription regulator protein
SMc03047	flgE flagellar hook protein
SMc03064	aglA Probable alpha-glucosidase protein
SMc03090	cheW3 Putative chemotaxis protein
SMc03105	dxr Probable 1-deoxy-D-xylulose 5-phosphate reductoisomerase protein
SMc03111	pmi Mannose-6-phosphate isomerase protein
SMc03151	Conserved Hypothetical protein
SMc03152	Hypothetical transmembrane protein
SMc03205	purU1 Putative Formyltetrahydrofolate deformylase protein
SMc03245	Putative Amidase protein
SMc03773	Conserved Hypothetical protein
SMc03784	Hypothetical transmembrane protein
SMc03802	Conserved Hypothetical protein
SMc03819	Conserved Hypothetical protein
SMc03836	tesA Putative Acyl-coA Thioesterase I protein
SMc03838	Hypothetical transmembrane protein
SMc03893	Putative Amino-acid transport system permease ABC transporter protein
SMc04026	gltD Probable glutamate synthase small chain protein
SMc04040	ibpA Heat shock protein
SMc04091	htpX Putative Protease transmembrane protein
SMc04213	dgkA Diacylglycerol kinase protein
SMc04346	ilvC ketol-acid reductoisomerase protein
SMc04865	Hypothetical protein

5. LIST OF FIGURES

6. LIST OF ABBREVIATIONS

% (v/v)	Volume percent
% (w/v)	Weight percent
× g	Times gravity
°C	Degrees Celsius
A	Adenine
Bp	Base pair
BPB	Bromphenol blue
BSA	Bovine serum albumin
C	Cytosine
CAS	Chromoazurol assay
cDNA	Complementary DNA
Cy3	Cyanine 3
Cy5	Cyanine 5
dATP	2'-deoxyadenosine 5'-triphosphate
dCTP	2'-deoxycytidine 5'-triphosphate
dGTP	2'-deoxyguanosine 5'-triphosphate
DNA	Deoxyribonucleic acid
DNAse	Deoxyribonuclease
dNTPs	Deoxynucleotide triphosphates
DTT	Dithiothreitol
dTTP	2'-deoxythymidine 5'-triphosphate
EDTA	Ethylenediaminetetraacetic acid
et al.	Et alii (lat.)/and others
EtOH	Ethanol
Fix^-	Non-nitrogen fixing phenotype
Fix^+	Nitrogen fixing phenotype
g	Grams
G	Guanine
GC	Sum of all the bases G and C
Gm	Gentamycin
h	Hour
H_2O	Water

H_2O_2	Hydrogen peroxide
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
Hsp	Heat shock protein
Kb	Kilobase
Km	Kanamycin
mg	Milligram
min	Minute
ml	Milliliter
mM	Millimolar
mRNA	Messenger RNA
nm	nanometer
Nm	Neomycin
Nod^-	Non-nodulating phenotype
Nod^+	Nodulating phenotype
O.D.	Optical density
PCR	Polymerase chain reaction
pH	Logarithm of reciprocal of hydrogen-ion concentration in gram atoms per liter
pmol	Picomol
RNA	Ribonucleic acid
RNAP	RNA polymerase
RNase	Ribonuclease
rpm	Rotations per minute
SDS	Sodium dodecyl sulfate
sec	Second
Sm	Streptomycin
T	Thymine
U	Unit
UV	Ultra-Violet
V	Volt
VMM	Vincent minimal medium
X-Gal	Bromo-chloro-indolyl-galactopyranoside
µg	Microgram
µl	Microliter
µM	Micromolar

7. TRANSCRIPTION PROFILES OF INDIVIDUAL GENES FOR *S. MELILOTI* WILD TYPE AND *RPOH1* MUTANT

Profiles for the expression of each individual gene in the microarrays throughout the pH shift time course were produced. For each gene, expression values for both for *S. meliloti* wild type and *rpoH1* mutant were plotted in the same graphic. The data can be found in the enclosed CD.

8. ACKNOWLEDGEMENTS

I would like to thank all in the Institute of Genetics, the Center for Biotechnology and the NRW Graduate School of Bioinformatics and Genome Research at Bielefeld University, who have made it possible for me to accomplish this work.

I am immensely grateful to the Prof. Dr. Alfred Pühler, for providing me with the opportunity to carry out my PhD work at the Chair of Genetics and for all the challenging and insightful discussions on the scientific research. I am likewise enormously grateful to Dr. Stefan Weidner for his supervision and guidance on all experiments and also on the writing of the paper and thesis, for his ideas and encouraging advice.

I would like to thank all of Dr. Stefan Weidner's and Dr. Andreas Schlüter's group members for all the support and company and for creating such a nice work environment. My most sincere gratitude and appreciation to Victoria Goedde for the outstanding and invaluable help in the laboratory work. A special thanks to Irene Krahn for all the delightful talks. I thank Eva Schulte-Berndt and Manuela Meyer for the skillful help with the microarray hybridizations and Epameinondas Fritzilas for the bioinformatics support.

A heartfelt thanks to my friends of the graduate school, the international Baptist church and in the international Bible study group. You were my family during my time in Bielefeld and your unyielding support meant a lot to me. I love and value you all greatly. I dedicate this thesis to my parents and sisters, who are my foundation and source of love and motivation.

"To him that is able to keep you from falling, and to present you faultless before the presence of his glory with exceeding joy, to the only God our savior, be glory and majesty, dominion and power, before all time and now and forever." Judas 1:24-25

9. STATEMENT

I hereby testify that I wrote this work myself and used no other source than the literature appropriately cited in the reference section.

Bielefeld, July 2011

Daniella Karine Cavalcanti de Lucena