UNIVERSIDAD DE GRANADA

Programa de Doctorado en Biología Fundamental y de Sistemas

CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

Insights into the function and activity mechanism of stress-induced small non-coding RNAs and the endoribonuclease YbeY in the legume symbiont *Sinorhizobium meliloti*

ALEXANDRA PEREGRINA LAVÍN

TESIS DOCTORAL

2017

Editor: Universidad de Granada. Tesis Doctorales
Autora: Alexandra Peregrina Lavín
ISBN: 978-84-9163-456-0
URI: http://hdl.handle.net/10481/48044

Insights into the function and activity mechanism of stress-induced small non-coding RNAs and the endoribonuclease YbeY in the legume symbiont *Sinorhizobium meliloti*

Memoria que presenta la licenciada en Biología

Dña. Alexandra Peregrina Lavín

como aspirante al grado de Doctor

Fdo: **Alexandra Peregrina Lavín**

VºBº del Director de la Tesis Doctoral

Fdo: Dr. José Ignacio Jiménez Zurdo

Científico Titular del CSIC

Universidad de Granada

2017

El doctorando / *The doctoral candidate* **[Alexandra Peregrina Lavín]** y los directores de la tesis / *and the thesis supervisor/s*: **[José Ignacio Jiménez Zurdo]**:

Garantizamos, al firmar esta tesis doctoral, que el trabajo ha sido realizado por el doctorando bajo la dirección de los directores de la tesis y hasta donde nuestro conocimiento alcanza, en la realización del trabajo, se han respetado los derechos de otros autores a ser citados, cuando se han utilizado sus resultados o publicaciones.

Guarantee, by signing this doctoral thesis, that the work has been done by the doctoral candidate under the direction of the thesis supervisor/s and, as far as our knowledge reaches, in the performance of the work, the rights of other authors to be cited (when their results or publications have been used) have been respected.

Lugar y fecha / *Place and date*: Granada, 12 de Junio de 2017

Director/es de la Tesis / *Thesis supervisor/s*: Doctorando / *Doctoral candidate*:

Firma / *Signed* Firma / *Signed*

Este trabajo de Tesis Doctoral ha sido realizado en el Departamento de Microbiología del Suelo y Sistemas Simbióticos (Grupo de Ecología Genética de la Rizosfera) de la Estación Experimental del Zaidín (CSIC-Granada).

Para la realización del siguiente trabajo, la Lda. Alexandra Peregrina Lavín fue financiada por las siguientes fuentes:

- Beca predoctoral JAE PREDOC adscrita al proyecto AGL2009-07925 del Ministerio de Ciencia e Innovación, actualmente conocido como Ministerio de Economía, Industria y Competitividad.

- Beca de movilidad de la Organización Europea de Biología Molecular (EMBO) para estancias breves, llevada a cabo en el Centro de Microbiología sintética (LOEWE-Zentrum für Synthetische Mikrobiologie, Marburg, Alemania) bajo la dirección de la Doctora Anke Becker. Periodo: Septiembre-Diciembre 2013.

- Beca de movilidad del Servicio Alemán de Intercambio Académico (DAAD) para estancias breves, llevada a cabo en el Centro de Microbiología sintética (LOEWE-Zentrum für Synthetische Mikrobiologie, Marburg, Alemania) bajo la dirección de la Doctora Anke Becker. Periodo: Septiembre-Diciembre 2014.

Parte de los resultados presentados en esta Tesis Doctoral han sido publicados en revistas internacionales:

Torres-Quesada, O., Oruezabal, R. I., Peregrina, A., Jofré, E., Lloret, J., Rivilla, R., Toro, N. and Jiménez-Zurdo, J. I. The *Sinorhizobium meliloti* RNA chaperone Hfq influences central carbon metabolism and the symbiotic interaction with alfalfa. *BMC Microbiology,* **(2010)** 10: 71.

del Val, C., Romero-Zaliz, R., Torres-Quesada, O., Peregrina, A., Toro, N. and Jimenez-Zurdo, J. I. A survey of sRNA families in α-proteobacteria. *RNA Biol,* **(2012)** 9(2): 119–129.

Torres-Quesada, O†., Reinkensmeier, J†., Schluter, J. P†., Robledo, M., Peregrina, A., Giegerich, R., Toro, N., Becker, A. and Jimenez-Zurdo, J. I. Genome-wide profiling of Hfq-binding RNAs uncovers extensive post-transcriptional rewiring of major stress response and symbiotic regulons in *Sinorhizobium meliloti*, *RNA Biol,* **(2014)** 11(5): 563–579 .

Saramago, M†., Peregrina, A†., Robledo, M†., Matos, R. G., Hilker, R., Serrania, J., Becker, A., Arraiano, C. M. and Jimenez-Zurdo, J. I. *Sinorhizobium meliloti* YbeY is an endoribonuclease with unprecedented catalytic features, acting as silencing enzyme in riboregulation. *Nucleic Acids Res,* **(2017)** 45(3):1371-1391.

Robledo, M†., Peregrina, A†., Millán, V., García-Tomsig, N.I., Torres-Quesada, O., Mateos, P.F., Becker, A, and Jimenez-Zurdo, J. I. A conserved α-proteobacterial small RNA contributes to osmoadaptation and symbiotic efficiency of rhizobia on legume roots. *Environ Microbiol,* **(2017)** DOI: 10.1111/1462-2920.13757.

†Co-first authors

RESUMEN

En la simbiosis rizobio-leguminosa, la transición del estado de vida libre en el suelo al interior de la célula dentro de la planta requiere una adaptación flexible de las bacterias invasoras a los cambios externos, lo que implica la expresión coordinada de complejas redes génicas. Por lo tanto, se espera que los sRNAs expresados por el microsimbionte sean elementos reguladores esenciales para la integración de diversas señales suelo-planta y la coordinación de una adecuada respuesta fisiológica bacteriana. A pesar de ello, el papel de la riboregulación en esta endosimbiosis mutualista ha permanecido en gran medida inexplorado. Recientemente, el RNoma no codificante del rizobio modelo *Sinorhizobium meliloti* ha sido ampliamente caracterizado mediante el uso de herramientas postgenómicas de alto rendimiento. El reto ahora es descifrar la función de las miles de diferentes moléculas de ARN expresadas por esta bacteria. En este trabajo, nosotros hemos profundizado en el conocimiento de dos sRNAs previamente identificados en nuestro grupo en el genoma de *S. meliloti* (del Val *et al*., 2007), y hemos explorado la participación que tiene en la riboregulación la proteína YbeY (*Sm*YbeY), muy conservada en bacterias.

El sRNA AbcR2 expresado en respuesta a estrés (Torres-Quesada *et al*., 2013), integra el regulón de RpoH1 (factor σ^{H1}) actuando en la regulación post-transcripcional de múltiples mRNAs transportadores que codifican proteínas implicadas en la captación de aminoácidos y otras fuentes de nitrógeno.

El sRNA NfeR1 está muy extendido entre las α-proteobacterias filogenéticamente relacionadas que interactúan con huéspedes eucariotas (del Val *et al*., 2012; Reinkensmeier and Giegerich, 2015). Su perfil de expresión y sus fenotipos asociados colocan a este sRNA como un nuevo regulador de respuesta al estrés salino que influye tanto en la osmoadaptación como en el comportamiento simbiótico general de *S. meliloti* en las raíces de alfalfa.

La endoribonucleasa YbeY de *S. meliloti* influye en el metabolismo del ARN, en las vías productoras de energía y en las funciones simbióticas codificadas por plásmidos. El perfil de ligandos de ARN y genes dependientes de *Sm*YbeY preve para esta RNasa una serie de sustratos dependientes e independientes de Hfq. Aunque los datos presentados aquí revelan que Hfq y *Sm*YbeY participar en redes de ARN ampliamente independientes, proporcionamos

pruebas de que el silenciamiento de genes dependiente de Hfq relacionado con la fijación de nitrógeno y la absorción de aminoácidos requiere de *Sm*YbeY.

ABSTRACT

In the rhizobia-legume symbiosis, the transition from a free-living state in soil to an intracellular residence within the plant host demands a flexible adaption of invading bacteria to external changes, which involves the coordinated expression of complex gene networks. Therefore, sRNAs expressed by the microsymbiont are expected to be essential regulatory elements for the integration of diverse soil and plant signals and the coordination of the appropriate physiological bacterial response. Despite this evidence, the role of riboregulation in this mutualistic endosymbiosis has remained largely unexplored. Recently, the non-coding RNome of the model rhizobia, *Sinorhizobium meliloti*, has been extensively characterized by diverse post-genomic high-throughput approaches. The challenge now is deciphering the function of the thousands of sRNA molecules of different type expressed by this bacterium. In this work, we deepened into the knowledge of two sRNAs, previously identified in our group in the *S. meliloti* genome (del Val *et al*., 2007), and we explored the involvement in riboregulation of the conserved bacterial YbeY protein.

The stress-induced AbcR2 sRNA (Torres-Quesada *et al*., 2013) integrates the regulon of RpoH1 (σ^{H1}factor) acting downstream in the post-transcriptional regulation of multiple transporter mRNAs coding for proteins involved in the uptake of amino acids and other nitrogen sources.

The salt- induced NfeR1 is widespread in phylogenetically related α-proteobacteria interacting with eukaryotic hosts (del Val *et al*., 2012; Reinkensmeier and Giegerich, 2015). Its expression profile and associated phenotypes place this sRNA as a novel regulator of a salt stress response influencing both osmoadaptation and the overall symbiotic performance of *S. meliloti* on alfalfa roots.

The *S. meliloti* endoribonuclease YbeY influence core RNA metabolism, energy-producing pathways and plasmid-encoded symbiotic functions. Profiling of the *Sm*YbeY-dependent genes and RNA ligands envisages a number of Hfq-independent and -dependent substrates for this RNase, e.g. the nitrogen fixation genes. Although the data presented here reveal that Hfq and *Sm*YbeY participate in largely independent RNA networks, we provide evidence that the

Hfq-dependent silencing of genes related to nitrogen fixation and amino acid uptake requires *Sm*YbeY.

INDEX

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

µg: Microgram

µl: Microliter

µm: Micrometer

ΦCIA: Phenol chloroform isoamyl alcohol

A: Adenine

Ap: Ampicillin

APS: Ammonium persulfate

asRNA: antisense sRNA

bp: base pair

BSA: Bovine serum albumin

C: Cytosine

cDNA: complementary DNA

Cm: Centimeter

CM: Covariance models

CoIP: Co-immunoprecipitation

DNA: Deoxyribonucleic acid

dNTP: Deoxynucleoside triphosphate (dATP, dCTP, dGTP, dTTP, dUTP)

dpi: Days post inoculation

DTT: Dithiothreitol

ECF: Extracytoplasmic function σ factor

EDTA: Ethylenediaminetetraacetic acid

***et al.*:** lat., and others

G: Guanine

GFP: Green fluorescent protein

IGR: Intergenic region

IS: Insertion sequence

Kb: Kilobase

kDa: Kilodalton

Km: Kanamycin

LB: Luria-Bertani medium

M: Molar (mol per liter)

Mb: Megabase

min: Minute

ml: Milliliter

mM: Millimolar

MM: Minimal medium

mRNA: Messenger ribonucleic acid

ms: Millisecond

ng: Nanogram

nt: Nucleotide

NTP: Nucleoside triphosphate (ATP, CTP, GTP, TTP, UTP)

OD: Optical density

ORF: Open reading frame

PCR: Polymerase chain reaction

RACE: Rapid amplification of cDNA ends

RBS: Ribosome binding site

RNA: Ribonucleic acid

rpm: Rounds per minute

rRNA: Ribosomal ribonucleic acid

RT: Room temperature

RT-PCR: Reverse transcription PCR

s: Second

SDS: Sodium dodecyl sulfate

sRNA: Small RNA

T: Thymine

TAP: Tobacco acid pyrophosphatase

Tc: Tetracycline

TEMED: Tetramethylethylenediamine

tRNA: Transfer ribonucleic acid

TSS: transcription start site

TY: tryptone yeast

UTR: Untranslated region

U: Uracil

UV: Ultraviolet

v/v: Volume per volume

V: Volt

WT: Wild-type

w/v: Weight per volume

w/w: Weight per weight

INTRODUCTION

I-1. THE BACTERIAL RNome: NON-CODING RNAs

I-1.1. THE NON-CODING RNome CONCEPT

Until the discovery of retroviruses, the central dogma of molecular biology stated that genes are transcribed to make ribonucleic acid (RNA) which in turn is translated into protein (Crick 1970). For years, the transcription of the bacterial genome was considered to result in the generation of three large groups of RNA molecules: messenger mRNA (mRNA), containing open reading frames (ORFs) that are translated into proteins, and two further types of RNA, referred to as ribosomal RNA (rRNA) and transfer RNA (tRNA), that are not translated into protein but are essential for protein biosynthesis carried out by ribosomes (Crick 1970; Temin 1985).

Given this information, it would appear that all cellular functions are performed by polypeptides encoded in ORFs, which represent 88-90% of the bacterial genome. However, post-genomic research, based on extensive genome comparisons and high-throughput sequencing, has challenged this central dogma of biology and is still revealing a high abundance and diversity of unexpected RNA molecules, which do not encode proteins but have diverse functions in maintaining bacterial physiology. In many cases, these RNAs are encoded in Intergenic Regions (IGR), the non-coding genomic zones flanking the annotated genes. Consequently, in the last years, the term non-coding RNome has been coined to allude the set of non-protein coding transcripts expressed in any organism. In bacteria these molecules are relatively short (50-400 nt), and therefore they are generically called sRNAs (Small non-coding RNAs) or ncRNAs (Non-coding RNAs) (Eddy, 2001) (Figure 1).

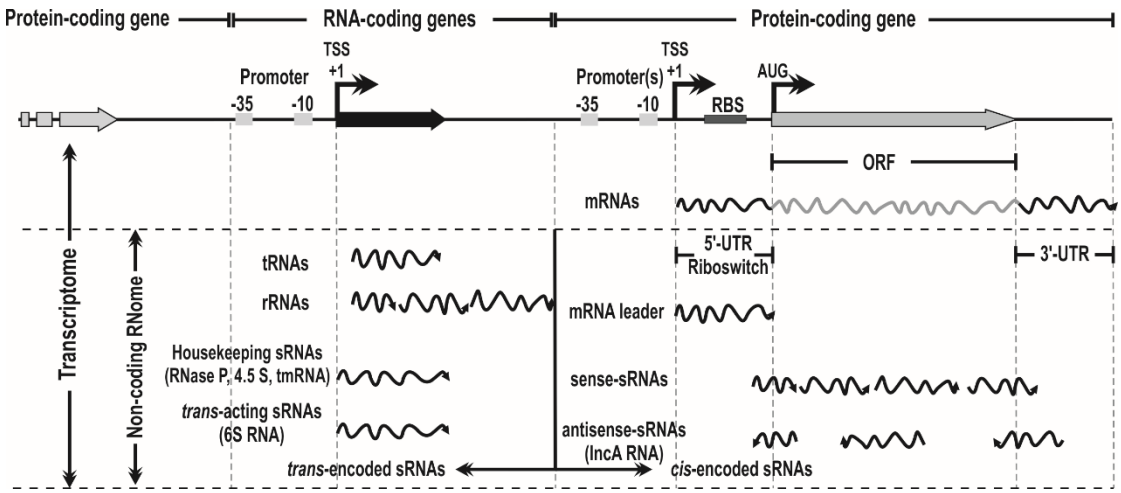

Figure 1. Non-coding RNome structure. Small untranslated transcript (sRNA) species expressed in bacteria. *Trans*-acting sRNAs are encoded in independent transcription units with recognizable promoter and termination signatures between annotated open reading frames (ORFs). *Cis*-encoded sRNAs are transcribed sense or antisense to mRNAs. *Cis*-sense sRNAs can either form chimaeras with the coding regions of mRNAs or be synthetized as short, discrete transcripts resulting from transcription attenuation (mRNA leaders), mRNA processing, or co-transcription with the full-length mRNA from alternative promoters. TSS, transcription start site; RBS, ribosome binding site (modified from Jimenez-Zurdo *et al.,* 2013).

I-1.2. DISCOVERY OF BACTERIAL sRNAs

In the 1970's, it was already observed that in some phages, plasmids and bacterial transposon genes, the non-coding DNA strand gives rise to transcripts complementary to part of or the whole mRNA sequence. These molecules, generically referred to as antisense RNAs (asRNAs), were discovered to be involved in the regulation of the propagation of extrachromosomal elements that encode them, such as replication of plasmids or phages, or compatibility between origins of plasmid replication (Tomizawa, 1984, 1985).

Soon after, the discovery of chromosomal sRNAs occurred rather casually through genetic tracing and radiolabeling of total RNA. The first sRNAs (approximately ten) were discovered in the model organism *Escherichia coli* (Wassarman *et al.,* 1999) and were found to have functions in diverse cellular processes. Some of these transcripts

are M1 RNA or RNasa P, tmRNA, SRP RNA, 6S RNA, and Spot42, among others. In the following sections, the function of some of them is briefly described.
More recently, in 2001, three independent working groups identified 31 new sRNAs encoded on the *E. coli* chromosome. The identification of these transcripts was achieved by comparative analysis of IGR sequences of this bacterium with the genomes of related enterobacteria followed by experimental verification by Northern blot and/or RACE mapping (Rapid Amplification of cDNA Ends) (Rivas *et al.*, 2001; Argaman *et al.*, 2001; Wassarman *et al.*, 2001).

Additional studies approached the experimental identification of sRNAs in this model bacterium by microarray hybridization or massive sequencing of cDNAs (Rivas *et al.*, 2001; Carter *et al.*, 2001; Huttenhofer *et al.*, 2006). All these works have laid the methodological basis of the bacterial RNomics and resulted in the development of new technologies for sequencing-based transcriptome profiling (essentially pyrosequencing) together with the availability of an increasing number of sequenced bacterial genomes (6893 in June 2016; https://www.ncbi.nlm.nih.gov/genome/browse/). Taken together, methodological advances in the RNomics field allowed the structural characterization of the non-coding RNome for a significant number of bacteria including some pathogens of clinical relevance (Vogel and Wagner, 2007).

The families and functional categories of RNAs generated by covariance models of the primary nucleotide sequences as well as secondary structures are reported in the Rfam database (http://rfam.xfam.org/).

I-1.3. FUNCTIONAL CLASSIFICATION OF sRNAs

The functional characterization of a representative group of recently identified sRNAs permits its classification into two categories of transcripts (Gottesman 2004, Storz 2002):

- **Transcripts with housekeeping functions.**
  This group includes RNAs that fundamentally control biosynthesis and secretion of proteins; tRNAs, rRNAs, tmRNA and SRP RNA. This division also includes all ribozymes such as RNase P.

- **Regulators of gene expression (riboregulators).**
  This is the largest group of sRNAs in bacteria. These sRNAs are divided into two main classes: *cis*-acting sRNAs and *trans*-acting sRNAs.

I-2. RNAs WITH HOUSEKEEPING FUNCTIONS

RNAs with essential functions were the first chromosomal RNAs discovered in *E. coli* because they are particularly abundant in prokaryotic cells and play important roles in the maintenance of homeostasis. Particularly, the latter feature allowed the identification of orthologues in phylogenetically distant bacteria. In this sRNA group it is necessary to emphasize the following important representatives:

I-2.1. RNA M1 (RNase P)

RNase P is an endoribonuclease responsible of the maturation of the 5'-ends of the precursor tRNAs. In *E. coli*, RNase P is composed of two subunits, RNA M1 and protein C5, which form a stable ribonucleoprotein complex (RNP). This ribozyme is not only conserved in bacteria but also in eukaryotes so that it is considered as essential sRNA in any living organism (Altman *et al.*, 1999; Frank and Pace, 1998).

I-2.2. tmRNA

The *transfer-messenger* RNA (tmRNA) is a type of RNA functioning as a transferor and messenger during a process called "*trans*-translation", which releases those ribosomes that are occupied by defective nascent polypeptides during the translation process (Keiler *et al*, 1996; Giller *et al.,* 2003). This molecule has two recognizable domains; the tRNA domain is recognized by the alanine-tRNA synthetase, which permits the binding of the tmRNA to the ribosome and subsequently to the stuck peptide. The second domain of the tmRNA messenger domain is then translated resulting in release of the ribosome and the protein with a final sequence recognized by cellular proteases that immediately degrade the aberrant protein (Lindell *et al.*, 2002).

I-2.3. SRP RNA (4.5 S RNA)

4.5 S RNA is associated to the Ffh protein (signal recognition) with which it forms an RNP complex called SRP responsible for recognizing the signal sequence of the nascent peptide that emerges from the ribosome for its secretion to the extracellular medium (Regalia *et al.,* 2002; Siu *et al.,* 2007).

I-3. RIBOREGULATORS (sRNAs)

Most recognized sRNAs have functions related to regulation of gene expression. These sRNAs can be either *cis*-encoded or *trans*-encoded, depending on their genomic location relative to the mRNA that they regulate (Georg & Hess, 2011; Jimenez-Zurdo *et al.,* 2013) (Figure 2).

I-3.1. *cis*-sRNAs

The biogenesis of *cis*-sRNAs (*cis*-encoded RNAs, antisense RNAs, or asRNAs) is diverse as they can be originated either from processing of the mRNAs by the cellular ribonucleases or the transcription of the non-coding strand of some genes (*i.e.* antisense sRNAs identified in extrachromosomal elements). Alternatively, co-transcription with an mRNA from a single promoter (*i.e.* riboswitches or mRNA leading regions) or from a differentially regulated alternative promoter can yield *cis*-sRNAs (Vitreschak *et al.,* 2002; Serganov *et al.,* 2009) (Figure 1).

Riboswitches and other mRNA leading regions are particular types of 5'-untranslated regions (5'-UTR) of the messengers to which they are associated. The regulatory mechanism for mRNA translation is based on their secondary structure. The remaining sRNAs in this group generally use an antisense mechanism by which they interact with complementary sequences in their target mRNA, thus influencing their translation and/or stability (Vitreschak *et al.,* 2002; Serganov *et al.,* 2009) (Figure 2).

I-3.2. *trans*-sRNAs

The majority of new sRNAs with assigned functions belong to the group of *trans*-sRNAs (*trans*-encoded RNAs, *trans*-acting RNAs) (Gottesman and Storz, 2011). These are encoded at loci far from the genomic location where the target messenger is transcribed (Figure 2), and the sRNA-mRNA mating is the major mechanism of action of these *trans*-sRNAs (Gottesman and Storz, 2011). *Trans*-sRNAs are expressed from a differentially regulated promoter and their transcription is usually terminated in a Rho-independent fashion (Jimenez-Zurdo *et al.,* 2013).

Notable and well-characterized exceptions from this general ribo-regulatory mechanism are the sRNAs 6S and CsrB, also classifiable as *trans*-sRNAs (Wassarman and Storz, 2000; Timmermans and Van Melderen, 2010). These sRNAs exert their

regulatory activity by sequestration and modification of specific protein activities due to presentation of sequences or motifs that they recognize in the cellular components with which they are normally associated, thus competing with them (Barrick *et al.*, 2005; Cavanagh *et al.*, 2008; Babitzke and Romeo, 2007).

In the following sections, two fundamental mechanisms for the regulation of gene expression mediated by *trans*-sRNAs will be described.

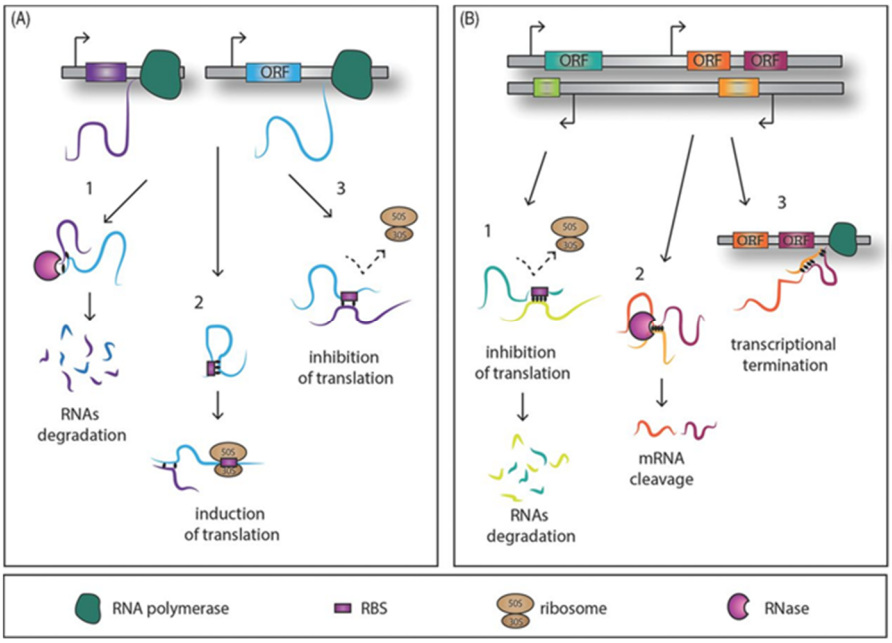

Figure 2. Simplified model of *trans*- and *cis*-acting sRNAs. (A1) *trans*-encoded sRNAs interact with their target RNA through imperfect base-pairing and hence promote RNase degradation of the double-stranded RNA molecules. **(A2)** Alternatively the sRNAs might affect the target translation positively **(A3)** and negatively by releasing or masking the ribosome binding site, respectively. **(B1)** Conversely, the ***cis*-encoded sRNAs** bind the mRNA target through full sequence complementarity, affecting translation, and end in the degradation of the sRNA-target RNA complex. **(B2)** A small *cis*-encoded RNA, antisense between two genes, can lead to mRNA cleavage or **(B3)** the transcriptional termination through a putative loop formation and consequently the cessation of the RNA polymerase activity (Taken from Oliva *et al.*, 2015).

I-3.2.1. *trans*-sRNAs that modify protein activity

In bacteria, two families of regulatory sRNAs that act by mimicking other nucleic acids have been characterized most extensively. The first, represented by the *E. coli* **6S RNA**, imitates a DNA promoter open complex and interacts with the RNA polymerase. It has been found in many other bacterial species, including Gram-positive and Gram-negative bacteria, although primary sequence conservation is very limited. Some organisms have multiple 6S RNAs (*i.e. Bacillus subtilis* and *Clostridium difficile*), although the purpose of this multiplicity is not yet known (Barrick *et al.,* 2005; Trotochaud and Wassarman 2005).

The function of this sRNA was understood many years after it had been first identified, as 6S was found to bind tightly to the RNA polymerase holoenzyme containing σ^{70} but not to the related σ^{S} holoenzyme (Hindley, 1967; Brownlee, 1971; Wassarman and Storz, 2000). This sRNA forms a double-stranded RNA hairpin with a critical bubble that mimics the DNA in an open complex promoter, binding at the active site of RNA polymerase. RNA polymerase was found to synthesize a short transcript encoded by the 6S RNA, starting within the bubble, evidencing that this sRNA can mimic DNA (Wassarman and Saecker 2006; Sharma *et al.,* 2010). Interaction of the 6S RNA with RNA polymerase requires region 4.2 of σ^{70} also involved in binding to the -35 region of promoter DNA (Cavanagh *et al.,* 2008; Klocko and Wassarman 2009). This RNA is processed from a longer mRNA, unlike other *E. coli* regulatory sRNAs, and it accumulates in stationary phase. Its transcription was found to require high levels of nucleotides and only occur during the transition from stationary phase into exponential growth when nucleotide levels increase. Once released, 6S becomes sensitive to degradation (Wassarman 2007; Wurm *et al.,* 2010).

The second family, named the **CsrB** family of sRNAs, includes sRNAs that regulate the CsrA/RsmA family of translation regulatory proteins by competing with mRNA targets. Members of this family are fairly widespread, found in both Gram-negative

and Gram-positive bacteria, and sometimes are also encoded in more than one copy per genome (White *et al.,* 1996; Babitzke and Romeo 2007; Sahr *et al.,* 2009). They seem to play central roles in the selection of bacterial life-styles between swarming, free-swimming, and biofilm formation, as well as modulating virulence, although the precise roles vary between bacterial species (Oliva *et al.,* 2015) (Figure 3).

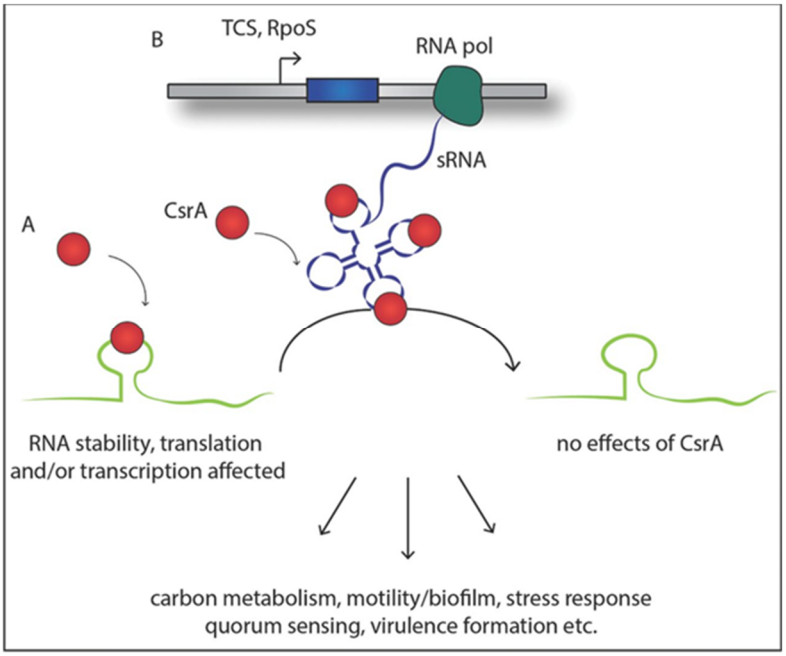

Figure 3. Simplified scheme of the general function of the Csr system. (A) The dimeric RNA-binding protein CsrA interacts with one or more A(N)GGA motifs of the target RNA, typically represented in a hairpin loop structure, leading to an alteration of the accessibility of the translation machinery, transcription anti-termination and/or the stability of the RNA. **(B)** Production of regulatory sRNA under the control of a two-component system and the alternative sigma-factor RpoS antagonizes CsrA function by binding CsrA with high affinity and hence abolishing the interaction with their target RNAs. This results in a global metabolic shift within the pathogen and among others to the activation of virulence-related proteins **(**Taken by Oliva *et al.,* 2015).

CsrB sRNAs were first identified in *E. coli* as critical posttranscriptional regulators of the switch between gluconeogenesis and glycolytic growth, inhibiting glycogen synthesis. There are reported cases of negative and positive regulation, but the mechanism for the latter remains unknown (Babitzke *et al.,* 2009; Wei *et al.,* 2001). Synthesis of the members of this sRNA family is dependent on the two-component system BarA/UvrY, and this regulatory cascade is conserved in many bacteria. In

addition, their stability is regulated: the turnover of the sRNAs is generally fast, but these sRNAs are completely stabilized in cells mutant for either RNase E, the essential endonuclease involved in degradation of many mRNAs, or CsrD, a protein with homology to proteins associated with synthesis and degradation of the signaling molecule cyclic di-GMP (Suzuki *et al.,* 2006).

I-3.2.2. *trans*-sRNAs acting by base pairing with target mRNAs

The main mechanism of riboregulation by *trans*-sRNAs is based on the complementarity matching of sRNA bases with 5'-UTR regions (around the Shine-Dalgarno sequence) of target mRNAs (Figure 4).

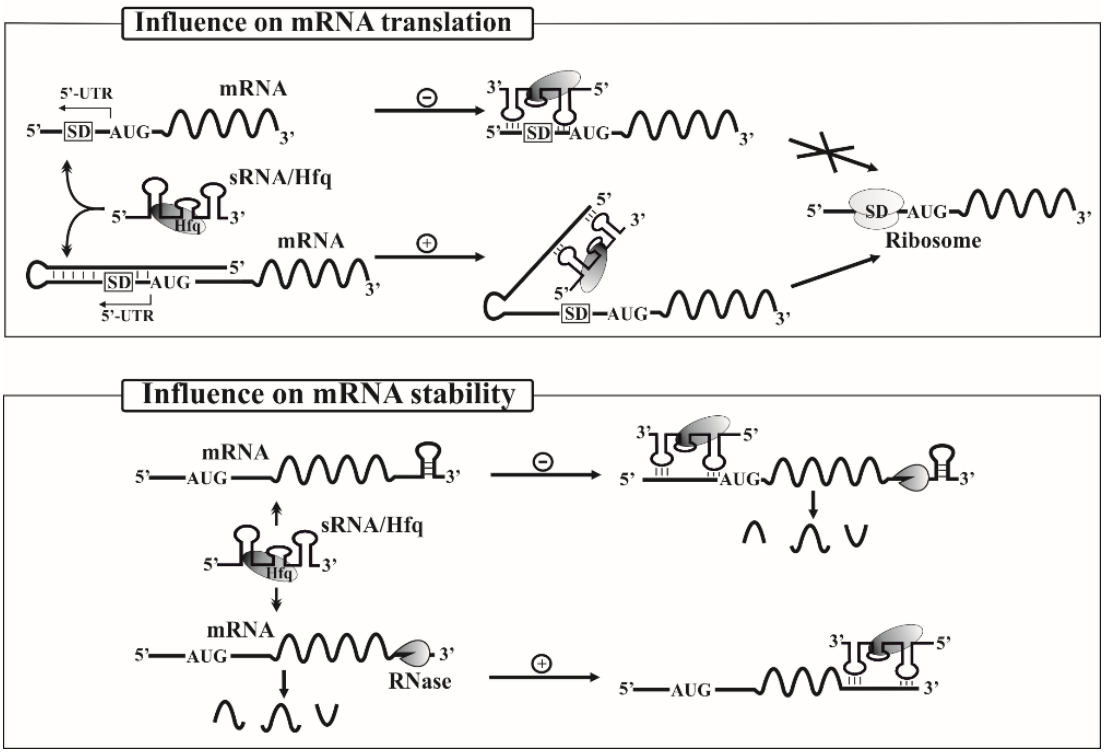

Figure 4. Main mechanism of *trans*-acting sRNAs.

This complementarity is often imperfect and involves short and discontinuous series of nucleotides in both molecules. It also may influence the translation of mRNA

and/or its stability. In most cases, the Shine-Dalgarno sequence of the mRNA is occluded, thus preventing accessibility of the ribosomes to the mRNA with inhibited translation as a consequence. In fewer cases, sRNA-mRNA complex formation relaxes translation-inhibitory secondary structures in the 5'-UTR region thereof, making it accessible to ribosomes and assisting translation. In addition, the sRNA-mRNA duplex may be a substrate or a signal recognized by different cellular RNases for mRNA degradation. By contrast, in other cases this mating may also have a positive effect on mRNA stability by protecting it from the action of ribonucleases (Majdalani *et al.,* 2005) (Figure 4).

The sRNA-mRNA interactions are generally facilitated and/or stabilized by a very abundant and widely distributed RNA-binding protein (~ 12 kDa) called Hfq (Valentin-Hansen *et al.,* 2004) (Figure 4), described in more detail below.

I-4. ADDITIONAL FACTORS FOR sRNA-MEDIATED REGULATION OF GENE EXPRESSION

I-4.1. THE RNA-BINDING PROTEIN HFQ

The Hfq protein was discovered almost 50 years ago in *E. coli* as a host factor required for the replication of the RNA phage Qβ (Franze de Fernández *et al.,* 1968). The rapidly growing genomic database helped to reveal that nearly half of the sequenced bacterial genomes and a few archaea encode a recognizable homolog of this protein, which is usually highly represented in the proteome repertoire. Genetic, biochemical and structural data evidenced a quaternary arrangement of Hfq into a hexameric toroid structurally related to the eukaryotic LSm family of RNA-binding proteins (Brennan and Link, 2007). The homo-hexameric Hfq ring exposes two different positively-charged surfaces (*i.e.* the proximal and distal faces), which constitute alternative binding sites that can discriminate between RNA molecules

(Figure 5). Studies on *Staphylococcus aureus* and *E. coli* Hfq revealed that the proximal face has a preference for uridine-rich RNA stretches, which seem to be accommodated around the pore in a constricted conformation that is stabilized by water molecules (Schumacher *et al.,* 2002; Sauer and Weichenrieder, 2011). As most sRNAs have typical bacterial Rho-independent terminators that usually contain a poly-U 3'-terminus, Hfq can interact with the terminators and influence sRNA stability (Vogel and Luisi, 2011). By contrast, the *E. coli* Hfq distal face presents an RNA-binding motif with preference for adenine-rich RNA segments. In this case, the RNA molecule is accommodated in a circular conformation along the distal surface. Hence, each Hfq ring is able to simultaneously bind two different RNA molecules or even a single molecule bridging both faces around the oligomer rim (Figure 5). If a sRNA binds on one face and a cognate target mRNA does so on the second face, this ternary complex will lead to productive RNA duplex formation (Wang *et al.,* 2013). In addition, Hfq offers a scaffold for the interaction with several other proteins (Sobrero and Valverde, 2012). An important interaction partner of Hfq is the major bacterial ribonuclease RNAse E, which engages in the formation of an atypical degradosome (Morita *et al.,* 2005). This silencing complex ensures the efficient modulation of the riboregulatory networks (Aiba, 2007).

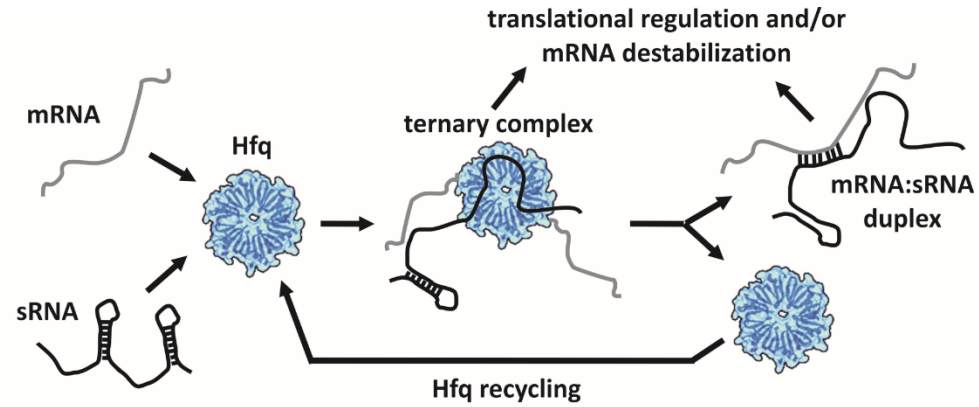

Figure 5. Hfq facilitates annealing of sRNA with their cognate mRNAs. The Hfq hexamer can bind different RNA molecules in either of its positively charged surfaces. RNA binding may result in secondary structure changes that promote base-pairing recognition and duplex formation due to proximity in the ternary complex. The bound

sRNA may block (or even facilitate) ribosome access to the mRNA; alternatively, Hfq-recruited RNAse may irreversibly downregulate mRNA expression (taken from Jiménez-Zurdo *et al.,* 2015).

The importance of Hfq for sRNA-mediated regulation was first evident in studies on OxyS RNA (Zhang *et al.,* 1998). By now it is known that Hfq interacts with most of the regulatory sRNAs as well as diverse mRNAs (Sittka *et al.,* 2008; Zhang *et al.,* 2003) and is required for the intracellular stability of many regulatory sRNAs (Valentin-Hansen *et al.,* 2004). Hfq turned out to have an RNA chaperone activity as changes in secondary structures of some sRNAs (*e. g.* OxyS and Spot42) as well as mRNAs (e. g. *sodB* and *ompA*) have been observed in structure probing experiments (Geissmann and Touati, 2004; Moll *et al.,* 2003).

Hfq interacts preferentially with *trans*-riboregulators but also with other RNA species such as mRNAs. Thus, it has emerged as a global post-transcriptional regulator of gene expression in bacteria. Deletion of the chromosomal *hfq* gene was early observed to impair multiple stress responses in *E. coli* under certain environmental conditions (Tsui *et al.,* 1994). Subsequently, similar reverse genetics approaches have confirmed the pleiotropic effects of the *hfq* mutation in a number of phylogenetically distant bacterial species with diverse lifestyles (Sobrero and Valverde, 2012). In animal pathogens (*e.g.* enterobacteria or *Brucella* species) the absence of Hfq has been shown to attenuate motility, secretion of virulence factors, host invasion or intracellular survival of bacteria (Gottesman, 2004; Robertson and Roop, 1999; Sittka *et al.,* 2007). These evidences suggest a universal role of Hfq in the establishment and maintenance of chronic intracellular residences within eukaryotic hosts.

I-4.2. RIBONUCLEASES

Besides the RNA chaperone Hfq, several ribonucleases (RNases) have been described to play important and even essential roles in post-transcriptional regulation (Régnier

and Arraiano, 2000; Arraiano and Maquat, 2003; Viegas *et al.,* 2007). Mutants deficient in RNases were utilized to characterize RNA decay mechanisms (Arraiano *et al.,* 1988; Viegas *et al.,* 2004; Viegas *et al.,* 2005). Since RNases are key modulators of RNA decay, the identification of the RNases that contribute to the decay of individual sRNAs is essential for a more general understanding of sRNA turnover *in vivo* (Viegas and Arraiano, 2008).

RNase E is one of the major endoribonucleases in *E. coli,* although it is not ubiquitous in bacteria. It cleaves single-stranded regions of structured RNAs with preference for 5'-ends and AU-rich sequences (Ow *et al.,* 2003). This endoribonuclease is also one of the main enzymes forming the degradosome, a multiprotein complex involved in the decay of many RNAs as well as in the processing of ribosomal and transfer RNAs (Carpousis *et al*., 1994; Arraiano and Maquat, 2003; Carpousis, 2002; Régnier and Arraiano, 2000). RNase E was also found to be important for coupled sRNA-mRNA degradation, to co-purify together with two sRNAs, SgrS and RyhB, and Hfq, and to be required for degradation of the mRNA targets *ptsG* and *sodB* (Afonyushkin *et al.,* 2005; Massé *et al.,* 2003; Morita *et al.,* 2006). The Hfq binding to RNase E was shown to take place at the C-terminal scaffold domain which is also required for RyhB-mediated degradation of *sodB* mRNA (Massé *et al.,* 2003). The RNase E-Hfq-sRNA RNP complex leads to translational repression and rapid target mRNA degradation. However, Hfq binding in the absence of RNase E and RNA-RNA interaction itself are sufficient to mediate translational repression, destabilization, and degradation of the target mRNA (Maki *et al.,* 2008; Morita *et al.,* 2006).

Another RNase which is involved in post-transcriptional regulation by bacterial sRNAs is the double-strand specific **RNase III**. Some of its main functions are rRNA processing (Babitzke *et al.,* 1993; Nicholson, 1999; Evguenieva-Hackenberg and Klug, 2000) and the decay of some mRNAs (Régnier and Grunberg-Manago, 1990; Santos *et al.,* 1997). The enzyme is also responsible for the cleavage of several bacterial and phage messengers. RNaseIII is a highly conserved enzyme specific for double-stranded RNAs and showing preference for continuous RNA duplexes of 20–

40 bps. Perfect antisense/sense RNA duplexes formed in sRNA/mRNA interactions constitute an optimal substrate for this enzyme (Lamontagne and Elela, 2004). In *E. coli*, an antisense interaction between the SOS-induced small RNA IstR-1 and its target *tisAB* was found to entail RNase III-dependent cleavage and thereby inactivates the mRNA for translation (Vogel *et al.,* 2004). Furthermore, the decay of RyhB *in vivo* was shown to be mainly dependent on RNase III in contrast to the RNase E-dependent turnover of its target *sodB* mRNA. Cleavage of RyhB by RNase III *in vitro* is facilitated upon base pairing with the *sodB* 5'-UTR (Afonyushkin *et al.,* 2005). Moreover, RNase III is important for regulation of several virulence factors in *Staphylococcus aureus* (Boisset *et al.,* 2007; Huntzinger *et al.,* 2005). In this case, coordinated action of RNase III is essential to degrade the mRNA and irreversibly arrest translation *in vivo*.

I-5. sRNAs IN α-PROTEOBACTERIA: *Sinorhizobium meliloti*

Sinorhizobium meliloti is a representative member of the family Rhizobiaceae within the order Rhizobiales of the α-proteobacteria. Soil-dwelling and diazotrophic *S. meliloti* either occurs in a free-living state or enters symbiosis with leguminous host plants from the genera *Medicago*, *Melilotus* and *Trigonella*.

Recent phylogenetic studies show that species of the genus *Sinorhizobium* are closely related to *Ensifer adhaerens* (Casida, 1982), so the genus *Sinorhizobium* has been reclassified as *Ensifer*. The name *Ensifer* was published in 1982 and the name *Sinorhizobium* was published in 1988. By the rules of the Bacteriological Code of the International Committee on Systematics of Prokaryotes, the older name (*Ensifer*) has priority (Young, 2003). In response to a request that the single extant species of *Ensifer* (*Ensifer adhaerens*) be moved to *Sinorhizobium*, a special subcommittee was formed to evaluate the request. It was ultimately ruled that *Ensife*r retained priority and that all *Sinorhizobium* species be transferred to the genus *Ensifer* (Lindstrom and

Martinez-Romero, 2002; Judicial Commission of the International Committee on Systematics of Prokaryotes, 2008). However, both terms continue to be used in published scientific literature, with *Sinorhizobium* being the more common one. Therefore, in this work we have preferred to use the name of *Sinorhizobium* for being the best known and accepted among the rhizobiologists.

I-5.1. BIOLOGY OF *S. meliloti*

I-5.1.1. The *S. meliloti* - *Medicago truncatula* symbiosis

The *S. meliloti-Medicago truncatula* symbiosis is a reference model system for studying rhizobia-legume interactions. The complex infection process for the establishment of this symbiosis includes a sequence of spatiotemporally coordinated steps schematized in Figure 6.

Several abiotic factors and environmental conditions are decisive for optimal growth and fitness of both plants and bacteria in soil, such as water availability, salinity, soil pH, temperature, heavy metals, nutrient deficiency and mineral toxicity (Triplett and Sadowsky, 1992). Subsequently, the mutualistic relationship between both partners is initiated after a series of complex exchanges of chemical signals between the bacteria and the legume plant. As a starting point, the legume plants secrete isoflavonoids that are recognized by the compatible bacterial NodD proteins, resulting in the initiation of the nodulation genes (nod genes) (Peck *et al.,* 2006; Chang *et al.,* 2009). Nod genes encode proteins that produce specific lipochito-oligosaccharides called Nod factors (NFs) that trigger the root infection process and initiate cell division in the root cortex. Other compounds with a signaling function, such as bacterial surface lipopolysaccharides (LPS), contribute to the infection process and establishment of an efficient symbiosis (Chang *et al.,* 2009). Nod factors induce certain morphogenetic responses in the plant such as curvature of the root hairs and the active division of root cortex cells forming the so called nodule primordium. The penetration of the bacteria

into the root occurs through infection threads generated in the curled root hairs (Yang *et al.*, 1994; Jones *et al.*, 2007). In infection threads, the bacteria divide and progress until they reach the nodule primordium where they are released into the plant cells. Then, differentiation of bacteria is initiated resulting in the formation of morphologically distinctive and polyploid bacteroids (Limpens *et al.*, 2005). The result of this symbiotic interaction is the formation of new organs, the nodules, on the legume roots. The formation of root nodules creates a basis for the symbiotic relationship between the legume plant and the N_2-fixing rhizobial bacterium. Once inside the nodules, the host provides carbon compounds to the bacteria as well as a favorable environment where competition with other soil microorganisms is absent (compared to the rhizosphere zone) (Oono *et al.*, 2011). In response, the rhizobia

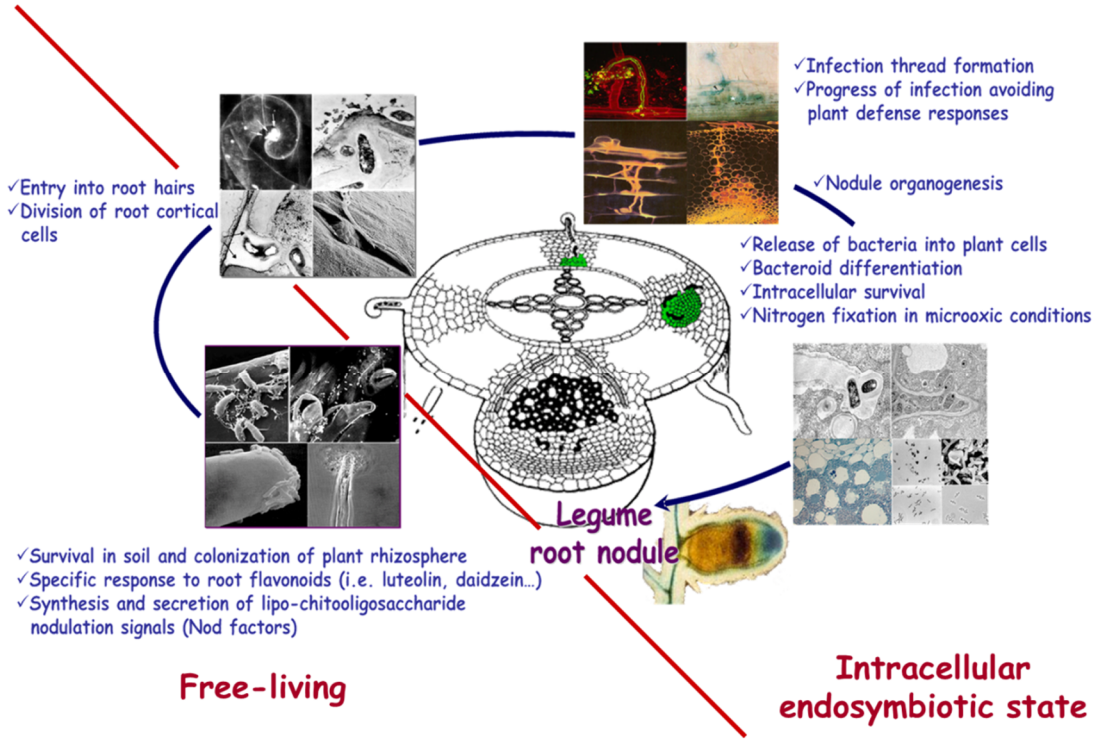

Figure 6. Stages of the *Rhizobium*-legume symbiosis. The depiction shows the cross section of a root at different stages of nodule organogenesis, illustrated with the photographs obtained by different microscopic techniques.

convert atmospheric dinitrogen (N_2) to inorganic nitrogen compounds, preferably NH_4^+, which is then incorporated into amino acids that can be utilized by the plant. This reaction is catalyzed by the nitrogenase, whose activity requires a low oxygen environment inside the nodule (Oldroyd, *et al.*, 2011).

The histidine kinase FixL is the enzyme which perceives low oxygen concentrations in the bacteroid membrane and transduces this signal through FixJ, whose phosphorylated form activates the transcription of *nifA* and *fixK*, both controlling expression of genes that encode structural proteins of the nitrogenase. Rhizobial genomes encode a ferro-molybdenum nitrogenase composed of two components, the homodimer nitrogenase reductase and the heterotetramer nitrogenase. This complex uses a total of 8 electrons and 16 molecules of ATP for the reduction of a molecule of dinitrogen resulting in the formation of 2 molecules of ammonium (Fischer, 1994; Schwarz *et al.,* 2009).

Symbiotic transitions thus demand continuous rhizobial adaptation to widely diverse abiotic and plant cues, which involves profound changes in gene expression and in the activities of enzymes as well as of transport proteins. Until very recently, regulation of gene networks underlying these adaptive responses was almost exclusively attributed to actions of transcription factors and alternative σ RNA polymerase holoenzymes, which are encoded in exceptionally large numbers by the rhizobial genomes. As attention is increasingly drawn to sRNAs as novel ubiquitous components of regulatory networks in bacteria (Waters and Storz, 2009; Beisel and Storz, 2010; Storz *et al*., 2011), it is therefore reasonable that sRNAs play also pivotal roles in the establishment of the rhizobia-legume interaction.

I-5.2. THE GENOME OF *S. meliloti* STRAIN 1021

Since the genome of *S. meliloti* is well characterized, this rhizobia specie is considered as ideal model for genomic analysis. Its genome contains a circular chromosome (3.65

Mb) and two megaplasmids, pSymA (1.35 Mb) and pSymB (1.68 Mb) (Galibert *et al.,* 2001) which host most of the genes necessary for the establishment of the symbiotic interaction (Figure 8). During the course of evolution, the genome of *Sinorhizobium* acquired new functions through acquisition of new transport and regulatory genes (Galibert *et al.,* 2001). The megaplasmid pSymA carries most of the genes required for nodulation and nitrogen fixation (*nod*, *nif*, and *fix* genes), carbon metabolism, transport, stress responses, whereas pSymB reveals a high number of genes involved in polysaccharide biosynthesis, also required for the establishment of symbiosis and other ecological specializations (Blanca-Ordóñez *et al.,* 2010).

The first sequence analysis of the *S. meliloti* 1021 genome resulted in the annotation of 6,206 ORFs, 54 tRNAs, 3 operons of rRNAs, a copy of tmRNA, as well as multiple insertion sequences (ISs) and repetitive extragenic palindromic (REP) elements. Among all 6,206 predicted protein-encoding genes, 3341 were found on the chromosome (Capela *et al.,* 2001), 1293 on pSymA (Barnett *et al.,* 2001), and 1570 on pSymB (Finan *et al.,* 2001). At the time of *S. meliloti* genome sequence determination, there was no experimentally prove of function for a vast majority of the predicted genes and 40% could not be placed into a functional category. Moreover, 8% were orphan genes, defined as those not found in any other sequenced genome. In 2009 a *S. meliloti* genome annotation update was published which incorporated information published from 2001 to 2008. With improved prediction tools, they identified 86 new putative genes, removed 66 previously predicted orphan genes and adjust the start positions of 360 coding regions. As a result, more than 71% of genes had then a predicted function (Becker *et al.,* 2009).

I-5.2.1. *S. meliloti* strains 2011, 1021 and 20112B3001

Rm2011, Rm1021 and 20112B3001 are three closely related *S. meliloti* strains, all of them capable of forming nitrogen-fixing nodules in *Medicago* hosts.

Rm2011 and Rm1021 are derivatives of the wild-type *S. meliloti* isolate SU47 separated by 25 years of growth under various media conditions in different laboratories (Brockwell and Hely, 1966; Casse *et al.,* 1979; Meade and Signer 1977). There are obvious phenotypic differences between both in terms of symbiotic growth and behavior (Wais *et al.,* 2002; Krol and Becker, 2004).

Rm1021 and Rm2011, both are known to have a disrupted *expR*, a gene required for swarming. *S. meliloti* Sm2B3001 is an Rm2011 derivative with restored *expR* gene on the chromosome (Bahlawane *et al.,* 2008; Nogales *et al.,* 2012). Both, Rm2011 (*expR*) and Sm2B3001 (*expR*+) have been used during the last years to establish the role of ExpR in the regulation of promoter activities in *S. meliloti* (Charoenpanich *et al.,* 2013; Nogales *et al.,* 2012).

In this work, all three strains were used indistinctly as reference of *S. meliloti*, focusing especially on Sm2B3001 (*expR*+) strain. Although, in general, there are no apparent inconsistencies in the results obtained in the different strains, some subtle differences in the expression profiles of specific sRNAs among strains could have their origin in punctual genomic.

I-5.3. THE *S. meliloti* non-coding RNome

Concerning the identification of non-coding RNAs, at the time of publication of the *S. meliloti* 1021 genomic sequence genes for only the above mentioned essential RNAs were annotated: tRNAs, rRNAs and tmRNA (Capela *et al.,* 2001).

In 2007, a pioneering study described a bioinformatics screening method for the prediction of sRNA genes in this genome (del Val *et al.,* 2007). This method is based on comparative analysis of the IGRs with the α-proteobacterial genomes phylogenetically related to *S. meliloti*. The alignments generated in these comparisons were then analyzed with the QRNA and RNAz programs. QRNA identifies base

substitution patterns in pairwise alignments likely corresponding to a conserved RNA secondary structure rather than to a conserved coding frame or other genomic features (Rivas *et al.*, 2001), whereas RNAz combines an estimation of thermodynamic stability with structure conservation for the RNA predictions (Washietl *et al.*, 2005). This analysis generated an initial list of 32 candidates as most probable sRNA genes.

Verification of QRNA/RNAz predictions by Northern hybridization and RACE mapping led to the identification of 8 sRNAs (Smr RNAs; *S. meliloti* RNA) encoded in independent transcriptional units. Seven of the identified sRNAs are differentially expressed in free-living and symbiotic bacteria, while homologs are present only in α-proteobacteria (Figure 7) (del Val *et al.*, 2007). These characteristics support the classification of these molecules as *trans*-RNAs with regulatory activity.

Following this study, 14 novel sRNAs were discovered and validated by combining several computational approaches with microarray, Northern and dot blot hybridizations (Ulvé et *al.*, 2007; Jiménez-Zurdo *et al.*, 2013). Computational predictions and microarray hybridization experiments were also used to screen the intergenic regions resulting in 14 candidates that were confirmed as novel small non-coding RNAs by Northern blot and/or microarray hybridizations (Valverde *et al.*, 2008; Jiménez-Zurdo *et al.*, 2013). Figure 8 shows the distribution of the sRNA loci identified in *S. meliloti* by these three first studies based on comparative genomics.

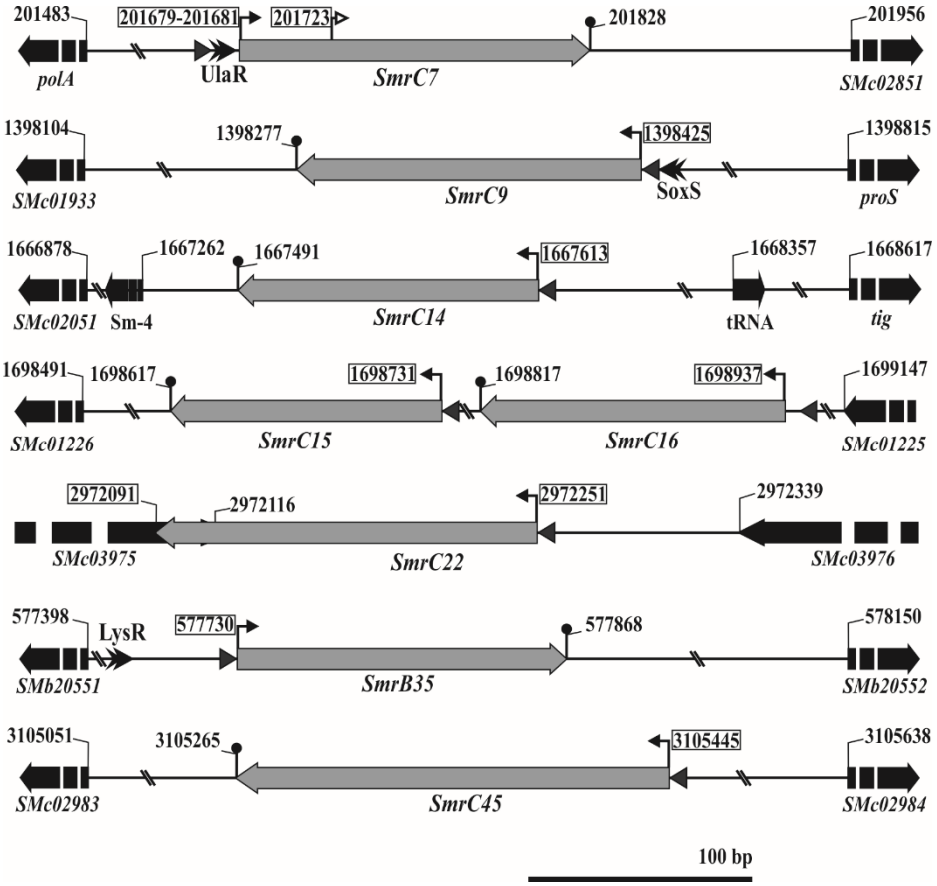

Figure 7. Genomic regions of the identified *S. meliloti* sRNA genes. The schematics (drawn to scale) summarize the bioinformatic predictions and the results of the experimental mapping. The Smr genes are represented by grey arrows and the flanking ORFs by the dotted black arrows. Numbers indicate coordinates in the *S. meliloti* 1021 genome database. Experimentally determined 5′- and 3′-ends of the Smr transcripts are boxed. 3′-ends of the differentially expressed sRNAs were assigned to the last U in the consecutive stretch after extended stem-loops of Rho-independent terminators, which are denoted by black dots above the horizontal lines. The white arrowhead indicates the processing site for SmrC7. Putative σ70 promoters are indicated by single arrowheads, and putative transcription factors binding sites by double arrowheads (taken from del Val *et al*, 2007).

In 2010, in a new experimental approach for screening sRNAs in *S. meliloti* applied deep sequencing technologies in combination with oligonucleotide microarray and chip hybridizations. This resulted in the identification of 1,125 sRNAs in a size range of 50 to 348 nucleotides and classified as *trans*-encoded sRNAs (173), *cis*-encoded antisense sRNAs (117), mRNA leader transcripts (379), and sense sRNAs overlapping

coding regions (456). These results suggest that in *S. meliloti* about 3% of the genes encode *trans*-sRNAs and about 2% *cis*-sRNAs. Determination of expression patterns provided further information on conditions for expression of a number of sRNA candidates. Sequence conservation analyses suggested strong similarities of a subset of *S. meliloti* sRNAs to regions in related α-proteobacteria (Schlüter *et al.*, 2010; Jiménez-Zurdo *et al.*, 2013). The different types of sRNAs expressed by *S. meliloti* as inferred from *in silico* and experimental approaches are depicted in Figure 1.

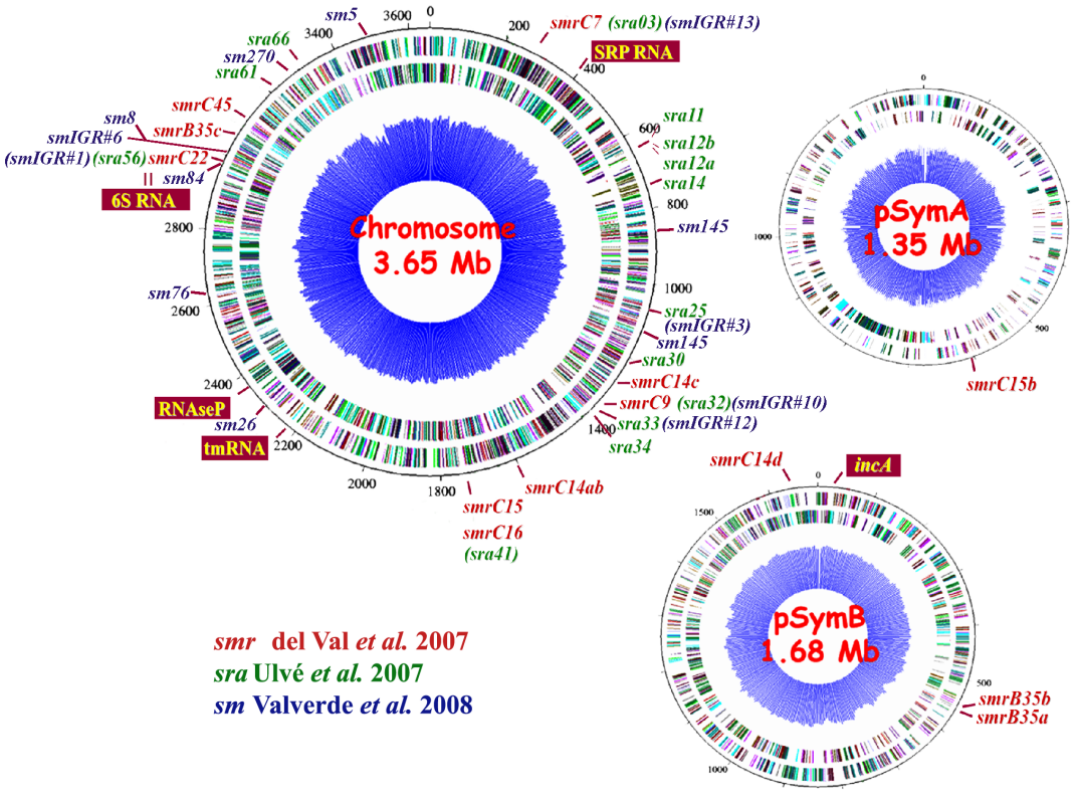

Figura 8. Annotated *S.meliloti* sRNAs. The sRNAs identified in Val *et al.* and subsequently by covariance analysis are in red. sRNAs which have multiple copies in the genome are denoted with lowercase letters. sRNAs identified in Ulvé *et al*. in green. In blue the sRNAs identified in Valverde *et al.* Known sRNAs orthologues are shown in red.

The biological function of the vast majority of sRNAs identified in *S. meliloti* remains unknown. Evidences for biological functions have been provided for only a few members, *i.e.* AbcR1, AbcR2, EcpR1, RcsR1and MmgR. (Torres-Quesada *et al.,* 2013, 2014; Robledo *et al.,* 2015; Lagares, *et al.,* 2017). These sRNAs are widely conserved in related α-proteobacteria and, with the exception of AbcR1, they accumulate in the cell under stress conditions. EcpR1 and RcsR1 target cell cycle and quorum sensing related mRNAs, respectively, in an Hfq-independent manner (Baumgardt *et al.,* 2015; Robledo *et al.,* 2015). AbcR1/AbcR2 and MmgR are all Hfq-dependent sRNAs functionally related to nutrient uptake and carbon metabolism, respectively (Torres-Quesada *et al.,* 2013; Borella *et al.,* 2016; Lagares *et al.,* 2017). However, none of these riboregulators is required for the establishment of an efficient symbiosis.

Besides, the *S. meliloti* genome encodes a recognizable Hfq protein, which function has been investigated in this organism through genetic, transcriptomic and proteomic approaches. Some of these studies demonstrate that the Hfq RNA chaperone is a global regulator of carbon metabolism in *S. meliloti* and contributes to the proper adaptation of these bacteria to the rhizosphere environment and intracellular propagation in the root nodule (Torres-Quesada *et al.,* 2010, 2014; Sobrero *et al.,* 2012; Jiménez-Zurdo *et al.,* 2013). However, recent studies in our laboratory indicated that only a small subset of *trans*-sRNAs expressed by *S. meliloti* bind Hfq (Torres-Quesada *et al*., 2014). Therefore, riboregulation may also involve other protein factors in this bacterium. It has been recently reported that a mutation in an eubacterial conserved gene encoding a homolog of the *E. coli* YbeY protein mimicked several phenotypes of the *S. meliloti hfq* mutants (Pandey *et al.,* 2011), and accordingly it was hypothesized that YbeY may serve Hfq-like roles in RNA-mediated gene regulation in *S. meliloti*.

OBJECTIVES

In the rhizobia-legume symbiosis the transition from a free-living state in soil to an intracellular residence within the plant host demands a flexible adaption of invading bacteria to external changes, which involves the coordinated expression of complex gene networks. Therefore, sRNAs expressed by the microsymbiont are expected to be essential regulatory elements for the integration of diverse soil and plant signals and the coordination of the appropriate physiological bacterial response. Despite this evidence, the role of riboregulation in this mutualistic endosymbiosis has remained largely unexplored. Recently, the non-coding RNome of the model rhizobia *S. meliloti* has been extensively characterized by diverse post-genomic high-throughput approaches. The challenge now is deciphering the function of the thousands of sRNA molecules of different type expressed by this bacterium.

With this background, the following objectives have been proposed in this study:

1. To deepen into the transcriptional regulation and targeting potential of the stress-induced AbcR2 sRNA.

2. To gain primary insights into the function of a novel, yet uncharacterized sRNA referred to as NfeR1.

3. To explore the involvement in riboregulation of the conserved bacterial YbeY protein.

MATERIALS AND METHODS

M-1. BACTERIAL STRAINS USED IN THIS STUDY

Table M-1 lists all bacterial species and strains used in this work. Strains are additionally listed in the appendix of each chapter according to their point of use.

Table M-1. Bacterial strains used in this study.

Strain	Relevant characteristics	Reference/ Source
S. meliloti		
Rm1021	Wild-type SU47 derivative; Nal^r, Sm^r	Meade *et al.*, 1982
Rm2011	Wild-type SU47 derivative; Nal^r, Sm^r	Casse *et al.*, 1979
Sm2B3001	Rm2011 expressing *expR* from the native locus; Nal^r, Sm^r	Bahlawane *et al*., 2008
1021ΔR1	Rm1021 *abcR1* deletion mutant; Sm^r, Er^r	Torres-Quesada *et al.*, 2013
1021ΔR2	Rm1021 *abcR2* deletion mutant; Sm^r, Er^r	Torres-Quesada *et al.*, 2013
1021ΔR1/2	Rm1021 *abcR1/2* deletion mutant; Sm^r, Er^r	Torres-Quesada *et al.*, 2013
VO3128 (*rpoH1*::*aadA*)	Rm1021 carrying an *rpoH1* gene-disrupting construct; $Spec^r$	Oke *et al*., 2001
AB3 (*rpoH2*::*aacCI*)	Rm1021 carrying an *rpoH2* gene-disrupting construct; Gm^r	Bittner and Oke, 2006
AB9 (*rpoH1*::*aadA rpoH2*::*aacCI*)	Rm1021 carrying *rpoH1* and *rpoH2* gene-disrupting constructs; $Spec^r$, Gm^r	Bittner and Oke, 2006
Sm*ΔnfeR1*	Sm2B3001 *nfeR1* deletion mutant; Sm^r, Er^r	This work
1021*hfq*^{FLAG}	Rm1021 carrying *hfq-3xflag* at the native locus	Torres-Quesada *et al.*, 2010
1021Δ*hfq*	Rm1021 *hfq* deletion mutant; Sm^r	Torres-Quesada *et al.*, 2010
Sm*hfq*^{FLAG}	Sm2B3001 carrying *hfq-3xflag* at the native locus	This work
SmΔ*hfq*	Sm2B3001 *hfq* deletion mutant; Sm^r	This work
Sm*ybeY*^{FLAG}	Sm2B3001 carrying *ybeY-3xflag* at the native locus	This work
SmΔ*ybeY*	Sm2B3001 *ybeY* deletion mutant; Sm^r	This work
E. coli		
DH5α	F^-, ø80d*lacZ*ΔM15, Δ(*lacZYA-argF*)U169, *deo*R, *recA*1, *endA*1, *hsdR*17(r_K^-, m_K^+), *phoA*, *supE*44, λ^-,*thi*-1, *gyrA*96, *relA*1	Bethesda Research Lab
S17-1	*recA pro hsdR RP4-2-Tc::Mu-Km::Tn7*	Simon *et al*., 1983

M-2. PLASMIDS USED IN THIS STUDY

Table M-2 lists all plasmid constructs used in this work. Plasmids are additionally listed in the appendix of each chapter according to their point of use.

Table M-2. Plasmids used in this work.

Plasmids	Description	Reference/ Source
pGEM-T Easy	Cloning vector for PCR products; Ap^r	Promega
pSRKKm	pBBR1MCS-2 derivative containing the *lac* promoter, *lacIq*, *lacZα*$^+$, Km^r	Khan *et al*., 2008
pSRK-C	Engineered pSRKKm lacking the LacIQ operator; Km^r	Torres-Quesada *et al*., 2013
pSRK-R2	pSRK_C carrying the *abcR2* coding sequence	Torres-Quesada *et al*., 2013
pSRKNfeR1	pSRK derivative constitutively expressing NfeR1; Km^r	This work
pBlueScript-KS II	Multicopy plasmid for *in vitro* transcription; Ap^r	Stratagen
pKS*nfeR1*	pBlueScript for *in vitro* transcription of *nfeR1;* Ap^r	This work
pK18mobsacB	Suicide plasmid in *S. meliloti*, sacB, oriV, Kmr	Schafer *et al*., 1994
pK18ΔnfeR1	Suicide plasmid for *nfeR1* deletion; Err, Kmr	This work
pK18Δ*hfq*	Suicide plasmid for *hfq* deletion; Km^r	This work
pK18*hfq*^{3xFLAG}	Suicide plasmid for Hfq tagging; Km^r	This work
pK18Δ*ybeY*	Suicide plasmid for *ybeY* deletion; Km^r	This work
pK18YbeY^{3xFLAG}	Suicide plasmid for YbeY tagging; Km^r	This work
pJB3Tc19	Low copy plasmid; IncP, Ap^r, Tc^r	Blatny *et al,* 1997
pJB-T1	pJB3Tc19 derivative in which Plac was replaced by the Rho-independent terminator T1; Apr, Tcr	This work
pJBNfeR1	pJB-T1 derivative expressing NfeR1 from its native promoter; Ap^r, Tc^r	This work
pBBR1MCS-2	pBBR1MCS derived with Km^r	Kovach *et al.*, 1994
pJBYbeY	pJB3Tc19 derivative expressing YbeY from its native promoter; Ap^r, Tc^r	This work
pBB-*egfp*	pBBR1MCS-2 derivative for generation of promoter- *egfp* fusions; Km^r	This work
pBB*syn::egfp*	pBBR1MCS-2 derivative constitutively expressing *egfp* ; Km^r	This work
pR_EGFP	Vector for generating of target mRNA-*egfp* translational fusions	Torres-Quesada *et al*., 2013
pR*oppA::egfp*	pR_EGFP expressing the *oppA::egfp* translational fusion; Ap^r, Tc^r	This work
pR*prbA::egfp*	pR_EGFP expressing the *prbA::egfp* translational fusion; Ap^r, Tc^r	This work

pR*SMa0495::egfp*	pR_EGFP expressing the *SMa0495::egfp* translational fusion; Ap^r, Tc^r	This work
p*nfeR1::egfp*	pBBR1MCS-2 derivative expressing a transcriptional fusion of the *nfeR1* promoter to *egfp*; Km^r	This work
p*nfeR1-40::egfp*	pBBR1MCS-2 derivative expressing a transcriptional fusion of a truncated *nfeR1* promoter (40 nt) to *egfp*; Km^r	This work
pGUS	Vector carrying the GUS reporter gene; Ap^r	García-Rodríguez and Toro, 2000
pRK2013	Mobilization helper plasmid in *E.coli* HB101; Km^r	Ditta *et al.*, 1980

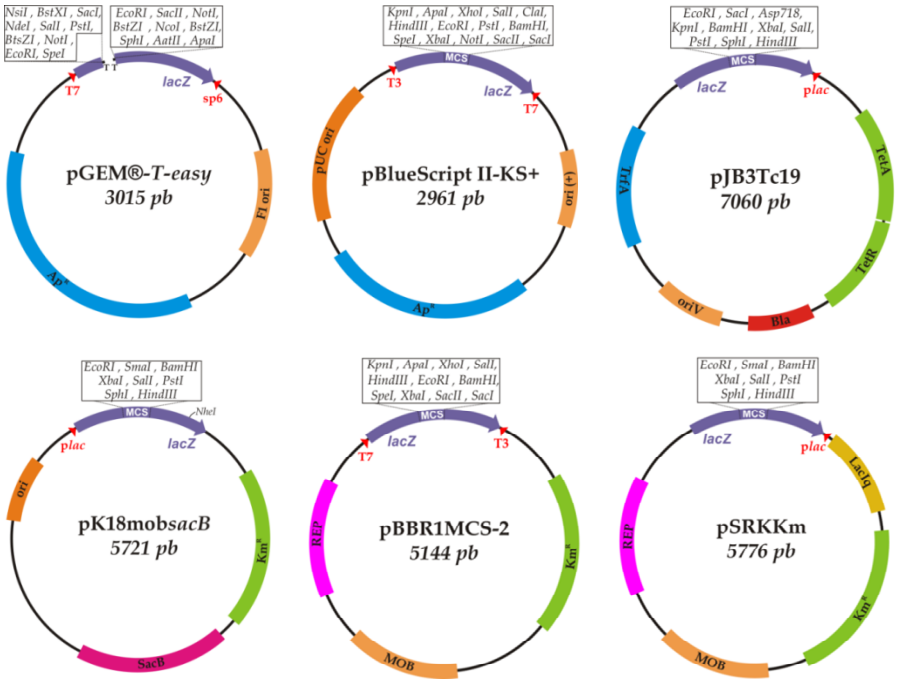

Figure M-1. Principal cloning plasmids used in this work. The red arrows indicate the promoter boxes and the universal primer annealing site.

M-3. OLIGONUCLEOTIDES USED IN THIS STUDY

Table M-3 lists all oligonucleotides used in this work. They are additionally listed in the appendix of each chapter according to their point of use.

Table M-3. Oligonucleotides used in this study.

Name	Sequence (5'-3')
T7 (universal)	TAATACGACTCACTATAGGGCGA
SP6 (universal)	GTATTCTATAGTGTCACCTAAATAGC
T3 (universal)	AATTAACCCTCACTAAAGGGAGA
M13-Forward (universal)	TGTAAAACGACGGCCAGTG
M13-Reverse (universal)	TCATGGTCATAGCTGTTTCC
SmrC14Fw	GGATCCATCGATCGGGCAGCGCAC
SmrC14Rv	GAGCTCGGAGGCAGAAATAAACAA
Smr14C1	AATCGCCTTTATGACGCCCGCCGGT
NfeR1 (Smr14C2)	TCCCGGTTGCCAATCAGATCAAGCA
Smr14C3	CACGGCGCCCGGCATTCGGTCGGTT
Smr14A1	CCACGGCGCAAGACGCCGATCGGTT
Smr14A2	TTCGATATGCGACGCACCTTTCCTC
Smr14B	GGTCAGGATCGAAAGCCCGGCGCAC
AbcR1	ACTGGGAGGAGAACGGAGCAAAGAT
AbcR2	GAGGAGAAAGCCGCTAGATGCACCA
secSRK	TTCCATTCGCCATTCAGGCT
FwSRK	ACTAAAGGGATCCAAAGCTGGAGC
RvSRK	GCTCACAATTGGATCCAACATACGAG
5-C14	TTGTGCAGTGCATCGATCAT
3-C14	TCTAGATTCACGTTGACAGTGCTCTT
5-C14i	GAGGTACCTCGCAGTGAAACCGAGAA
3-C14i	GAGGTACCAACCCCCGGATTTTACCA
T1_F	AGCAAAGAGCCGCCACGGCGCAGCCTCCGCGCCGT GGCGGTTTTTTA
T1_Rv	AGCTTAAAAAACCGCCACGGCGCGGAGGCTGCGCC GTGGCGGCTCTTT
P14C2_H	AAGCTTATTCTGTGATCATTCGGCGC
P14C2Fw	ACTAGTATTCTGTGATCATTCGGCGC
P14C2Rv	TCTAGAGCTGCCCGATCGATGATTGG
P14C2_54	CTAGTGCCCCTGGTAAAATCCGGGGGTTCGGCCTAT ATTCCAATCATCGATCGGGCAGCT
P14C2_54i	CTAGAGCTGCCCGATCGATGATTGGAATATAGGCCG AACCCCCGGATTTTACCAGGGGCA
SecC14	AAACAAGCCGCCCCGGGTAT
SecC14i	GAGGAGTGTTGCCAATCCAT
SMc01642_F	GGATCCGAACAGCGCGGATAACGCGCAA
SMc01642_R	GCTAGCTTTGCCGAGCATGACCTGAC
SMa0495_F	GCTAGCCATTGCAACCGCCGACCCCA
SMa0495_R	GGATCCTAGAAGGCATCGAATTTCCA
SMb21196_F	GGATCCTGGTGCTTCCGTGCAAGCAG
SMb21196_R	GGATCCGTTCGTTTGGGCCTGATATC
5HfqMut	TCTTCATCACCGCTGCTACC
3HfqMut	AACGATCATGCCGTGAACGA

YbeY_F1	ATTGCTGGAAGAGCGATTGC
YbeY_R1	ATCCTCGAGTTGCTCGCGTA
YbeY_iR	GTCCAAAAGCTTCATGATAAACGCGGCCGC
YbeY_iF	GTTGGGAAGCTTTAACAGTTTGGAACGATG
YbeY_RE	CGGAATTCATCCTCGAGTTGCTC
YbeY_FS	GCGAATTCTGACGTCGTCGCAACCA
YbeY_F2	CGCGTTTCATATGACGGCATTGG
YbeY_R2	CGTTCCGGATCCTTAATGCGGG
YbeY_MutF	GAGGCGCTGCAGATACTCAA
YbeY_MutR	GATGATGTGGATTTGCTGCC
YbeY_PrF	AAGCTTGATGTTCCTGACCCGTCTCG
YbeY_PrR	GAATTCCCGGCTGTGTCTTGAAGTCG
YbeY_XbaI	TCTAGAATGCGGGGGTTGGTCCCC
YbeY_RK	GGTACCATCCTCGAGTTGCTCGCGTA
5'-*ybeY*Mut	GAGGCGCTGCAGATACTCAA
3'-*ybeY*Mut	GATGATGTGGATTTGCTGCC
nifAFw	TCGTCTTGAGACCACGCTTA
nifARv	CATGACTTGGTCTATTGCGG
fixKFw	TCCATCGAGGTCGAACACCT
fixKRv	CATTTCGCCTGGGAGATGAA
16SFw	GGCTAGCGTTGTTCGGAATT
16SRv	TCCGATCCAGCCGAACTGAA

M-4. MICROBIOLOGICAL TECHNIQUES

M-4.1. CULTURE MEDIA AND ANTIBIOTICS

M-4.1.1. Culture media

M-4.1.1.1. Complete medium for *S. meliloti* (TY)

The complete medium used for the routine growth of *S. meliloti* was TY (Beringer, 1974) prepared in deionized water:

$CaCl_2 \cdot 2H_2O$	0.9 g/l
Tryptone (DIFCO)	5 g/l
Yeast extract (DIFCO)	3 g/l

For the preparation of the solid medium, **1.6% agar** (PANREAC) was added. The sterilization was performed in an autoclave for 20 min at 120 °C.

M-4.1.1.2. Minimal medium for *S. meliloti* (MM)

The following modification of the minimal medium (MM) described by Robertsen *et al.*, 1981 was prepared in deionized water to grow *S. meliloti*:

K_2HPO_4	0.3 g/l
KH_2PO_4	0.3 g/l
$MgSO_4 \cdot 7H_2O$	0.15 g/l
$CaCl_2$	0.05 g/l
$FeCl_3$	0.006 g/l
NaCl	0.05 g/l
Sodium glutamate	1.1 g/l
Manitol	10 g/l
Biotin	0. 2 mg/l
Calcium Pantothenate	0. 1 mg/l

To solidify the medium, **1.6% purified agar** (OXOID) was added. The pH was adjusted to 6.8-7.0 and the medium was autoclaved for 20 min at 120 ° C.

M-4.1.1.3. Complete medium for *E. coli* (LB)

The Luria-Bertani medium (LB) (Sambrook *et al.*, 1989) was prepared with deionized water and used as the routine culture medium for *E. coli* strains:

NaCl	5 g/l
Triptone (DIFCO)	10 g/l
Yeast extract	5 g/l

To solidify the medium, **1.6% agar** (PANREAC) was added. The medium was autoclaved at 120 °C for 20 min.

As a restrictive culture medium for coliforms, **Endo Agar ™** LES (DIFCO) was used. It was used for the verification of possible contaminations by *E. coli* after the conjugations. It was prepared by dissolving 51 g per liter of deionized water, heating and sterilization in an autoclave at 120 °C for 20 min.

M-4.1.2. Antibiotics

Addition of antibiotics to the culture media was done from 100-fold concentrated solutions prepared in deionized water (Km, Ap) or in water-50% EtOH (Tc). Solutions prepared in water were sterilized by filtration using Minisart® NML (Sartorius) units of 0.2 μm pore size. The concentrations used for the different antibiotics are shown in Table M-4.

Table M-4. Antibiotic concentrations used in this work for *E. coli* and *S. meliloti*.

Antibiotic	Effect	*E.coli*	*S. meliloti*
Streptomycin. Sm (Roche)	Bactericide	200 mg/l	200 mg/l
Ampicillin, Ap (Sigma)	Bacteriostatic	200 mg/l	200 mg/l
Kanamycin, Km (Sigma)	Bactericide	50 mg/l	180 mg/l
Tetracyclin, Tc (Sigma)	Bactericide	10 mg/l	10 mg/l
Rifampicin, Rf (Sigma)	Bactericide	50 mg/l	50 mg/l
Erythromycin, Er (Sigma)	Bacteriostatic	60 mg/l	100 mg/l
Nalidixic acid, Nal (Sigma)	Bactericide	-	10 mg/l
Gentamicin, Gm (Sigma)	Bactericide	50 mg/l	200 mg/l
Spectinomycin, Spc (Sigma)	Bacteriostatic	50 mg/l	200 mg/l

M-4.2. CULTIVATION CONDITIONS

The different ***S. meliloti*** strains were grown at 28 °C. Its generation time in liquid medium was approximately 2.5 h at 180 rpm, in orbital shaker.

E. coli was incubated at 37 °C and its generation time in liquid medium was about 30 min at 180 rpm, in orbital shaker.

M-4.2.1. Expression in growth phase

Exponentially and **stationary** growing bacteria were obtained by incubation in the selected growth medium in each case to OD_{600} 0.6 and 2-2.4, respectively.

Growth curves were recorded with a microplate reader (Victor X3 Multilabel Plate Reader, PerkinElmer).

For **phenotypic analysis**, growth of the tested strains in the selected media was assessed by absorbance measurement (OD_{600}) and bacterial plate counting. Bacterial growth on agar was also analyzed.

M-4.2.2. Expression during abiotic stresses

For the characterization of gene expression in *S. meliloti*, several abiotic stresses were mimicked by modifying the selected medium as follows:

Osmotic stress:

The **osmotic upshift** was imposed by adding 400 mM NaCl to exponential cultures, which were incubated for a further 1 h upon stress exposure. This salt concentration and incubation time of the bacteria represent those conditions which were shown to

trigger most changes in the transcriptome of *S. meliloti* in response to the saline stress (Domínguez-Ferreras *et al*., 2006). **Moderate salinity** was also generated by supplementing TY or MM media with 50 mM NaCl in which the bacteria were grown from the beginning of the experiment.

Oxidative stress:

To mimic oxidative stress conditions, 1 mM H_2O_2 was added to exponential cultures, which were incubated for a further 1 h upon stress exposure. This concentration of H_2O_2 can be considered sublethal and its effect on the adaptation of bacteria to oxidative stress was previously experimentally determined in our laboratory.

Acid/alkaline pH stress:

Acid and alkaline stresses were generated by resuspension of bacteria (grown to OD_{600} 0.6) in MM, modified by addition of either **HCl to pH 5.8** or **NaOH to pH 8.5**. In all cases cultures were incubated for a further 1 h upon stress exposure.

Cold/heat stress:

Cold and heat shocks were applied to exponentially growing bacteria in the selected medium during 1 h by shifting the temperature from 28 °C to **20 ºC** and **42 ºC**, respectively.

Low oxygen tension (microaerobiosis):

In order to generate conditions with low tension of oxygen, 20-ml cultures in 100-ml Erlenmeyer flasks grown in TY medium to an OD_{600} of 0.6 were subjected to a

mixture of **2% Oxygen/98% argon** for 10 min and incubated for a further 4 h in that atmosphere. This condition was previously verified in our laboratory to induce microaerobiosis, essentially shown by measuring β-galactosidase activity in *S. meliloti* bearing transcriptional fusions of the *fixK* and *nifA* promoters to the *lacZ* reporter gene (Ditta *et al.*, 1987; Foussard *et al.*, 1997)

M-4.3. CONSERVATION OF BACTERIAL CULTURES

Freezing was used for the long-term preservation of bacterial cultures. For maintenance of cell viability during the storage period, 20% (v/v) glycerol was used as cryoprotectant. Sterile glycerol in cryotubes was mixed with cultures grown to late logarithmic phase. The vials were rapidly frozen and stored at -80 °C.

M-4.4. MOBILIZATION METHODS OF EXOGENOUS DNA IN BACTERIA

M-4.4.1. Transformation of *E. coli*

M-4.4.1.1. Preparation of competent cells

A type of cell, ready to capture exogenous DNA, is used to introduce circularized DNA. Its preparation requires chemical or physical methods that alter the cell wall of the bacteria.

M-4.4.1.1.1. Chemical methods (RbCl):

E. coli cells were cultured in 100 ml of LB medium until log phase (OD_{600} 0.4). Growth was then stopped by incubating the culture on ice for 15 min. After this, the

bacteria were pelleted by centrifugation (6000 rpm, 10 min, 4 °C). Carefully, the pellet was resuspended in 32 ml of sterile, pre-cooled RF1 solution [RF1 (per 100 ml): 1.2 g RbCl; 0.99 g $MnCl_2 \cdot 4H_2O$; 0.294 g potassium acetate; 0.15 g $CaCl_2 \cdot H_2O$; 11.9 ml glycerol; distilled water to 100 ml, pH 5.8]. The bacterial suspensions were incubated on ice for 15 min and then centrifuged at 6,000 rpm for 10 min. The supernatant was removed and the cells were resuspended in 4 ml of RF2 solution precooled at 4 °C [RF2 (per 50 ml): 0.1046 g morpholino propanesulfonic acid (MOPS); 0.06 g RbCl; 0.55 g $CaCl_2 \cdot H_2O$; 5.95 ml of glycerol; distilled water to 50 ml, pH 6.8]. The bacterial suspensions were aliquoted into 100 µl aliquots in pre-cooled tubes at 4 °C and immediately frozen in liquid nitrogen. Competent cells were kept for a limited time period at -80 °C. The transformation efficiency ranged from 1 to $5*10^5$ cells/µg DNA.

M-4.4.1.1.2. Electrocompetent cells:

For this protocol, the complete removal of salts from the culture medium was required to avoid disturbance during the electroporation process. In addition, a cryoprotective agent (glycerol) was added which ensured viability of the conserved cells at -80 °C. It consists in successive washes with decreasing volumes of 10% glycerol; we proceed from 500 ml of culture to 0.5 OD_{600}, whose metabolism is delayed by incubation on ice for 20 min. Cells were collected at 6,000 rpm for 15 min and 4 ° C, resuspended in 100 ml of 10% cold glycerol and repeatedly centrifuged. In the next step, cells were resuspended in 20 ml of glycerol to finish, after this second washing, with the bacteria resuspended in 2 ml of 10% glycerol. Cells were distributed in 50 µl aliquots which are immediately frozen in liquid N_2. The cells were stored at -80 °C for a limited period. The transformation efficiency was approximately 10^6 cells/µg DNA.

M-4.4.1.2. Transformation of competent cells

M-4.4.1.2.1. Transformation by thermal shock

Competent cells prepared by chemical methods were thawed on ice (~20 min). DNA was added under sterile conditions and incubated with the cells for further 20-25 min. The cells were subjected to a thermal shock of 42 ° C for 90 s, then returned to ice for 5 min to allow DNA uptake by the cells. 900 µl fresh LB was then added and cells were incubated for 1 h at 37 ° C (under slow shaking conditions) for the recovery of the cells. After this, the bacterial suspensions were dispensed on selective LB agar.

M-4.4.1.2.2. Electrotransformation

In this protocol, the cells had also to be thawed on ice. After that, the DNA to be transformed was added. The cell suspension was transferred to an electroporation cuvette and subjected to an electrical pulse according to the manufacturer's guidelines. We used an Eppendorf 2510 electroporator set up to a pulse of 1800 V for 3-5 ms. The subsequent procedure equaled that of chemically competent cells as described in the previous section. It was important that the applied DNA was free of salts. To achieve this, digestion or ligation reaction mixtures were dialyzed against water using 0.025 mm VSWP nitrocellulose filters (MILLIPORE).

M-4.4.2. Conjugational transfer of DNA to *S. meliloti*

The mobilization of plasmids between the *S. meliloti* receptor strain and the *E. coli* donor strain was performed mainly by triparental mating (Ditta *et al*., 1980). In this type of conjugation, the mobilization genes present in the *E. coli* HB101: pRK2013 helper strain are required. In the case of biparental mating, *E. coli* S17-1 (Simon *et al*., 1983) was used as the donor strain.

Conjugations were performed by mixing equal amounts of the bacterial strains and incubating them on solid TY medium at 28 °C for 12-16 h. Transconjugants were selected on MM agar supplemented with the appropriate antibiotics Cells were grown at 28 °C until appearance of visible colonies (after 2-3 days) and a new selection was made on MM agar. Possible contamination with coliforms was checked in parallel through the use of solid Endo-Agar. In addition, several cycles of cultivation were performed on TY agar supplemented with nalidixic acid and streptomycin (increased to 600 mg/l) to eliminate any present *E. coli*.

M-4.4.3. Selection of allelic exchange and co-integration

The introduction of genomic modifications was performed by double recombination according to the method described for Gram negative bacteria (Schäfer *et al*., 1994). For this purpose, an internal fragment of the gene to be mutated was modified/deleted, previously cloned into the pGEM®-T-Easy plasmid with the appropriate restriction sites. Procedures and design of these constructions is described in the section on molecular biology techniques in this chapter.

Subsequently, the fragment with the mutated gene was subcloned into the suicide plasmid pK18mobsacB, which was then conjugated to *S. meliloti* by triparental mating. Transformants carrying the plasmid in the chromosome were selected by growth on both Km-supplemented TY agar and TY agar containing 10% sucrose (Panreac).

To select for double recombinants, colonies showing sucrose sensitivity were cultured in liquid TY medium at 28 °C and grown to an OD_{600} of 1. From this culture, dilutions were dispensed on TY agar containing 10% sucrose. The raised colonies were replicated in TY medium with and without Km, finally to select those ones sensitive

to Km. Loss of the original copy of the gene in the selected colonies was subsequently confirmed by PCR and Southern blot.

In the case of the presence of resistance markers in the strains carrying the modified allele, the corresponding antibiotic (erythromycin) was added to the selective medium of the second selection at a concentration of 100 mg/ml.

M-5. MOLECULAR BIOLOGY TECHNIQUES

M-5.1. DEOXYRIBONUCLEIC ACIDS (DNA) MANIPULATION

M-5.1.1. DNA isolation

M-5.1.1.1. Extraction of total DNA with a commercial kit

For this purpose the commercial kit *Real Pure Extrac. DNA Genómico* (REAL). 1 ml of *S. meliloti* culture was collected and washed with 0.2 ml of 0.1% sarcosyl in TE. Lysis was performed by adding 0.6 ml of "Lysis Solution" and incubating for 5 min at 80 °C. The cell lysate was cooled to RT and then 3 µl of "RNAse Solution" was added. The samples were mixed by inverting and then incubated for 1 h at 37 °C. After this, 0.2 ml of "Protein Precipitation" solution was added and the mix was shaken vigorously for 20 s. After incubation on ice for 10 min, the samples were centrifuged at 13,000 rpm speed for 5 min and the supernatant was collected in a new tube. The DNA was precipitated by adding 0.6 ml of cold isopropanol, inverting the whole mixture and incubation for 10 min at -20 °C. After centrifugation for 10 min, the pellet was washed with 0.2 ml 70% EtOH and resuspended in 50-100 µl of distilled water.

M-5.1.1.2. Colony lysis

Colony lysis is a very quick strategy to obtain DNA in a sufficient quality for its PCR amplification. ***E. coli*** colonies grown on selective LB agar were picked with a sterile toothpick, stirred up in 25 µl PCR reagent mixture and subjected to a standard PCR program.

In the case of ***S. meliloti***, the selected colonies were picked under sterile conditions and resuspended in 100 µl of 0.1% sarcosyl in TE. The cells were pelleted by centrifugation for 2 min at 13,000 rpm, the supernatant was discarded and the cell pellet was resuspended in the same volume of double-distilled water (miliQ). The mixture was boiled for 3 min and then centrifuged again. The supernatant was used in PCR reactions.

M-5.1.1.3. Extraction of plasmid DNA by alkaline lysis

Isolation of plasmid DNA was performed following the protocol modified by Sambrook *et al.* (Birnboim and Doly, 1979; Ish-Horowicz and Burke, 1981; Sambrook *et al.*, 1989). 1.5 ml of bacterial culture grown to exponential phase were collected and centrifuged at 13000 rpm for 2 min. The cells were washed with 500 µl of 0.1% sarcosyl in TE and centrifuged again. After discarding the supernatant, the cells were resuspended in 100 µl of 4 mg/ml lysozyme dissolved in solution I (50 mM glucose, 25 mM Tris-HCl pH 8, 10 mM EDTA, stored at 4 ° C). After incubation for 5 min at RT, 200 µl of solution II (0.2 M NaOH, 1% SDS) was added. The preparation was homogenized by shaking the tube inversely several times and incubation on ice for 5 min. 150 µl of solution III (5 M potassium acetate pH 4.8) was then added for neutralization and precipitation of proteins. The tubes were immediately inverted and incubated on ice for 5 min to minimize DNA degradation by cellular DNases. The cell debris was sedimented by centrifugation for 5 min at 13,000 rpm and the supernatant

was collected in a new tube. A 25:24:1 (ΦCIA) Phenol:Chloroform:Isoamyl Alcohol volume (~400 µl) was added and the lysate was mixed vigorously. After centrifugation for 5 min at 13,000 rpm, the upper aqueous phase was collected and mixed with an equal volume of ΦCIA (eliminates possible phenol residues in the preparations). The mixture was centrifuged for 5 min and the aqueous phase was transferred to a fresh tube. Precipitation of plasmid DNA was achieved by adding 2.5 volumes of cold 100% EtOH and incubating the lysate for 5 min at toom temperature. The precipitated DNA was centrifuged and the pellet was washed with 200 µl of 70% EtOH. After drying, the pellet was resuspended in 25 µl of deionized water containing 10 µg/ml RNase A. RNase solution was prepared by diluting the enzyme (lyophilized, Sigma) in a solution containing 10 mM Tris-HCl pH 7.5 and 15 mM NaCl. The DNA extracts were boiled for 15 min, cooled at RT and finally stored at -20 °C.

M-5.1.1.4. Minipreparation of plasmid DNA by precipitation with magnesium salts

This technique is based on the method of Studier, 1991 that allows the rapid isolation of *E. coli* plasmids. 1.5 ml of bacterial cultures was collected by centrifugation and, after removing the supernatant, the cells were resuspended in 100 µl of deionized water. Lysis was achieved by adding 100 µl of a solution containing 0.1 M NaOH, 10 mM EDTA and 2% SDS. The mixture was homogenized by stirring and boiled for 2 min. 50 µl of 1 M $MgCl_2$ was added and stirred vigorously. After complete homogenization (appearance of white precipitate), the mixture was incubated on ice for 2-5 min and centrifuged for 1 min at 13,000 rpm. 50 µl of 5 M potassium acetate pH 4.8 were added and the mixture was inversely vortexed to avoid detachment of the sediment. Then, the samples were incubated for 5 min on ice and centrifuged for 5 min. The supernatant was transferred to a new tube and mixed with 0.6 ml of -20 °C pre-cooled 100% EtOH to precipitate plasmid DNA. The mixture was incubated at RT

for 5 min and centrifuged for 5 min. The pellet was washed with 200 µl of 70% EtOH, dried in vacuum and finally resuspended in deionized water containing 10 µg/ml RNase.

M-5.1.2. Evaluation of the quantity and quality of extracted DNA

The calculation of the concentration and purity of the nucleic acids was estimated following the spectrophotometric method described in Sambrook *et al.* (1989) and using a NanoDrop® ND-100 spectrophotometer, which integrates all these data from 1-2 µl of isolated DNA. The ratio of absorbance values at 260 nm and 280nm was used as an estimation of the DNA purity, with values below 1.8 being considered as indicators of protein or phenol contamination.

M-5.1.3. DNA digestion by endonucleases

Complete digestion of DNA was carried out at optimal temperature and buffer conditions according to the manufacturer's guidlines (Roche and New England Biolabs). Briefly, 1-5 U of enzyme/µg DNA was used for digestion. In the case of simultaneous digestions with two restriction enzymes, a buffer compatible for both (supplier's instructions) was used, whereas in the case of incompatibility, digestion was performed sequentially. Salts were removed by the use of commercial kits, such as QIAquick PCR Purification Kit (Qiagen), or by phenolization and precipitation of the DNA.

M-5.1.4. Dephosphorylation reaction

In the case of compatible cohesive or blunt ends, the removal of phosphate groups present at the 5'-ends of linearized plasmid DNA is useful to prevent re-ligation. Dephosphorylation of 5 µg of digested plasmid was carried out using 1 U of alkaline phosphatase (CIP, Calf Intestine Phosphatase; Roche) in the supplied buffer for 1 h at 37 C. The reaction was stopped by adding 55 µl of a solution containing STE 1X buffer [1 mM Tris-HCl pH 8, 100 mM NaCl and 1 mM EDTA] and 1% SDS, followed by incubation for 15 min at 68°C. To obtain higher ligation efficiency, phenolization and precipitation of the DNA was performed.

M-5.1.5. PCR (polymerase-chain-reaction)

Depending on the purpose, different polymerases were used: Phusion® DNA polymerase (NEB), Accuprime (Invitrogen), Certhamp® (Biotools) or Pfu polymerase (Biotools), all representing high-fidelity polymerases. For routine testing, a Taq polymerase purified in our laboratory was used. The amplification was performed with an Eppendorf Mastercycler® thermocycler, and the program used varied depending on the objective.

The reactions were generally performed in a final volume of 25 µl containing 0.1 to 100 ng of template DNA, 25 pmol of each of the specific primers, 0.1 mM deoxynucleotide triphosphate (increased to 1 M for templates like total DNA or cell lysates), the buffer specified for each application and 2 units of DNA polymerase.

Purification of amplified products (primer and salt removal) was performed using MicroSpin™ S-300HR (GE Healthcare) columns or using the QIAquick PCR Purification Kit (Qiagen).

M-5.1.6. DNA cloning

This section includes different protocols commonly used for DNA cloning.

M-5.1.6.1. Isolation of DNA fragments from agarose gels

For the isolation of restricted DNA fragments separated by horizontal agarose gel electrophoresis, the commercial QIAquick Gel Extraction Kit (Qiagen) was used following the manufacturer's instructions. To this end, the DNA was stained with either ethidium bromide or Bio-safe gel Red®, visualized with an UV trans-illuminator (taking care of minimized exposure of DNA to UV light) and cut from the gel.

M-5.1.6.2. Ligation of digested DNA fragments

The covalent linkage of two linear DNA molecules was performed by the addition of 1 U T4 DNA ligase (Roche) in the supplied buffer. The most common molar ratio of vector to insert was 1:3. The reactions were prepared in a final volume of 10-20 µl and incubated at 14-16 °C overnight.

M-5.1.6.3. Adenylation of PCR products

For the ligation of PCR-amplified DNA, vectors which allow the cloning of fragments with protruding adenine ends produced by the Taq polymerase are available. Nevertheless, some polymerases with 3'-5' exonuclease activity produce blunt ends. In these cases, it is necessary to adenylate the PCR product. The adenylation reaction is

catalyzed by Taq polymerase (5 U) in the presence of 0.2 mM dATP (70 °C for 30 min). This product can be used directly in the ligation reaction.

M-5.1.7. Sequencing and sequence analysis

For the classic Sanger sequencing, the DNA sequencing/genomics service of the "Estación Experimental del Zaidín" or the Institute of Parasitology and Biomedicine "López-Neyra" was used. These services amplify samples with a PE-9600 thermocycler (Perkin-Elmer) and perform sequencing with the ABI 373 XL Stretch (Perkin-Elmer) sequencer using the ABI PRISM Big Dye™ Terminator Cycle Sequencing Ready Reaction (Perkin-Elmer) kit.

Sequence peaks were displayed using the program Chromas Lite 2.01 and Sequence Scanner v1.0. Routine work with sequences (restriction target search, primer design, sequence comparison, ORF search, etc.) was performed with the Clone Manager Professional Suite® program.

M-5.1.8. DNA electrophoresis

DNA electrophoresis is used for separating macromolecules according to their size and charge. Nucleic acids, with negative charge, migrate towards the positive pole in the applied electric field.

M-5.1.8.1. Electrophoresis in non-denaturing agarose gels

Horizontal electrophoresis was run in 1% agarose gels (*SeaKem*® agar, Cambrex / Iberlabo) prepared with 1x TAE: [40 mM TrisHCl, 2 mM Na_2EDTA, 0.11% (v/v) glacial acetic acid]. The applied voltage was 120 V, although certain procedures

required reduced voltage. A 6x solution (1 µl per 5 µl sample) of 0.25% (w/v) bromophenol blue and 45% (v/v) glycerol was used as the loading buffer.

DNA staining was performed by immersing the gels in a solution containing either ethidium bromide or 1 mg/ml bio-safe *gel Red*® in water for 15-20 min. These intercalants allow visualization of DNA or RNA upon exposure to UV light of 365 nm of wavelength. Image acquisition and processing for the gels was performed with *Gelprinter* and *Gelstation* from TDI.

M-5.1.8.2. Electrophoresis in non-denaturing polyacrylamide gels

For separation of DNA fragments smaller than 200 bp, vertical electrophoresis with polyacrylamide gels was used. In this case, the electrophoresis buffer was 1X TBE [0.089 M Tris-HCl pH 8, 0.089 M boric acid and 0.002 M EDTA]. We worked with 40% (29:1) solutions of acrylamide:bis-acrylamide (Bio-Rad). The percentage and thickness of gels varied depending on specific purpose, while amounts of added polymerizing agent (TEMED, Sigma) and the catalyst (ammonium persulfate, APS; Sigma) were adjusted accordingly. The voltage at which the electrophoresis was performed ranged from 150-200V.

Staining with silver nitrate (Heuer *et al*., 1997) was used for the development of these gels. The gel was fixed with a solution containing 10% (v/v) EtOH and 0.5% (v/v) glacial acetic acid (two times for 3 min).Then, the gel was incubated in a 0.2% silver nitrate solution in water and under agitation for 30 min. It was washed four times with water to remove excess of silver and incubated in a developing solution [10% (w/v) sodium tetrahydroborate, /0.4% (v/v) formaldehyde, /1.5% (w/v) NaOH] until appearance of visible bands. The gel was then washed with distilled water and incubated in a solution containing 0.75% (w/v) sodium carbonate. Prior to drying, the gel was kept in a solution of 25% EtOH and 10% glycerol overnight.

M-5.1.9. DNA molecular weight markers

Molecular weight markers were in all cases supplied by companies:

- **Marker II**: λ phage DNA digested with the HindIII enzyme. It consists of eight fragments ranging from 100 bp to more than 20 kbp (Roche, Universidad Autónoma de Madrid).
- **Marker III**: phage λ DNA digested with HindIII and EcoRI enzymes. In this case, the pattern of digestion consists of 13 fragments ranging from 100 bp to more than 20 kbp (Roche).
- **Marker Φ29**: phage DNA with the same name digested with HindIII that releases 14 fragments: 72, 156, 273, 453, 579, 611, 759, 1150, 1331, 1933, 2201, 2498, 2899 and 4370 bp (Universidad Autónoma de Madrid).
- **Marker pGEM**: pGEM-T plasmid DNA digested with the HinfI and EcoRI enzymes, which generate 15 fragments ranging from 36 bp to ~2500 bp (Promega).

M-5.1.10. DNA-DNA hybridization (Southern Blot)

Enzyme-digested and size-fractionated total DNA in agarose gels was transferred and fixed to a positively charged Biodine® membrane. An alkaline transfer system (1 M NaOH) was used for this purpose (VacuGene™ XL system from pharmacies). Prolonged exposure (15-20 min) of the gels to UV light in the trans-illuminator was used to break the DNA. To avoid salt deposition, the membrane was soaked in deionized water and then equilibrated with 20x SSC solution [1x SSC: 150 mM NaCl, 15 mM sodium citrate, pH 7.0] for 5 min. The transfer system was prepared according to the manufacturer's instructions and avoiding the formation of air bubbles between the membrane and the gel. Vacuum was generated in the transfer system, which was

covered with 1 M NaOH. A constant pressure of around 55 mbar was applied for 1-2 h. The membrane was then washed with a 2x SSC solution for 10 min under gentle agitation and dried at RT. To attach the DNA to the membrane, the membrane was wrapped in a 3MM Whatman paper, and we applied heat (120 °C for 35 min) and vacuum (70 cm Hg). The membrane was stored at RT until imaging.

M-5.1.10.1. Radioactive labeling DNA probes

In this type of experiments, the labeling is performed with dCTP [α-^{32}P] (PerkinElmer). The probe was synthesized using the commercial kit "*MegaPrime™ DNA labeling system*" from Amersham-GE healthcare™.50 ng of the probe were taken and annealed for 10 min at 65ºC with random primers provided in the kit. The radioactive isotope was added to the annealed product by the addition of 1 U Klenow polymerase and subsequent incubation for 1 h at 37 ° C. The labeled probe was purified with S300 columns (GE-healthcare) and labeling efficiency (counts per minute, cpm) was estimated with a Scintillation Counter "Liquid Scintillation Analyzer Tri-Carb 1500, Δ Packard".

M-5.1.10.2. Hybridization, washing and developing

This process was executed in a Hybridiser® HB-1D (Techne) hybridization oven, following the protocol of Church and Gilbert, 1984. The filter was incubated in a prehybridization solution (20 ml/250 cm^2) [0.5 M Na_2HPO_4 pH 7.2, 0.34 % (v/v) H_3PO_4 (*Church Buffer),* 1 mM EDTA, 0.5 % (w/v) SDS], for 2 h at 68 °C. The probe was denatured for 10 min at 95 °C and then transferred to ice (5-10 min). The entire amount of the denatured probe (50 ng) was added to the prehybridization solution and incubated at 68 °C overnight.

On the next day, washing and developing of the membrane were performed. A quick first wash was made with 2x SSC/0.1% (w/v) SDS solution to remove the unincorporated probe. Then, two successive washing steps with 1x SSC/0.1% (w/v) SDS and 0.1x SSC/0.1% (w/v) SDS), each for 15 min at 68 °C, were performed to remove probe bound non-specifically to the membrane.

For developing, the membrane was wrapped in transparent plastic and exposed to a screen (*Imagin Plate 2040, Fujifilm*), either for 4 h (short exposure) or 24 h (long exposure).

M-5.2. RIBONUCLEIC ACIDS (RNA) MANIPULATION

M-5.2.1. Extraction of RNA from *S. meliloti*

M-5.2.1.1. Extraction of total RNA from bacterial cultures

Growth of 25 ml of bacterial cultures was topped by addition of 1/5 volumes of STOP solution (95% EtOH, 5% phenol). Cultures were then centrifuged for 10 min at 5000 rpm and 4 °C. Pellets were frozen in liquid nitrogen and stored at -80 °C.

Pellets obtained from exponentially growing cultures were resuspended in 1 ml of lysis solution (1.4% SDS, 4 mM EDTA, 500 µg proteinase K) preheated at 65 °C (3 ml for stationary growth phase cultures). Bacterial suspensions were incubated at 65 °C for 10 min, vortexed and placed on ice. 500 µl (1.5 ml for stationary growth phase cultures) of 5 M NaCl at 4 °C was added to the cell lysates for precipitation of proteins. After incubation for 10 min on ice, the cell debris was pelleted by centrifugation for 20 min at 10,000 rpm and 4 °C. The aqueous phase was transferred to a new tube and precipitated with 4.5 ml of cold 100% EtOH (7 ml for stationary growth phase cultures). After incubation at -80 °C for at least 1 h, cell lysates were

centrifuged for 30 min at 10,000 rpm and 4 °C and EtOH was removed prior to DNase I treatment.

M-5.2.1.2. Extraction of total RNA from nodules

1-2 g of nodules was pulverized in a mortar with liquid nitrogen. Then, they were thawed and lysed directly in the mortar by adding 4 ml of NTES (100 mM NaCl, 10 mM Tris-HCl pH 7.5, 1 mM EDTA, 1% (w/v) SDS, 100 mM β-mercaptoEtOH). The lysate was transferred to a 15 ml Falcon tube and phenolized with Phenol:Chloroform:Isoamyl Alcohol (ΦCIA) (pH 8 in Tris Buffer). The mixture was slightly vortexed and centrifuged for 15 min at 5,000 rpm and 4 ° C. The aqueous phase was transferred to a fresh tube and mixed with 300 µl of 3 M AcNa (pH 5.2) and 5 ml isopropanol. After incubation at -80 °C for at least 1 h, lysates were centrifuged for 20 min at 10,000 rpm prior to DNase I treatment.

M-5.2.1.3. DNase I treatment

The precipitates obtained from the lysis (both, bacteria or nodules) were resuspended in 440 µl milliQ water, transferred to 1.5 ml tubes and mixed with 50 µl of DNase 10x buffer (500 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$), 5 µl RNase inhibitor RNasin 40U/µl (Roche) and 5 µl DNase I (Roche). Reaction mixtures were incubated for 1 h at 37 ºC.

After the reaction was phenolized with 1 volume of ΦCIA pH 4.3 and centrifuged for 10 min at 13,000 rpm and 4 °C, the aqueous phase was transferred to a new tube. A further 500 µl extraction was performed with cold isoamyl chloroform (24:1), centrifuged 5 min at 13,000 rpm and4 °C, followed by collecting the aqueous phase and dividing it into two 1.5 ml tubes. The RNA was precipitated by adding 20 µl of

3 M AcNa pH 5.2 and 1 ml of cold 100% EtOH. After incubation at -80 °C for at least 1 h, the reaction mixtures were centrifuged for 30 min at 13,000 rpm and 4 ° C. The pellets were then washed with 500 µl of 70% EtOH. After centrifugation for 10 min, the pellet was dried at RT using a *speedvac* and resuspended in 50-100 µl of milliQ water. A NanoDrop® ND-1000 spectrophotometer was used to determine the RNA concentration in each sample.

M-5.2.1.4. Extraction of total RNA from *S. meliloti* with a commercial kit

Efficient labeling of the RNAs used for the hybridization of the microarrays included use of the commercial kit *RNeasy Protect Bacteria Mini Kit* (Qiagen) for extraction of total RNA.

Since *S. meliloti* is a mucosal bacterium100 µl of 0.4 mg/ml lysozyme in TE was added to 1.5 ml of culture in order to digest the cellular cell wall and thus improve cell lysis. The reaction was incubated for 5 min at RT followed by addition of 350 µl RLT buffer (mixed with β-mercaptoethanol according to the manufacturer´s instructions), vigorous shaking and centrifugation for 2 min. The supernatant was transferred to a new tube supplemented with 250 µl of 100% EtOH. The solution was applied to the provided columns (one column per 3 ml of culture) and centrifuged for 30 s at 13,000 rpm. The columns were washed once with 700 µl of RW1 buffer, twice with 500 µl of RPE buffer and centrifuged for 2 min to remove residual EtOH. Finally, the RNA was eluted in two steps by addition of 30 µl of water to the resin, incubation for 1 min at RT and centrifugation for 1 min.

The technology on which the column is based minimizes the amount of DNA attached to the resin. However, this method required RNA completely free of DNA (which may interfere in the efficient hybridization of microarrays). To achieve this, digestion of remaining DNA present in the isolated RNA was performed with DNase I (Qiagen). 14 µl RDD buffer and 6 µl DNase I were added to 120 µl isolated RNA and

the reaction mixtures were incubated for 1-2 h at 28 °C. To the 140 µl reaction, 490 µl of RLT buffer (with β-mer) and 350 µl of 100% EtOH were added. It was then mixed by pipetting, applied to a column and centrifuged for 30 s. It is washed twice: first with 350 µl of RW1 and then with 500 µl of RPE. As described above, any remaining wash solutions were removed by further centrifugation for 2 min and eluted in two steps to yield a final volume of 50 µl.

In order to concentrate and further purify the obtained RNA, we used Microcon 30 columns (Millipore). The samples were adjusted to a final volume of 500 µl with water, applied to the columns and mixed by pipetting. Centrifugation was performed for ~8 min to reduce the volume to ~35 µl. Finally, the column was placed in a new tube and centrifuged for 2 min.

Although RNA is stable at -20 ° C, it is advisable to store it at -80 ° C.

M-5.2.1.5. Co-immunoprecipitation (CoIP)

The *S. meliloti* wild-type strain and its derivative with a chromosomally-encoded FLAG epitope-tagged (Hfq or YbeY) protein were grown in 200 mL of the selected medium. Cells were collected by centrifugation for 10 min at 6,000xg and 4 °C. The pellet was washed twice with 24 ml of 0.1% (w/v) sarcosyl in TE and two additional times with 50 ml cold PBS. The cells were resuspended in 8 ml of lysis buffer (FLAG® Immunoprecipitation Kit, Sigma) supplemented with 0.2 mM phenylmethylsulfonyl fluoride (PMSF), and incubated at 4 °C with shaking. To increase lysis efficiency, bacterial suspensions were sonicated three times for 15-20 s with the *Branson Sonifier* sonicator. The cell lysates were centrifuged for 10 min at 12,000xg and 4 °C. The supernatant of each sample (**soluble cellular fraction**) was applied to 40 µL α-FLAG M2 affinity gel (FLAG® Immunoprecipitation Kit, Sigma) , previously washed according to the manufacturer's instructions, and incubated at 4 °C

overnight with shaking. Next, the samples were centrifuged for 1 min at 8200xg and 4 °C. The supernatants (**fraction not retained by the anti-FLAG antibody**) were removed and kept for analysis by Western blot (described below). The precipitates (**fraction bound to the anti-FLAG antibody**) were collected and transferred to *Sigma-prep*™ spin columns (SC1000, Sigma), previously cooled on ice. The columns were centrifuged for 1 min at 8200xg and 4 °C. Subsequently, the resin retained in the columns was washed three times with 500 µl of 1x wash buffer (FLAG® Immunoprecipitation Kit, Sigma) and centrifuged for 1 min at 8200xg and 4 °C. For the elution of the RNA-protein complexes, the resin was resuspended in 3 µl of 30 µg/µl FLAG peptide solution (FLAG® Immunoprecipitation Kit, Sigma) and incubated 30 min at 4 °C with gentle agitation. Finally, the columns were centrifuged for 30 s at 8,200xg and 4 °C. The eluates, containing either the Hfq^{FLAG}-RNA or $YbeY^{FLAG}$-RNA complexes, were treated with 5 U DNase I and mixed with phenol-chloroform. The CoIP-RNA was resuspended in 10 µl of RNase-free water and quantified with a NanoDrop® ND-100 spectrophotometer. Different steps of the procedure were verified by Western blotting of all protein fractions (cell lysate, soluble protein fraction, non-retained and retained proteins) using anti-FLAG antibodies.

M-5.2.2. RNA electrophoresis in denaturing agarose gels

In order to assess quality and quantity, the total RNA was separated in 1.3% denaturing agarose gels (agarose *SeaKem® LE*, Cambrex/Iberlabo) in 1x MOPS [4x: 80 mM MOPS, 20 mM sodium acetate and 4 mM EDTA; adjusted to pH 7 with NaOH] and 1.875% (v/v) formaldehyde. It was important to melt the agarose without the MOPS, as it is sensitive to light and heat. Once the agarose was melted, MOPS and formaldehyde were added in a gas extraction hood. A 6x solution (1 µl per 5 µl of sample) containing 0.25% (w/v) bromophenol blue and 45% (v/v) glycerol was added

to the samples. The electrophoresis was run in 1x MOPS buffer and with a usual voltage of 100 V. To visualize nucleic acids, an intercalating agent (ethidium bromide or Bio-safe gel Red®) was added to the buffer in a final concentration of 1% (v/v). Gels were documented as described in M-5.1.8.1.

M-5.2.3. DNA-RNA hybridization (Northern blot)

For detection of specific transcripts in an RNA sample (total or from a co-immunoprecipitated fraction) from *S. meliloti*, we used the Northern blot hybridization with a specific radiolabeled probe.

M-5.2.3.1. Electrophoresis and RNA transfer

RNA was separated in urea-polyacrylamide denaturing gels (6% acrylamide, 7 M urea, 1x TBE). The measures of the gel were 16x16 cm. For the preparation of the samples, 10-20 µg of RNA was resuspended in 2x loading buffer (0.3% bromophenol blue, 0.3% xylene cyanol FF, 10 mM EDTA pH 7.5, 97.5% deionized formamide). The samples were then loaded into the wells of the denaturing gel. As marker, γ-^{32}P-ATP-labeled pGEM was used and mixed with the loading buffer in a 1:1 ratio.

Pre-electrophoresis at 450 V in 1x TBE was performed for 30-60 min until the gel reached a temperature of approximately 50 °C. Prior to loading, nucleic acids were denatured by incubation at 95 °C for 5 min, followed by a rapid transfer to ice. The electrophoresis was performed for 1 h at 450 V until the bromophenol blue reached the gel border. The gel was removed from the electrophoresis chamber; the marker lane was cut to dry it and then documented with a PhosphorImager.

For the RNA transfer, six 3MM Whatman papers and one membrane (Zeta-Probe, Hybond N^+ Amersham Biosciences) were cut according to the gel size and wetted with 1x TBE. One of the 3MM papers was glued to the gel avoiding the formation of

bubbles and two further 3MM were placed on it. With the help of the three papers, the gel were removed from the electrophoresis glass, and then followed by placement of the membrane and three further 3 MM papers on the gel. The stack was then placed on a LKB-Amersham transfer unit with an applied current of 0.5-3 mA/cm^2 membrane for 30-60 min (e.g. 150 mA for 60 min for a standard transfer towards the positive pole). After the transfer was completed, the RNA was fixed to the membrane by incubating the filter for 1 h at 80 °C in a vacuum.

M-5.2.3.2. Hybridization

Oligonucleotides labeled at their 5'-end with γ-^{32}P-dATP were used as probe. The labeling was performed with the polynucleotide kinase (NEB) and by mixing the following components in a final volume of 10 µl: oligonucleotide (50 pmol), γ-^{32}P-ATP (20 µCi), 10x polynucleotide kinase buffer and enzyme (5 U). The reaction mixtures were incubated for 1h at 37 °C, adjusted to a volume of 25 µl with Milli-Q water, and purified with G-25 columns (Bio-Rad).

Pre-hybridization was performed at 42 °C for 30 min with 20 ml of pre-hybridization solution (0.5 M sodium phosphate buffer, pH 7.2, 7% SDS, 10 mM EDTA) for a 16x16 cm membrane. The probe was denatured by heating at 95 ° C for 5 min, and 10^6 cpm/ml were added to the hybridization solution (typically the entire volume of a common label of 50 pmol oligonucleotide) to hybridize at 42 °C overnight. The probe was removed and the membrane was washed successively with 2x SSC/0.1% (w/v) SDS for 1 min, 2x SSC/0.1% (w/v) SDS for 15 min, 1x SSC/0.1% (w/v) SDS for 15 min and 0.1x SSC/0.1% (w/v) SDS for 15 min at 42 °C. The membrane was covered with plastic and exposed over to the PhosforImager screen overnight. The images were acquired with a *Personal Molecular Imager® FX* scanner equipped with the *Quantity One®* program.

M-5.2.4. Hybridization of microarrays

M-5.2.4.1. Sm14KOLI microarrays

Sm14kOLI DNA microarray slides were provided by the University of Bielefeld (Germany). These contained 14,976 duplicated hybridization probes fixed on epoxy resin, and distributed in 48 arrays of 26 rows and 24 columns. These hybridization probes represented 6,208 oligonucleotides of 70 nt complementary to the different genes encoded in the genome of *S. meliloti* Rm1021. They also contained 8,080 oligonucleotides from 50 to 70 nt of length complementary to 2,881 intergenic regions, some of which were represented by more than one oligonucleotide complementary to both strands of the DNA (Becker *et al*., 2009). The microarray slides also contained a series of controls that correspond to 17 oligonucleotides of 70 nt that determined the specificity of hybridization (the identity of their sequences varied in a range of 70-90%). The negative controls were 70-mer oligonucleotides of sequences not present in the *S. meliloti* Rm1021genome. Besides, they also included empty spots (negative control) and spots containing buffer that promoted attachment of oligonucleotides (positive control). Finally, there were also four 70-mer oligonucleotides complementary to genes from different organisms (*gusA*, *lacZ*, *nptII*, *aacC1*).

M-5.2.4.2. cDNA synthesis and labeling

The used protocol was an adaptation of that described by Becker and Rüberg in the manual of the *S. meliloti* Sm6KOligo microarray.

10-20 µg of RNA was annealed with 10 ng of random hexamers in a final volume of 16 µl. The mixture was incubated for 10 min at 70 °C, placed for 5 min on ice and supplemented with the following components:

5x first strand buffer	6 µl
0.1 M DTT	3 µl
50x dNTPs (with aa-dUTP)	0.6 µl
RNase inhibitor (40 U/µl)	0.5 µl
SuperScript II (200 U/µl)	1.5 µl

50x dNTPs containing aminoalyl-dUTP was prepared using a 100 mM aa-dUTP stock solution. The latter was prepared by mixing 1 mg aa-dUTP, 17 µl H_2O and 0.68 µl 1 M NaOH. The pH was adjusted to 7 and the dNTPs/aa-dUTP mixture was stored at -20 °C. The concentrations were as follows: 25 mM dATP, 25 mM dCTP, 25 mM dGTP, 5 mM dTTP and 20 mM aa-dUTP (aa-dUTP/dTTP molar ratio of 4:1).

cDNA synthesis was carried out by the reverse transcriptase SuperScript II for 2 h at 42 °C, while adding additional enzyme after 1 h of reaction time. After this, the samples were stored at -20 °C.

The RNA was removed by the addition of 15 µl 0.2 M NaOH (mixing by pipetting) followed by incubation at 70 °C for 10 min. Subsequently, the solution was neutralized by adding 15 µl 0.2 M HCl. The cDNA was rapidly purified using CyScribe GFX Purification kit (GE Healthcare). 450 µl of capture buffer was added to the tube and transferred to the CyScribe GFX column. The column was washed three times with 600 µl 80% EtOH and the cDNA was eluted with 60 µl 0.1 M $NaHCO_3$ (pH 9.0).

The applied microarray technique was based on an indirect labeling of the cDNA, which is achieved by coupling a fluorophore to the aminoalyl groups present in the generated cDNA. The fluorophores Cy3 and Cy5 were prepared according to the manufacturer`s instructions. Briefly, fluorophores were mixed in 10 µl water-free DMSO, divided into 1 µl aliquots in opaque tubes and dried for 1 h in a *speedvac*. Both, Cy3/Cy5 were stored in bags with desiccant pearls and at -20 ºC. The

fluorophores were resuspended in 60 µl eluate containing the cDNA, and incubated 1-2 h in darkness at RT. Blocking of uncoupled reactive groups was performed by treatment with 4.5 µl 4 M hydroxylamine for 15 min in the dark at RT (mixing by pipetting). Unincorporated fluorophores were cleansed with a CyScribe GFX Purification kit (GE Healthcare) as follows: 600 µl capture buffer was added to the Cy5-labeled sample, mixed by pipetting, transferred to the Cy3-labeled sample and applied to a CyScribe GFX column. The mixture was washed three times with 600 µl washing buffer (kit) and eluted with 60 µl elution buffer (kit), each added to the center of the filter. The samples were incubated in a vertical position in the dark for 5 min.

The probes were concentrated in a *speedvac* to an approximate volume of 5 µl, followed by addition of 45 µl DIG Easy Hyb solution and 1 µl salmon sperm DNA (5 µg/µl).

M-5.2.4.3. Pre-treatment of microarrays

Prior to hybridization, microarray slides needed to be pre-equilibrated. Solutions were prepared freshly for every hybridization. Equilibrations were performed in darkness, at RT, with continuous agitation and with the following sequence of incubations: in solution 1 (0.001% Triton X-100) for 5 min; twice in solution 2 (0.0032% HCl) for 2 min; in solution 3 (100 mM KCl) for 10 min; and in milliQ water for 1 min. The microarray slides were then kept in blocking solution (1x QMT Blocking solution (Nexterion) supplemented with 0.0075% HCl) pre-warmed to 50 °C.

M-5.2.4.4. Hybridization and washing

The labeled cDNA was denatured at 60 °C for 5-10 min, followed by incubation at 37 °C for 30 min. The samples were added to the center of the slide (pre-heated to 50 ºC)

avoiding the formation of bubbles. The samples were dispersed homogenously before covering the microarray slides with glass plates. Samples were placed inside the hybridization chamber and incubated in a 42 °C water bath for at least 16 h.

After the hybridization was finished, a quick first wash was performed and the allowing removal of the coverslip in 2x SSC/0.2% (w/v) SDS at 42 °C. The slides were then incubated with shaking in 2x SSC/0.2% (w/v) SDS solution at 42 °C for 5 min. Successive washes of 5 min were made in 0.2x SSC/0.1% (w/v) SDS solution at 42 °C (vertical shaking) and two additional times in 0.2x SSC at RT. Finally, an additional wash was performed with 0.1x SSC (18 ° C) for 1 min. Slides were immediately dried by centrifugation for 5 min at 1200 rpm using an A-2-DWP rotor.

M-5.2.4.5. Scanning and data processing

Image acquisition with individual sensitivities for each fluorescence channel was performed with a Genepix® 4000B microarray scanner equipped with the GenePix Pro6.1 software.

The online platforms ArrayLIMS (https://arraylims.cebitec.uni-bielefeld.de/cgi-bin/login.cgi?cookietest=1) (Küster *et al.* 2007) and EMMA 2.8.2 (https://emma.cebitec.uni-bielefeld.de/cgi-bin/login.cgi?cookietest=1) (Dondrup *et al.*, 2009), both developed at the biotechnology center (CeBiTec) of the University of Bielefeld, were used for analysis of data derived from the used microarray . These programs allowed the normalization of the data and statistical treatment.

The two most significant parameters to be considered during the analysis of the data were the **A value**, which determines the average intensity corresponding to a hybridization point calculated as a function of the different replicates, and the **M value**, which reflects the balance between the two fluorescent samples. The calculation of A was performed according to the following formula: $A=0.5*\log_2$

(R_i*G_i), where R_i corresponds to the signal intensity minus the non-specific background intensity for the problem sample, and G_i is also the subtraction of intensities, but for the control sample. The M value represents the $\log_2$ ratio of R_i and G_i.

M-5.2.5. RT-PCR (reverse-transcription polymerase-chain-reaction)

The presence of specific transcripts in an RNA sample was assessed by RT-PCR. The general process of annealing and RT reaction was similar to that discussed in section 5.2.4.2., "cDNA synthesis and labeling", with the only exception thatcDNA synthesis was carried out with dUTP instead of aa-dUTP. For the hydrolysis of RNA, 10 U of RNase H (Roche) was used.

The resulting cDNA was stored at 4 °C and used for amplification by standard PCR.

M-5.3. PROTEIN MANIPULATIONS

M-5.3.1. Protein electrophoresis in SDS gels

In order to separate extracted proteins according to their molecular weight, we used vertical electrophoresis in denaturing acrylamide/bis-acrylamide gels containing SDS (SDS-PAGE) and according to the methodology developed by Laemmli (1970).

The gels were used either for direct visualization of the proteins by staining or for detection of individual proteins (Western blot analysis).

M-5.3.1.1. Preparation of SDS gels

The volumes given in Tables M-5 and M-6 correspond to two denaturing gels of 0.75 mm thickness in the REAL® Sub System Mini 10x10 cm Dual System (REAL).

Table M-5. Composition of separating gels for SDS-PAGE.

SEPARATING GEL (lower)	10%	12%	15%
Acrylamide/bis-acrylamide (40%)	3	3.6	4.5
MilliQ	5.8	5.2	4.14
1.5 M Tris-HCl pH 8.8	3	3	3
10% (w/v) SDS	0.12	0.12	0.12
10% (w/v) APS	0.12	0.12	0.12
TEMED	0.012	0.012	0.012

Numbers refer to volumes (in ml).

First, the lower gel was prepared by mixing the components as indicated in Table M-5. Three drops of isopropanol were then rapidly added (it allows a fully horizontal polymerization on its surface).

Table M-6. Composition of stacking gels for SDS-PAGE.

STACKING GEL (upper)	5%
Acrylamide/bis-acrylamide (40%)	0.375
MilliQ	2.225
1 M Tris-HCl pH 6.8	0.38
10% (w/v) SDS	0.03
10% (w/v) APS	0.03
TEMED	0.005

Numbers refer to volumes (in ml).

After polymerization for 30 min, the upper gel was prepared as indicated in Table M-6 and the well comb was placed prior to solidification of the gel matrix.

M-5.3.1.2. Electrophoresis

We used the electrophoresis buffer described by Laemmli (1970) as buffer for denaturing SDS-PAGE in the REAL® Sub-Mini 10x10 cm Dual System (REAL) electrode chamber:

Running buffer (10x)

0.124 M Tris-HCl (30.25 g)
1.252 M glycine (144 g)
5 % (w/v) SDS (20 g)
H_2O c.s.p. 1 l

The pH was adjusted to 8.8 with concentrated HCl. Samples were prepared by adding 1/3 volumes of loading buffer with the following composition:

Sample buffer (3x)

0.01 % (w/v) bromophenol blue
12 % (w/v) SDS
300 mM Tris-HCl pH 6.8
60 % (v/v) glycerol

Protein samples were mixed with the Sample buffer, supplemented with 10% DTT or 600 mM β- mercaptoEtOH, and denatured by boiling for 3 min. After that, the samples were cooled on ice for 5 min and loaded onto the gel. The electrophoresis was run at 150-200 V until the bromophenol blue reached the lower gel border.

For the rough estimation of the molecular weight of the proteins resolved in the gels, two molecular weight standards were used: a marker for low-molecular weight proteins composed of phosphorylase b (97.4 kDa), bovine serum albumin (66.2 kDa),

ovalbumin (45 kDa), carbonic anhydrase (31 kDa), trypsinogen (21.5 kDa), and lysozyme (14.1 kDa) (Bio-Rad®), and a marker ranging from 10 to 250 kDa (*Kaleidoscope Molecular Marker,* Bio-Rad®).

M-5.3.1.3. Detection of proteins in polyacrylamide gels

Protein detection was performed by the following methods for fixation and staining:

Staining with Coomassie blue

Subsequent to electrophoresis, the gel was immersed in a EtOH:acetic acid:water solution (46:45:9) together with 0.25% (w/v) *Coomassie R-250 Brilliant blue* (Bio-Rad®) for 15-30 min at RT with gentle agitation. Destaining was performed with an EtOH:acetic acid:water solution (5:7:78). To further facilitate removal of the dye a sponge piece was repeatedly added to the solution until distinct protein bands became visible.

Staining with silver nitrate

After electrophoresis, the gel was subjected to fixation and staining following a modified procedure described in Blum *et al*. (1987). The gel was immersed for 30 min with gentle agitation in 100 ml of a solution composed of EtOH, acetic acid and milliQ water (4:1:5). Subsequently, the gel was incubated for 30 min in 100 ml of a sensitizing solution containing 0.2% (w/v) sodium thiosulfate, 30% (v/v) EtOH and 6.8% (w/v) sodium acetate in milliQ water. After this pre-treatment, the gel was washed three times for 5 min with milliQ water and, incubated for 20 min in 100 ml of 0.25% (w/v) silver nitrate solution with gentle agitation and in the dark. The silver impregnation was removed by washing twice for 1 min with milliQ water. The gel was finally developed by adding 100 ml of a solution containing 2.5% (w/v) sodium carbonate and 0.004% (v/v) formaldehyde. When the protein bands appeared, the gel was incubated in a solution containing 1.46% Na_2EDTA (w/v) for at least 10 min to

completely stop the development. Finally, the gel was washed three times with milliQ water for 5 min.

M-5.3.1.4. Conservation of gels

To preserve the gels for short periods of time, they were immersed in milliQ water and maintained at 4 °C. In the case of preservation for long periods of time (more than 2-3 months), 0.2% (w/v) sodium azide was added to the water.

M-5.3.2. Immunological detection of proteins (Western blot)

For the detection of an individual protein in a cell extract, Western blot analysis was used. This method is based on an incubation of immobilized proteins with a primary antibody that specifically binds the protein of interest present in the extract. In this work, a monoclonal antibody against the FLAG epitope (*Mouse anti-FLAG® antibody*, Sigma) was used, which was recognized by a secondary antibody conjugated with horse radish peroxidase. It is finally developed by the addition of a substrate, which is converted by the enzyme resulting in a detectable light signal.

M-5.3.2.1. Transfer of proteins to PVDF membranes

After electrophoresis, the proteins were transferred to 0.45 μm pore *BioTRace PVDF* (Sigma) membranes using the Mini Trans-Blot® Electrophoretic Transfer Cell (Amersham-GE Healthcare™) system. Prior to immobilization of proteins on PVDF, the membrane was activated with 100% methanol for 10 s. Both, the gel and the membrane were equilibrated in transfer buffer (25 mM Tris pH 8.3, 192 mM glycine, 20% methanol) for 5 min with gentle agitation. Then, on a transfer cassette, the

following order of assembly was arranged: sponge and 3MM Whatman paper soaked in transfer buffer, the pre-equilibrated polyacrylamide gel, the pre-equilibrated PVDF membrane and another 3MM Whatman paper and sponge soaked in transfer buffer. The stack was placed on the electrode module in the correct orientation, which was then placed in an electrophoresis chamber supplemented with transfer buffer. The transfer was carried out at a current of 0.8 mA/cm^2 membrane for 50 min. To check for complete protein transfer, the PVDF membrane was stained with a solution containing 0.5% (w/v) Ponceau red (*Ponceau* S, SIGMA®)and 2% (v/v) acetic acid for 5 min at RT. Bands became visible after washing with distilled water until the adequate contrast was obtained. The dye was completely removed by successive washes with distilled water.

M-5.3.2.2. Immunological detection

The PVDF membrane with the transferred proteins was subjected to blocking of non-specific antibody binding sites by incubation in a solution containing 1.5% (w/v) skimmed milk powder in blocking buffer (20 mM Tris-HCl pH 7.8, 0.12 M NaCl) for 1h at RT with gentle agitation. The membrane was then three times washed with milk-free blocking buffer for 10 min. After this step, the membrane was incubated in 10 ml of 1000x-diluted primary anti-FLAG antibody in blocking buffer with skim milk for 1h at RT with gentle agitation with. After incubation, the excess of the primary antibody was removed by several 10 min washes with blocking buffer. The membrane was then incubated in 50 ml of 50,000x-diluted horse radish peroxidase-conjugated anti-mouse IgG antibody (Sigma) in blocking buffer containing skim milk for 1h at RT. The membrane was washed several times with blocking buffer to remove excess of the secondary antibody. 10 ml luminol buffer (50 mM Tris-HCl pH 8.6, 150 mM NaCl) was prepared and then a solution containing the following components was added to it in the dark: 0.04% (w/v) 3-aminophthalhydrazine (luminol, Sigma), 100 µl

4-Iodophenol (from a stock containing 10 mg in 1 ml DMSO, preserved at 4 °C) and 3.2 µl H_2O_2 (30%). The membrane was incubated in this solution for 2 min, quickly covered with plastic and placed on a photo cassette. An autoradiographic film (Kodak X-Omat®) was placed thereon and exposed 5-15 min. After this, the film was developed using the developer and fixer of TETENAL® at the dilutions and incubation times recommended by the manufacturer.

M-5.3.3. Quantitative proteomics

The experimental strategy followed in this work was based on the technique known as SILAC (*Stable Isotopic Labeling of Amino acids in Cell Culture*) (Ong, *et al.*, 2002), adaptated for the study of microbial proteomes (Sobrero, *et al.*, 2012; Lagares *et al.*, 2017). The different steps of the procedure are described below.

M-5.3.3.1. ^{15}N isotope-labeling, bacterial lysis and cellular sub-fractionation of proteins

Bacterial strains were cultured to exponential growth phase (OD_{600} 0.5) in MM containing either $^{15}C_5H_9NO_4$ (isotope-labelled L-glutamic acid) or $^{14}C_5H_9NO_4$ (unlabeled L-glutamic acid), as the sole nitrogen source. Three independent cultures of each strain were processed as biological replicates. During cell harvesting, both strains were mixed and cell pellets were collected for proteomic comparison. Aliquots of each culture grown to OD_{600} 0.5 were mixed with the corresponding control samples for proteomic comparison. The suspensions were incubated on ice for 5 min and centrifuged for 5 min at 9,000 rpm and 4 °C. The bacterial pellets were stored at -80 °C.

Protein extraction and preparation were performed through cell sub-fractionation in a **periplasmic**, **cytosolic** and **membrane** fraction, with the aim of reducing the complexity of the samples for analysis by mass spectrometry. The pellets were resuspended in 3 ml of lysis buffer (10 mM $MgCl_2$, 10 µg DNase I, 5 µg RNase A, 20 mM Tris-HCl pH 8.0) and centrifuged for 5 min at 9,000 rpm and 4 °C. At this point, the **periplasmic fraction** was considered in the supernatants (ready for the acetonic precipitation, described below in section 5.3.3.2), and the cytosolic and membrane fractions were in the pellet. For isolation of the cytosolic and membrane fractions, the pellets were resuspended in 3 ml lysis buffer, incubated on ice for 15 min and subsequently subjected to French press at 20,000 psi three times. To discard non-lysed cells, the cell extract was centrifuged for 5 min at 9,000 rpm and 4 °C. The extracts containing the cell debris were ultracentrifuged for 90 min at 160,000xg and 4 °C. The supernatants contained the **cytosolic fraction** (soluble) and the pellets contained the **membrane fraction** (insoluble). The pellet was finally carefully washed and resuspended in 150 µl of 10 mM Tris-HCl pH 8.0 containing 1% (w/v) SDS.

Total protein concentration was determined with a colorimetric Bradford Assay (*Bradford Protein Assay Ready-to-use*, BioRad, EEUU) (Bradford, 1976), using a bovine serum albumin (BSA) solution as standard. The quality of the protein preparations was analyzed by SDS-PAGE.

M-5.3.3.2. Acetonic precipitation of protein sub-fractions

To reduce the levels of contaminating agents (salts, detergents, low molecular weight cellular components) that could affect downstream applications, the proteins present in each fraction were precipitated with acetone. 100 µg of protein from each fraction were precipitated by the progressive addition of three aliquots of two volumes of cold acetone, while homogenizing the suspension after each addition. The sequential addition of acetone caused the solution to become turbid without the formation of

large protein aggregates, which improves the performance of subsequent digestion (Duan, *et al*., 2009). The mixtures were incubated overnight at -20 °C and centrifuged for 20 min at 12,000xg and 4 °C. After discarding the supernatants, the pellets were air-dried for 3-5 min and stored at -20 °C.

M-5.3.3.3. Tryptic digestion of proteins and peptide treatment

The tryptic digestion of the extracted proteins was performed to obtain peptides for chromatographic separation and identification by mass spectrometry. Each pellet was resuspended in 50 µl buffer (50 mM Tris pH 8.5), and1:60 w / w trypsin was added to the resulting solution (Qu *et al*., 2010).The mixtures were vigorously shaken for 30 s, briefly centrifuged to precipitates the content at the bottom of the tube, and incubated at 37 °C for 2 h with 120 rpm shaking. A second addition of 1.7 µg trypsin was then performed and the mixture was incubated for 5 h with shaking. Reduction of the peptides was achieved by adding 0.7 µmol DTT to the mixtures and incubation at 95 °C for 10 min. Protection of the free thiol groups (alkylation) was carried out by the addition of 2.8 µmol of iodoacetamide and incubation at 37 °C for 1 h. To titrate the excess of iodoacetamide, 2.8 µmoles of DTT was added and the solutions were incubated at 37 °C for 1 h. 4 µg of trypsin was finally added and the digestion was completed by incubating the solutions at 37 ° C overnight.

M-5.3.3.4. Separation of peptides and protein identification by mass spectrometry

The mass spectrometric analysis was performed by the mass spectrometry service of the department of chemistry of the University of Marburg (Germany) using the *Orbitrap Velos Pro* (Thermo-Scientific, USA). An *Ultimate nanoRSLC-HPLC system* (Dionex) equipped with a nano-C18-RP column was connected to the spectrometer by

a *Proxeon nano-spray* system. 1 µl of trypsin-treated protein solution was injected into the chromatographic column. The capture was carried out with a flow of 6 µl/min of an aqueous sample solution containing 0.05% (v/v) formic acid. Separation of tryptic peptides was performed by a gradient of solvent B containing 0.05% (v/v) formic acid in 80% (v/v) acetonitrile. The gradient was generated at a flow rate of 300 nl/ min. Initially, the 4% (v/v) solvent B was used for 5 min, then the concentration was linearly increased to 45% (v/v) over a time period of 30 min and to 95% (v/v) within the final 5 min. The eluent of the column was sprayed directly onto the preheated capillary of the mass spectrometer with a potential of 2300 V. An initial scan with a resolution of 6000 was performed on the *Orbitrap* mass analyzer with at least three independent MS/MS scans. At the same time, the "dynamic exclusion" function was used for 30 swith the linear ion trap and a resolution of 7500 in the detection by the *Orbitrap*.

M-5.3.3.5. Comprehensive bioinformatic analysis of spectral data by QuPE

Spectral data were loaded onto the QuPE server (Albaum, *et al*., 2009) of the University of Bielefeld (Germany). An initial pre-processing of the spectra was carried out by applying a filter according to the pre-determined parameters by the server. From the filtered spectra, the identity of the peptides was mapped using the Mascot® search engine in the *S. meliloti*-Decoy database. For the annotation of the peptides that were subject to the quantitative analysis, an FDR (*false discovery rate*) analysis was performed, and only those that exceeded the threshold value of 0.05 were filtered. The quantification of the intensities of the light and heavy isotopic variants of each peptide was performed using the RelEX-linear exclusion algorithm (MacCoss, *et al*., 2003) (using parameters pre-determined by the QuPE server). A *t* Student statistical analysis was performed, considering only the single peptides (with A values > 3.5) annotated for each protein *i*, for the estimation of the corresponding M_i value.

The A value was calculated as the product of the peak intensity of the isotopic variants of each peptide in the chromatograms constructed from the spectra register. The M_i value represents the $\log_2$ ratio between the abundance of protein *i* in the problem sample / abundance of protein *i* in control sample).

M-6. PLANTS ASSAYS

M-6.1. STERILIZATION AND GERMINATION OF SEEDS

The *M. sativa* (alfalfa) seeds were immersed in a solution of 2.5% $HgCl_2$ for 9 min with agitation. Under sterile conditions, they were washed five to six times and kept for 1-2 h in distilled water. Finally, they were washed two to three times with sterile distilled water and placed in Petri dishes with agar-water medium (1.5% agar in deionized water). The germination was carried out in the dark at 28 ° C until the seedlings reached a length of about 10 mm (approximately after 24 h).

M-6.2. NUTRIENT SOLUTIONS FOR PLANTS

We employed a mineral solution derived from the one described by Rigaud and Puppo (1975) and composed as follows:

Macroelements (in 1 l distilled H_2O):

KH_2PO_4	0.2 g
$MgSO_4 \cdot 7H_2O$	0.2 g
K_2SO_4	0.2 g
$CaSO_4$	0.12 g
Ferric EDTA	0.05 g

Microelements (in 1 l distilled H_2O):

$Na_2MoO_4 \cdot 2H_2O$	4 mg
H_3BO_3	18.6 mg
$MnSO_4 \cdot 4H_2O$	22.6 mg
$ZnSO_4 \cdot 7H_2O$	29 mg
$CuSO_4 \cdot 5H_2O$	24 mg

The pH of the solution was adjusted to 7 with KOH. The solution was autoclaved at 120 °C for 20 min.

M-6.3. CULTIVATION OF PLANTS

M-6.3.1. Axenic cultivations in tubes

The axenic cultivation of alfalfa plants was performed according to the technique described by Olivares *et al*. (1980). 20x200 mm tubes containing a filter paper and 10 ml Rigaud and Puppo mineral solution were sterilized by autoclaving at 120 °C for 20 min. When newly germinated alfalfa seedlings reached about 10 mm in length, they were placed in the tubes (1 plant/tube) under aseptic conditions. The tubes with the plants were left in darkness for 1-2 days. To prevent roots from being directly affected by light, the tubes were covered in their lower half with opaque paper and taken to the plant growth chamber, maintained under the following conditions: 500 $\mu E\ m^{-2}\ s^{-1}$ (wavelength: 400-700 nm) light intensity, a photoperiod of 16/8 h light/darkness, 23/17 ºC of temperature day/night and 50% humidity.

After 7-10 days of plant growth, plants were inoculated with investigated *S. meliloti* strains in 1 ml suspensions of approximately 10^6 cells/ml. Uninoculated plants served as a control.

M-6.3.2. Axenic cultivations in agar plates

For the culture of plants on solid medium, sterile 6x6 cm plates with Rigaud and Puppo mineral solution supplemented with 1.5% nitrogen-free bacteriological agar were used. Seeds were sterilized and germinated according to the protocol described in section 6.1 and 1-2 days old seedlings were placed on the solid medium (not more than six plants per plate). The plates were placed in a plant growth chamber with the conditions described in the previous section, in an upright position slightly inclined and partially covered with an opaque paper.

After 7 days of plant growth, plants were inoculated with a bacterial suspension of 10^6 cells/ml grown in TY medium. This suspension was washed two times with distilled water, and 10 ml were inoculated on each plate. The plates were left horizontally in sterile conditions for 1 h and after this the excess solution was removed. The plates were re-placed in the plant growth chamber and incubated for 30 days until plant harvest.

M-6.3.3. Cultivation in Leonard jars

The Leonard jars is a plant culture device described by Leonard (1943) and consisting of two containers. The upper vessel was filled with vermiculite and the lower one with Rigaud and Puppo mineral solution, connecting both with strips of filter paper and all sterilized by autoclaving at 120 °C for 20 min. In each jar, sterilized and germinated seeds were seeded and inoculated with *S. meliloti* strains in 1 ml suspension of 10^6-10^7 cells/ml. Uninoculated plants served as a control. Once inoculated, the jars were covered with a layer of sterile perlite that avoids possible environmental contamination and reflects the light, avoiding excessive heating of the seeds that could alter their development. The seal between the upper and lower components of the jar

was sealed with parafilm and the jars were wrapped in opaque paper to prevent light exposure of the roots.

The jars were moved to a plant growth chamber maintained under the conditions described in section 6.3.1.

M-6.4. NODULATION KINETIC ANALYSIS

In order to assess the symbiotic potency of *S. meliloti* strains over time, two parameters were monitored: the number of nodules per plant and the percentage of nodulated plants. For this purpose, tubes with plants grown in hydroponic culture were placed as described in section 6.3.1. Once inoculated, nodules and plants with at least one nodule (nodulated plants) were daily counted for the first 15 days and every 2-3 days until the experiment was completed (after 28 days). In addition, nitrogen-fixing nodules were monitored by determining the count and day of appearance of pink-coloured nodules on each plant root.

At 28 days post-inoculation (dpi), nodules were subjected to microscopy by preparing longitudinal sections of the nodules. Harvested nodules were immersed in buffer containing 50 mM sodium citrate tribasic dehydrate (pH 7) and cut using a Leica VT1000S vibratome. The sections were then placed on microscopy slides with one drop of the buffer and sealed with cover slips.

M-6.5. COMPETITIVE NODULATION ASSAYS

To measure the degree of nodulation competitiveness, twelve to 24 tubes were prepared with 7-10 days old plants and inoculated with 1 ml of cell suspensions composed of 1:1 mixtures of a GUS-tagged *S. meliloti* wild-type strain and a non-tagged mutant derivative. In control experiments, plants were co-inoculated with the tagged and non-tagged wild-type bacteria (García-Rodríguez and Toro, 2000). Both strains were adjusted to a final concentration of 10^6 cells/ml. Inoculated plants were

maintained for a period of 30 days in a plant growth chamber under the conditions described above. Roots were collected 30 dpi, washed in distilled water and immersed in staining solution [1 mM 5-bromo-4-chloro-3-indolyl-D-glucuronic acid (X-Gluc), 1% (w/v) SDS, 50 mM sodium phosphate buffer (pH 7.5)] at 37 ºC in the dark to visualize GUS activity in nodules (van Dillewijn *et al*., 2001). The nodules occupied with the pGUS3 strain displayed bluish staining compared to white nodules occupied by the strain without plasmid. The degree of competitiveness was calculated as the ratio of the number of nodules occupied by the mutant strain and that occupied by the wild-type strain. The statistical error was calculated according to the following formula:

$$E = 1.96\sqrt{\frac{P(1-P)}{N}} \times 100$$

E = Statistical error

P = Percentage of occupied nodules (100% set to 1)

N = Total number of recorded nodules

M-6.6. EVALUATION OF PLANT MEASURES

Fresh plants were used for measuring shoot lengths, determined 30-35 dpi.

M-6.7. ISOLATION OF BACTEROIDS

Bacteroids were isolated essentially as previously described (Finan *et al*., 1983) with modificationsof methods used by Planque *et al*. (1978) and Laane *et al.* (1978).Ten to 16 nodules per strain were harvested 28 dpi and immersed in 200 µl MMS buffer (40 mM MOPS pH 7, 20 mM KOH, 2 mM $MgSO_4$, 0.3 M sucrose). Then, the nodules were crushed with a pestle. 800 µl MMS buffer was added, the crude extracts were sedimented and the supernatants were collected. Then, the samples were centrifuged

for 10 min at 500xg and 4°C. The pellet was discarded and the supernatants were centrifuged for 10 min at 2,000xg and 4°C to sediment the bacteroids. After the supernatant was carefully removed, the final pellet was suspended in 50 µl MMS buffer.

2.5 ul of each sample was spotted on 1 % agarose pads and directly used for microscopy.

M-6.8. MICROSCOPY

To assess the endosymbiotic phenotypes of gene deletion mutants, whole wild-type- and mutant-induced nodules on alfalfa roots were prepared as described above and closely inspected with a binocular NIKON C-DSD230.

Longitudinal nodule sections and bacteroids on 1 % agarose pads were prepared as described above and visually examined by brightfield microscopy using a Nikon microscope Eclipse Ti-E equipped with differential interference contrast (DIC) CFI Apochromat TIRF oil objective (100x; numerical aperture of 1.49) with AHF HC filter set F36-525 EGFP (excitation band pass 472/30 nm, beam splitter 495 nm, and emission band pass 520/35 nm filters). Images were acquired with an Andor iXon3 885 electron-multiplying charge-coupled device (EMCCD) camera. Image acquisition and adjustment was done with Nikon NIS elements 4.0 software.

M-7. BIOINFORMATIC STRATEGIES FOR PREDICTION OF mRNA TARGETS

Due to the complexity of the interactions between sRNAs and mRNAs, the experimental revealing of sRNA regulons can be costly and time intensive. The development of algorithm strategies during the last years has significantly contributed

to the identification of new mRNA targets.

M-7.1. IntaRNA

IntaRNA (Interacting RNAs) is one of the assets of the Freiburg RNA Tools webserver (http://rna.informatik.uni-freiburg.de/IntaRNA/Input.jsp) (Smith *et al*., 2010). It predicts interacting regions between two RNA molecules by incorporating the accessibility of both interaction sites and the presence of a seed interaction; both features are commonly observed in sRNA–mRNA interactions (Richter *et al*., 2012). IntaRNA can also be applied to non-whole genome screens using smaller sets of RNA molecules as input. Thus, it is also applicable to RNA–RNA interaction predictions for eukaryotic systems (Starczynowski *et al*., 2011).

For the generation of prioritized lists of candidates, IntaRNA essentially combines two calculated parameters *in silico*: the energy of sRNA-mRNA hybridization and the accessibility for mating of the complementary sequences in each molecule. In our analysis, the program inputs were the complete sequences of the sRNAs studied in this work, and a FASTA file comprising all nucleotide sequences comprised between position -250 and +150 with respect to the annotated start codon of every ORF in the *S. meliloti* genome. We therefore assumed the sRNA-mRNA interaction in the 5'-UTR region of the target messenger, which operates predominantly in riboregulation in bacteria. As an additional parameter for generating predictions, a minimum of eight consecutive nucleotides matching in the sRNA and mRNA target was imposed on both molecules.

M-7.2. CopraRNA

CopraRNA (Comparative prediction algorithm for small RNA targets) is a recent asset of the Freiburg RNA Tools webserver (http://rna.informatik.uni-freiburg.de/ CopraRNA/Input.jsp). It incorporates and extends the functionality of the previously

existing tool IntaRNA (described above), in order to predict targets, interaction domains and consequently the regulatory networks of bacterial sRNA molecules (Wright *et al.*, 2013, 2014). The CopraRNA prediction results also provide extensive post-processing methods such as functional enrichment analysis and illustration of interacting regions. It is a comparative method that constructs a combined sRNA target prediction for a set of given organisms. Employing a statistical model, this tool computes whole genome target predictions by combining whole genome IntaRNA target screens for homologous sRNA sequences from distinct organisms. Individual evolutionary distances between the organisms and the statistical dependencies in the data are considered and corrected within the workflow of the algorithm (Wright *et al.*, 2013, 2014). The interaction energies are fitted to a general extreme value distribution and transformed into *P* values to normalize for organism-specific GC-content and dinucleotide frequency. These *P* values are combined for orthologous genes into a single *P* value per conserved interaction. Orthologous genes are determined based on the respective amino acid sequences (Uchiyama *et al.*, 2007), while genes that are present in less than 50% of the investigated genomes remain unconsidered. Two aspects require specific normalization. First, CopraRNA normalizes for the degree of overall dependency to account for the non-independent *P* values that result from the general sequence conservation between related organisms. Second, the individual dependencies have to be calculated because, in most cases, the considered organisms will not be equidistant from each other. Input data must be supplied in FASTA format. For CopraRNA, the FASTA file should represent three or more homologous sRNA sequences from distinct organisms. Homologous sRNA sequences may be retrieved from databases such as NCBI via BLAST (Altschul *et al.*, 1990) or from Rfam (Burge *et al.*, 2013). While only three input sequences are mandatory, it is suggested using at least five, if available. For each sequence, CopraRNA requires a RefSeq ID of its affiliated organism as FASTA header. If several RefSeq IDs correspond to replicons of one organism, any of these IDs may be supplied. A maximum of eight input organisms is possible. One of these species must be selected as central reference

(organism of interest) for post-processing and annotation (Wright *et al*., 2013, 2014).

RESULTS AND DISCUSSION

CHAPTER 1

TRANSCRIPTIONAL REGULATION AND TARGETING POTENTIAL OF THE AbcR2 sRNA

1.1. INTRODUCTION

Rhizobial genomes are predicted to encode an unusually large repertoire of ATP-binding cassette (ABC) transporters depending on a periplasmic solute binding protein (SBP), which guarantees bacteria to cope with the oligotrophy of soil (e.g. 200 ABC genes in *Sinorhizobium meliloti* compared to 67 in *Escherichia coli*) (Galibert *et al*., 2001; Mauchline *et al*., 2006). In *S. meliloti*, Hfq influences the expression of a large fraction of ABC transporter genes and the corresponding mRNAs are major Hfq-binding transcripts (Gao *et al*., 2010; Torres-Quesada *et al*., 2010). Therefore, many of the *S. meliloti* Hfq-dependent mRNAs from transporter genes are expected to be post-transcriptionally regulated by the concerted activity of Hfq and its cognate partner sRNAs (Gao *et al*., 2010; Torres-Quesada *et al*., 2010).

Recently, mRNAs encoding the periplasmic component of ABC transport systems have been shown to be targeted by at least two sRNAs, which were therefore named AbcR1 and AbcR2 (Torres-Quesada *et al*., 2013, 2014; Overlöper *et al*., 2014; Harfouche *et al*., 2014; Jiménez-Zurdo and Robledo, 2015). Both *trans*-acting sRNAs are homologous, strictly Hfq-dependent and expressed in phylogenetically close α-proteobacterial species (e.g. *S. meliloti*, *Agrobacterium tumefaciens* and *Brucella abortus*) (Torres-Quesada *et al*., 2013, 2014). AbcR1/2 belong to the family of sRNAs designated αr15, whose members occur in multiple copies in bacterial genomes classified in the Rhizobiaceae and Brucellaceae families of the order Rhizobiales (del Val *et al*., 2012). In *S. meliloti*, AbcR1 and AbcR2 are tandemly encoded in the same intergenic region but exhibit divergent unlinked regulation (Figure 1.1, A) (Torres-Quesada *et al*., 2013). AbcR1 is transcribed in actively dividing bacteria, either in culture, rhizosphere or within the invasion zone of mature alfalfa nodules, whereas AbcR2 expression is induced upon entry of bacteria into stationary growth phase and under a number of abiotic stresses (Figure 1.1, A). Only deletion of AbcR1 results in a discrete growth delay in rich medium, but both AbcR1 and AbcR2 are dispensable for

the establishment of an efficient symbiosis (Figure 1.1, A) (Torres-Quesada *et al.*, 2013; Harfouche *et al.*, 2014).

At the beginning of this work, one single mRNA encoding the branched-chained amino acid transporter LivK was identified as AbcR1 target. However, LivK accumulation did not seem to be influenced by AbcR2 expression (Torres-Quesada *et al.*, 2013, 2014). Their divergent unrelated expression profiles, the exclusive contribution of AbcR1 to a growth phenotype and their specific targeting potential were thus evidences for independent regulatory functions of AbcR1 and AbcR2 sRNAs in *S. meliloti* (Torres-Quesada *et al.*, 2013). Their *A. tumefaciens* homologs also have a different targeting potential but are co-regulated. AbcR1, but not AbcR2, silences a suite of ABC transporter mRNAs, including the one encoding the SBP of the plant-derived quorum sensing signal γ-amino butyric acid (GABA), thus suggesting a function of this sRNA in phytopathogenesis (Wilms *et al.*, 2011). In *B. abortus*, AbcR1 and AbcR2 similarly regulate a set of uncharacterized amino acid and polyamine transporters, and deletion of both sRNA *loci* is required to attenuate virulence, which suggests redundant functions (Caswell *et al.*, 2012). In all cases, AbcR1/2-mediated regulation likely proceeds by a canonical mechanism involving base-pairing with nucleotides within and in the vicinity of the Shine-Dalgarno (SD) sequence of the target mRNA, which would interfere with translation initiation. Two conserved anti-Shine Dalgarno (aSD) motifs (M1 and M2) (Figure 1.1, B) that are predicted to remain unpaired in the α-proteobacterial AbcR1 and AbcR2 homologs have been identified as the targeting domains of these sRNAs (Torres-Quesada *et al.*, 2013; Jiménez-Zurdo and Robledo, 2015). AbcR1 and AbcR2 are thus typical examples of bacterial sRNAs that regulate large arrays of mRNAs, which anticipates that the regulatory potential of these *S. meliloti* sRNA homologs remains largely unexplored (Torres-Quesada *et al.*, 2014).

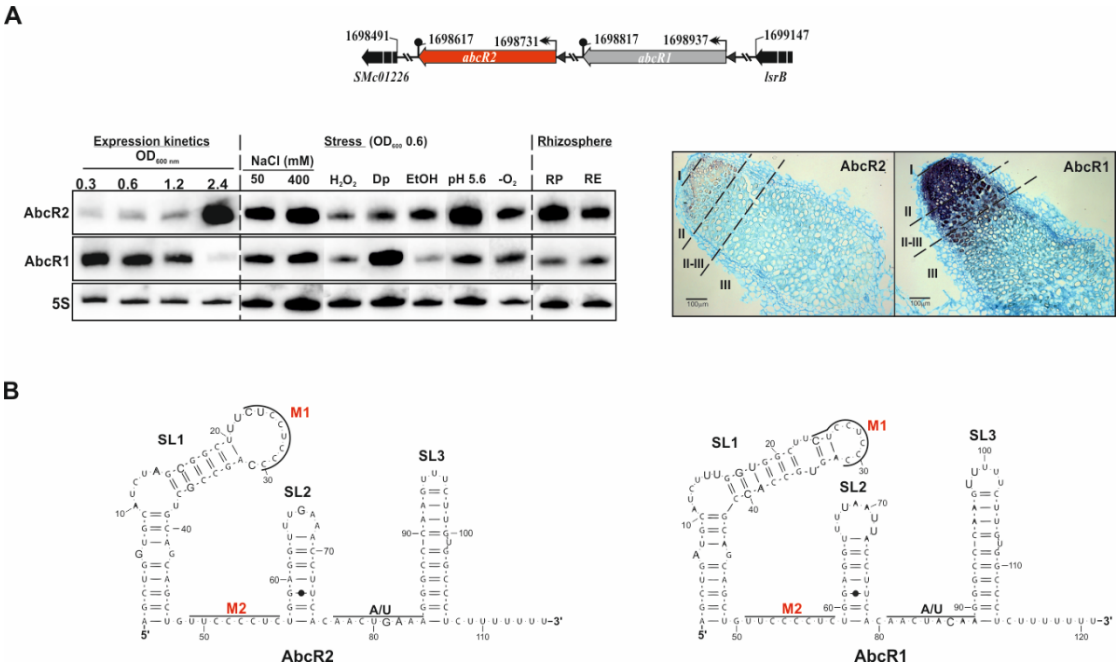

Figure 1.1. The *S. meliloti* AbcR1 and AbcR2 sRNAs. (A) AbcR1/2 expression in *S. meliloti* Rm1021. In upper panel, genomic region of the AbcR1 and AbcR2 sRNA loci in the chromosome of the reference strain Rm1021, indicating their flanking genes and relevant coordinates. In the left panel, Northern blot detection of AbcR1/2 transcripts in total RNA obtained at different OD_{600} (indicated above the panel) during Rm1021 growth in rich medium (expression kinetics), under different stresses and in rhizosphere-like conditions as indicated on top. 5S rRNA probing was used as RNA loading control. $-O_2$, microoxic conditions; RP, Rigaud and Puppo medium; RE, root exudates (plant presence). In the right panel, *in situ* hybridization of sections of *M. sativa* mature nodules occupied by Rm1021 with digoxigenin (DIG)-labeled riboprobes targeting AbcR1 and AbcR2. Zones of typical indeterminate nodules (I, II, II-III and III) are indicated. Bar represents 100 µm. **(B)** Predicted secondary structures of AbcR1/2. Numberings denote relative nucleotide positions from the 5'-end of each molecule. The two conserved anti-Shine Dalgarno (aSD) motifs (M1 and M2) are indicated in red. SL, stem-loop domain (Modified after Torres-Quesada *et al.*, 2013).

In this chapter, we present novel data about the transcriptional regulation and mRNA targets of the stress-induced AbcR2 sRNA.

1.2. EXPERIMENTAL SETUP

1.2.1. CONSTRUCTION OF *S. meliloti* MUTANTS AND DERIVATIVE STRAINS

1.2.1.1. Generation of the *S. meliloti* Rm1021Δ*abcR2* mutant and its derivative overexpression strain

In this work, the *S. meliloti* Rm1021Δ*abcR2* mutant was used, which was complemented with plasmid pSRK-R2 (constitutively expressing AbcR2) (Figure 1.2) (Torres-Quesada *et al.,* 2013). The **Rm1021Δ*abcR2*** mutant was constructed by replacing the chromosomal sRNA locus with an erythromycin resistance cassette and further verified in strain Rm1021 by Torres-Quesada *et al.,* 2013. Plasmid **pSRK-R2**, constitutively expressing *abcR2*, was generated by engineering the mid-copy pBBR1MCS-2 derivative pSRKKm (Khan *et al*., 2008) and mobilized to Rm1021 and Rm1021Δ*abcR2* mutant by conjugation (Figure 1.2, A). All strains were verified by Northern analysis (Figure 1.2, B) by Torres-Quesada *et al.,* 2013.

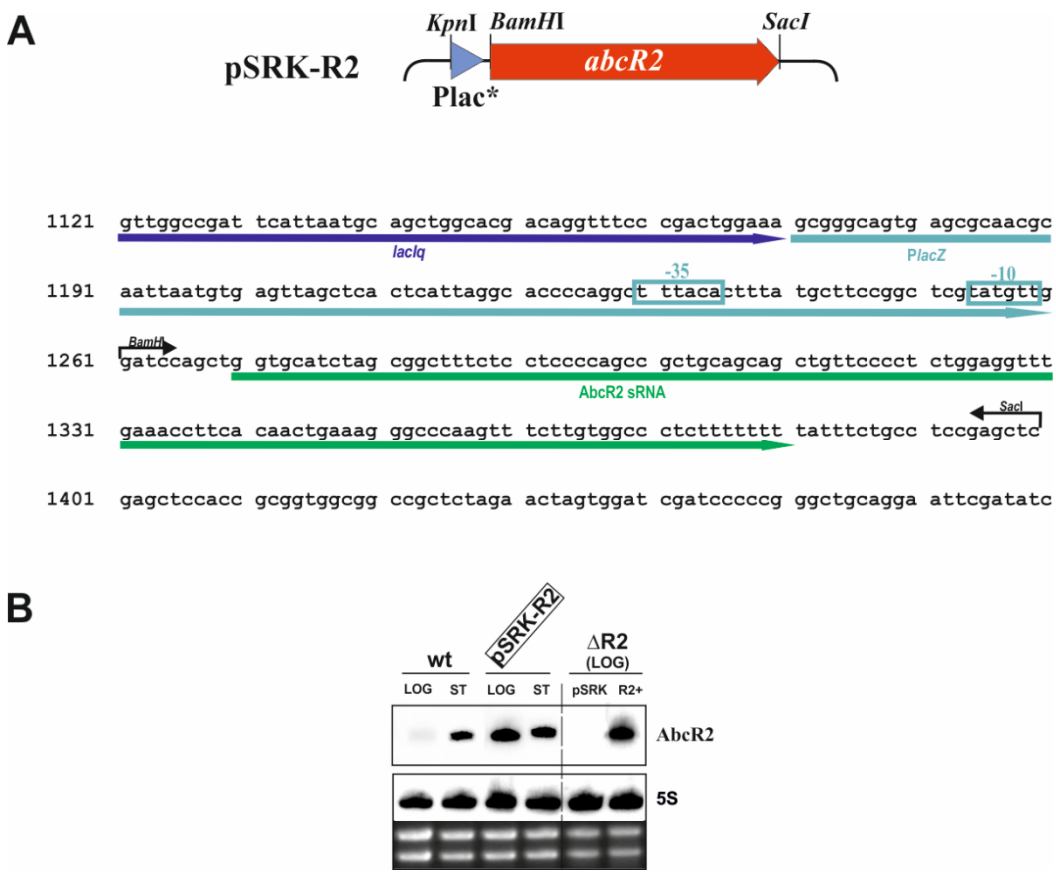

Figure 1.2. Constitutive AbcR2 (over)expression. (A) Upper panel, diagram of the genetic constructs tested to express AbcR2 from a modified *lac* promoter (Plac*) on pSRK. Relevant restriction sites for cloning the full-length AbcR2 *locus* are indicated (taken from Torres-Quesada *et al*., 2013). Lower panel, section of the nucleotide sequence of pSRK-R2. The sRNA was cloned as *BamH*I/*Sac*I fragment. In light blue, the *lacZ* promoter (P*lacZ*) is shown upstream of the AbcR2 sRNA (in green), and the -35 and -10 boxes are indicated. **(B)** Northern hybridization analysis of total RNA extracted from *S. meliloti* Rm1021 wild-type strain (wt) and the transconjugant harboring pSRK-R2 grown to exponential (LOG) and stationary phases (ST) in TY broth. The last two lanes in each panel correspond to RNA samples from the Rm1021 and Rm1021Δ*abcR2* deletion mutant transformed with pSRK, pSRK-R2 (R2+) as indicated on top. The hybridization signal corresponding to the 5S rRNA and the ethidium bromide MOPS-formaldehyde gel with the 23S and 16S RNAs are shown below the panel (modified after Torres-Quesada *et al*., 2013).

1.2.1.2. Double-plasmid Reporter Assay

To assess regulation of ***oppA (SMb21196)***, ***prbA (SMc01642)*** and ***SMa0495*** mRNAs by AbcR2 *in vivo* we used a reporter assay based on that developed by Urban and

Vogel for enterobacteria (Urban and Vogel, 2007). Our system is based on the co-expression in the same cell of two compatible plasmids transferred by conjugation to the appropriate recipient *S. meliloti* Rm1021 derivative. These plasmids included the medium-copy number plasmid pSRK-R2 expressing the full-length sRNA from a modified *lac* promoter (Plac*) (see above) and low-copy reporter plasmids (pR*oppA::egfp*, pR*prbA::egfp*, pR*SMa0495::egfp*) derived from the IncP broad host-range vector pJB3Tc19 (Blatny *et al*., 1997) carrying translational fusions of the *oppA*, *prbA* and *SMa0495* upstream regions to *egfp* under the control of a constitutive synthetic promoter (P_{syn}) (Giacomini *et al*., 1994).

The mentioned reporter plasmids were constructed by Torres-Quesada *et al*., 2014, and for this work they were individually transferred by conjugation to an *S. meliloti* Rm1021 AbcR1/2 double deletion mutant (Δ*R1/2*) (Torres-Quesada *et al*., 2013) harboring plasmids pSRK-R2 (constitutively expressing AbcR2 sRNA), or the empty vector pSRK (Torres-Quesada *et al*., 2013). Four double transconjugants for each pSRK-target fusion combination were grown in TY medium to exponential growth phase and fluorescence of 100 µl of bacterial cell cultures in 96-well microtiter plates (Greiner) was measured with a Tecan Infinite M200 reader (Tecan Trading AG). Background fluorescence was determined from strains harboring pSRK and the empty pR_EGFP plasmid and subtracted from the fluorescence produced by target fusions.

1.2.2. ANALYSIS OF THE AbcR2 PROMOTER

The AbcR1/2 promoters analysis was performed in collaboration with Dr. Coral del Val (University of Granada) and as described in https://en.wikipedia.org/wiki/%CE%91r15_RNA#Promoter_Analysis. To identify binding sites for known transcription factors we used the fasta sequences provided by **RegPredict** (Novichkov *et al*., 2010) (http://regpredict.lbl.gov/regpredict/help.html), and used those position weight matrices (PSWM) provided by **RegulonDB** (Gama-Castro *et al*., 2011) (http://

regulondb.ccg.unam.mx). We built PSWM for each transcription factor from the RegPredict sequences using the **Consensus/Patser program**, choosing the best final matrix for motif lengths between 14-30 bp with a threshold average E-value < 10^{-10} for each matrix that was established. Moreover, we searched for conserved unknown motifs using **MEME algorithm** (Bailey and Elkan, 1994) (http://meme-suite.org/) and used relaxed regular expressions (i.e. pattern matching) over all αr15 homologous promoters.

Based on established promoter models (Schlüter *et al*., 2013), profile searches for putative promoters specific for the σ factors RpoD, RpoE2, RpoH1, RpoH2, RpoH1/2, and RpoN were performed for the *abcR1/2* genes using **PoSSuMsearch** (Beckstette *et al*., 2009).

The logo of the consensus sequence of this motif was generated at http://weblogo.berkeley.edu/logo.cgi. Promoter sequence alignments were generated with **ClustalW** implemented in **BioEdit** (http://www.mbio.ncsu.edu/BioEdit/bioedit.html).

The *S. meliloti* Rm1021 mutant strains used for Northern (Material and Methods) and promoter analyses were **VO3128 (*rpoH1*::*aadA*)**, **AB3 (*rpoH2*::*aacCI*)** and double mutant **AB9 (*rpoH1*::*aadA rpoH2*::*aacCI*)**, each carrying gene-disrupting constructs at the respective native genomic locations (Oke *et al*., 2001; Bittner and Oke, 2006). All three mutant strains were provided by Dr. Melanie J Barnett (Stanford University).

1.2.3. ANALYSIS OF THE AbcR2 REGULON

1.2.3.1. Microarray-based transcriptomics

Total RNA was obtained with the RNeasy Mini Kit (Qiagen), which was applied to Rm1021 and Rm1021Δ*abcR2* bacteria cultured to exponential growth phase (OD_{600} 0.5) (four independent cultures per strain) in MM followed by an osmotic upshift

(Material and Methods). cDNA synthesis, Cy3- and Cy5-labeling, competitive hybridization of wild-type and mutant RNA to Sm14kOLI microarrays, image acquisition and data analysis were performed as described in Material and Methods. Signal intensities obtained for the wild-type strain grown exponentially in MM and subjected to the salt shock was additionally compiled in an *in silico* microarray experiment. Normalization and t-statistics were carried out using the EMMA 2.8.2 microarray data analysis software (Dondrup *et al.*, 2009). Genes with P-value ≤ 0.05 and M ≥ 1.0 or ≤ -1.0 were included in the analysis. The M value represents the $\log_2$ ratio between both channels. Functional categories of the differentially expressed genes were established according to the *S. meliloti* Rm1021 genome sequence annotation (Galibert *et al.*, 2001).

1.2.3.2. Quantitative proteomics

Strains Rm1021 and Rm1021Δ*abcR2* were cultured to exponential growth phase (OD_{600} 0.5) in MM followed by an osmotic upshift for 1.5 hours in medium containing either $^{15}C_5H_9NO_4$ (isotope-labelled L-glutamic acid) or $^{14}C_5H_9NO_4$ (unlabeled L-glutamic acid), as the sole nitrogen source. Three independent cultures of each strain were processed as biological replicates. During cell harvesting, both strains were mixed and cell pellets were collected for proteomic comparison. The complete extraction procedure and analysis of the periplasmic fraction of proteins are described in Material and Methods.

1.3. RESULTS

1.3.1. TRANSCRIPTIONAL REGULATION OF THE AbcR2 sRNA IN *S. meliloti*

In *S. meliloti*, transcriptional control of adaptive responses to changing environments can be exerted by a large set of transcription factors (8.7% of the *S. meliloti* protein-coding genes) and 15 alternative RNA polymerase holoenzymes (σ factors), of which eleven belong to the extracytoplasmic function (ECF) subfamily (Galibert *et al*., 2001). RpoH1/2 ($\sigma^{H1/2}$ homologous to *E. coli* heat shock σ^{32}) and the ECF σ factor RpoE2 (σ^{E2}, functionally analogous to the enterobacterial global stress regulator RpoS) have been shown to be major regulators of the general (stationary phase), heat shock, and hyperosmotic stress responses (Ono *et al*., 2001; Oke *et al*., 2001; Sauviac *et al*., 2007; Flechard *et al*., 2010; Barnett., *et al* 2012; Schlüter *et al*., 2013).

The practically opposite expression profiles of AbcR1 and AbcR2 in *S. meliloti* Rm1021 (Figure 1.1, A) suggested a differential regulation of both transcripts. To better understand their expression, we further investigated the transcriptional regulation of these sRNAs by closer inspection of their promoter regions (Figure 1.3). DNA sequence stretches were collected, which reached up to 100 bp upstream of the predicted transcription start site (TSS) of αr15 loci, and exhibited the closest homology to AbcR1 and AbcR2. This group of sequences included AbcR1/2 homologs encoded in diverse *Sinorhizobium*, *Rhizobium*, and *Agrobacterium* genomes (del Val *et al*., 2012). The alignment of a subset of ten representative sequences homologous to AbcR1 revealed a conserved sequence stretch extending up to 80 bp upstream of the TSS of all loci analyzed. In addition, it evidenced recognizable σ^{70}-dependent promoters showing a -35/-10 consensus motif CTTGAC-N_{17}-CTATAT (Figure 1.3, A, upper panel), which has been previously shown to be widely conserved among several other genera in the α-subgroup of proteobacteria (MacLellan *et al*., 2006), and mostly operates during the exponential growth of bacteria (Bar-Nahum and

Nudler, 2001; Hofmann *et al.*, 2011; Haakonsen *et al.*, 2015). Nevertheless, the alignment of eight representative sequences homologous to AbcR2 suggested that transcription of this sRNA depends on the alternative RNA polymerase σ factor RpoH1, showing a conserved -35/-10 consensus motif CTTGAA-N_{16}-CCTATAT (Figure 1.3, A, bottom panel).

The diverse conserved motifs identified in this *in silico* analysis suggested a differentially regulated transcription of both sRNAs (Figure 1.3, A), and therefore, support their observed differential expression profiles (Figure 1.1, A) (del Val *et al.*, 2007; Torres-Quesada *et al.*, 2013).

Similar promoter consensus motifs of the *S. meliloti* σ^{70}, σ^{H1}, σ^{H2} and $\sigma^{H1/2}$ factors have been noted (Barnett *et al.,* 2012; Schlüter *et al.*, 2013). To support the results from our promoter *in silico* analysis and to define the contribution of $\sigma^{H1/2}$ to the regulation of *abcR2* expression, we performed Northern blot analysis by probing the RNA obtained from wild-type *S. meliloti* Rm1021 and *rpoH1* (VO3128), *rpoH2* (AB3), and *rpoH1 rpoH2* (AB9) mutant strains grown in four different conditions (i.e. exponential and stationary growth phases, and salt and heat shocks) in TY broth and MM (Material and Methods) (Figure 1.3, B). The Northern hybridization confirmed the stress-induced expression of the transcript in TY and MM, and revealed the dependence of AbcR2 on the σ^{H1} factor.

Therefore, we conclude that both sRNAs are differentially regulated and that the alternative RNA polymerase σ factor RpoH1 is required for *abcR2* transcription.

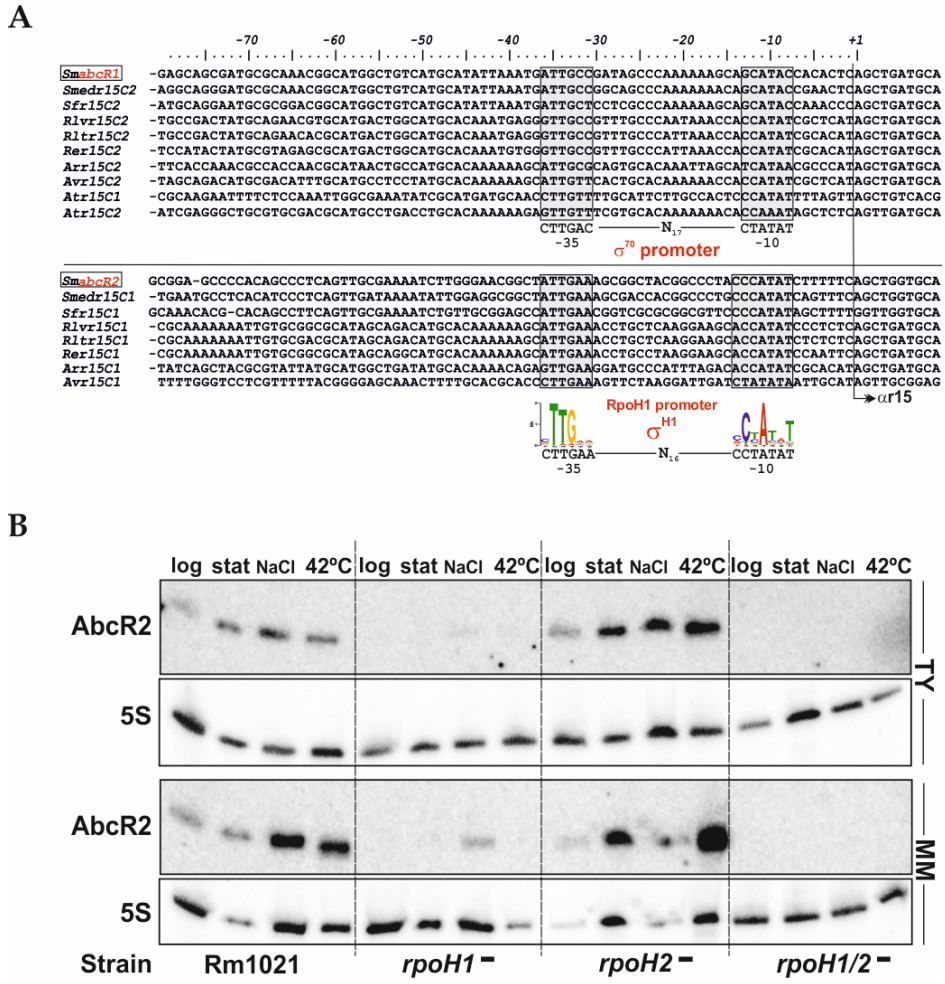

Figure 1.3. Transcriptional regulation of AbcR2 in free-living bacteria. **(A)** In upper panel, sequence alignment of the promoter regions of *Sinorhizobium meliloti* (*Sm*) AbcR1 homologs in type strains of *S. medicae* (*Smedr*), *S. fredii* (*Sfr*), *Rhizobium leguminosarum* bv. *viciae* (*Rlvr*), *R. leguminosarum* bv. trifolii (*Rltr*), *R. etli* (*Rer*), *Agrobacterium radiobacter* (*Arr*), *A. vitis* (*Avr*) and *A. tumefaciens* (*Atr*). In bottom panel, sequence alignment of the promoter regions of *S. meliloti* (*Sm*) AbcR2 homologs in type strains of the same bacteria except *A. tumefaciens* (*Atr*). Numbers on top stand for nucleotide positions with respect to the experimentally determined AbcR1/2 transcription start sites (+1). Consensus sequences of the σ^{70} and σ^{H1} promoter signatures motifs (logos) are indicated at the bottom of the alignment. **(B)** Northern blot probing of total RNA obtained from *S. meliloti* Rm1021 wild-type and *rpoH1*, *rpoH2* and *rpoH1*/2 mutant strains grown in conditions indicated on top of the panels. 5S rRNA was probed as RNA loading control. Log, logarithmic growth phase (OD_{600} 0.4-0.5); Est, stationary growth phase (OD_{600} 2.5); NaCl, salt shock (400 mM for 1 h); 42ºC, heat shock (for 1 h); TY, complex tryptone-yeast medium; MM, minimal medium.

1.3.2. TARGETING POTENTIAL OF AbcR2 IN *S. meliloti*

1.3.2.1. Computational comparative prediction of AbcR2 mRNA targets

The identity of the mRNA target(s) is a key to decipher *trans*-sRNA functions and activity mechanisms in bacteria. We used the CopraRNA algorithm (Comparative prediction algorithm for small RNA targets) (Wright *et al*., 2013) (Material and Methods) to predict mRNA partners of AbcR2.

The nucleotide sequence of AbcR2 and its closest homologs in *S. meliloti* Rm1021, *S. medicae* WSM419, *S. fredii* NGR234, *A. tumefaciens* C58 (now called *A. fabrum*), *R. etli* CFN42 and *R. leguminosarum* bv. trifolii WSM1689 were used as queries in these predictions. This search of the targets of AbcR2 homologs in other bacteria returned a list of possible messengers highlighting enrichment for mRNA targets that encode components of ABC transporters (Figure 1.4). Moreover, this *in silico* prediction suggested that the regulation of targets translation by AbcR2 involves the interaction of the TSS in mRNAs with the conserved aSD motifs, M1, but also with an alternative aSD sequence, M2, (Figure 1.4) that remains single-stranded between the two first hairpins of AbcR1 and AbcR2 (Figure 1.1, B).

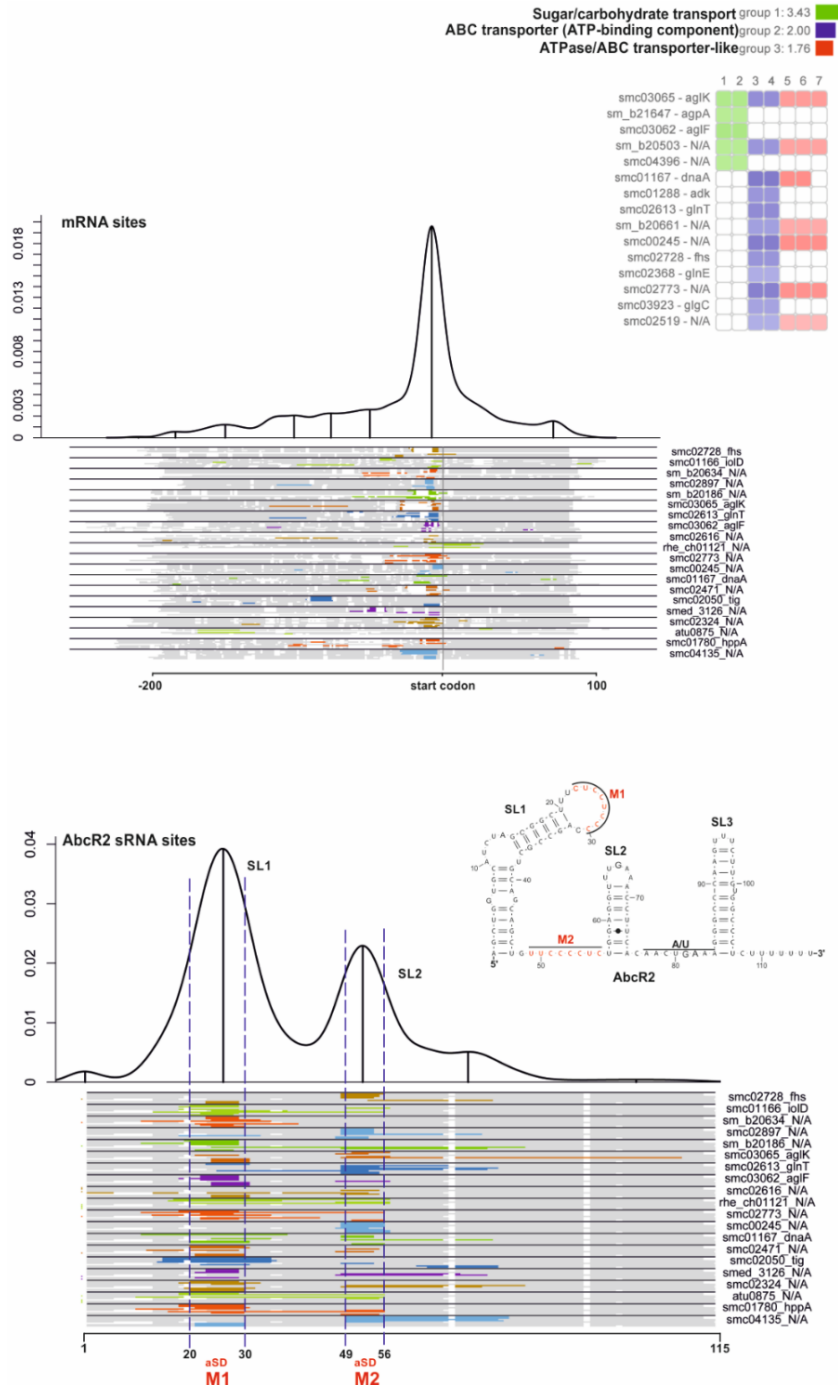

Figure 1.4. CopraRNA prediction of AbcR2 mRNA targets. Depiction of the putative interaction domains within the predicted mRNA targets of AbcR2 (mRNA sites) and within the sRNA (AbcR2 sRNA sites). The plots combine all predictions with a *P* value ≤ 0.01 in all included homologs. Local maxima indicate distinct interaction domains and are marked with upright lines. The schematic alignment of homologous sRNAs and targets at the bottom show the predicted interaction domains. The aligned regions are displayed in grey, gaps in white, and predicted interaction regions in color (color differences are for contrast only). The locus tag and gene name (if available) of a representative cluster member are given on the right.

1.3.2.2. The AbcR2-dependent periplasmic proteome

Subsequent to the *in silico* analysis that suggested AbcR2-mediated regulation of mRNAs targets encoding periplasmic components of ABC transporters (Figure 1.4), we performed the comparative and quantitative proteomic approach for the periplasmic fraction of Rm1021 and Rm1021Δ*abcR2* bacteria grown in MM subjected to osmotic upshift (Figure 1.5). The experimental strategy followed in this work was based on the technique known as SILAC (Stable Isotopic Labeling of Amino acids in Cell Culture) (Ong, *et al*., 2002), as described in more detail in Material and Methods.

The peptides of our samples were quantified and standardized in the two experimental conditions tested. The obtained M = 0 centered distribution indicated that only a small fraction of the periplasmic proteome responds to AbcR2 activity (Figure 1.5).

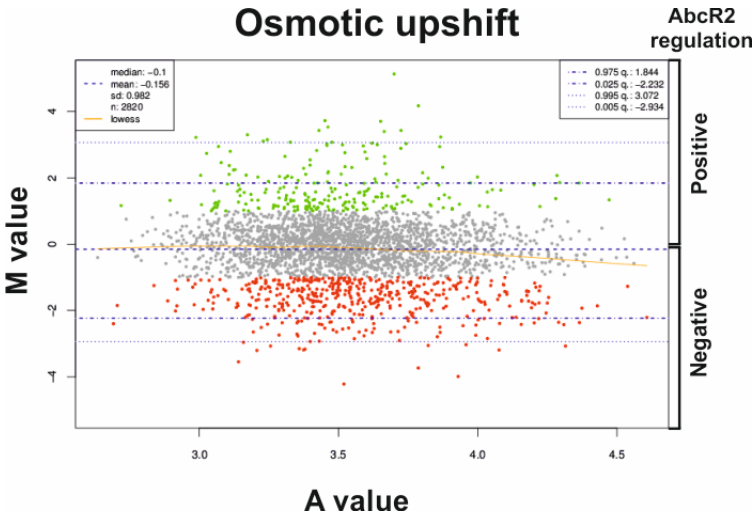

Figure 1.5. Quality control analysis of the proteomic approach. Representations of the M-A values of each peptide in MM with osmotic upshift. In grey color are represented the peptides with values -1 < M < 1. The repressed (M ≥ 1) or accumulated (M ≤ -1) peptides in the mutant strain are indicated in green and red, respectively.

The periplasmic proteome profiling evidenced that of the total proteins identified, 17 were down-regulated in the Rm1021Δ*abcR2* mutant, and 46 up-regulated, corresponding 18 of these 46 to ABC transporters (Table 1.1) that therefore represent putative AbcR2 mRNA targets.

Table 1.1. Transport proteins negatively regulated by AbcR2.

Gene ID	Putative substrate	Copra RNA	Hfq-bound[a]	Hfq regulation[a]	RpoH1 regulation[b]
SMa1462	Amino acids	No	Yes	No	No
SMa1860	Amino acids	#89	Yes	No	No
SMc01966	Spermidine/putrescine	No	Yes	Negative	No
SMc01827	Uracil/uridine	#112	Yes	Negative	No
SMc01597	Amino acids	No	Yes	Negative	No
phoD	Phosphate	No	No	No	No
SMc02259	Amino acids	No	Yes	Negative	No
potF	Putrescine	No	Yes	Negative	No
prbA	Proline betaine	No	Yes	Negative	No
SMa0392	Amino acids/polyamines	No	Yes	Negative	Negative
dppA2	Dipeptides	#27	Yes	Negative	No
SMc03864	Amino acids	No	No	No	No
nrtA	Nitrate	No	No	No	No
SMc02378	Glycine betaine	No	Yes	Negative	No
oppA	Dipeptides/amino acids	#127	Yes	Negative	Negative
SMa1755	Polyamines	#103	Yes	No	No
choX	Choline/nitrogen compounds	#73	Yes	Negative	No
SMc02873	Sugars	No	No	No	No

[a]Results obtained from a genome-wide profiling of Hfq-binding RNAs (Torres-Quesada *et al.*, 2014).
[a]The RpoH1 regulon is described in Barnett *et al.*, 2012.
In red, mRNAs predicted also by CopraRNA and further analyzed *in vivo*.

1.3.2.3. Mining the Hfq CoIP-RNA for AbcR2-mRNA regulatory pairs

Enrichment by Hfq CoIP combined with misregulation in the *hfq* mutant may indicate post-transcriptional control of certain mRNAs by Hfq-dependent sRNAs. In *S. meliloti*, targeting of the *livK* mRNA by the AbcR1 sRNA was the only sRNA-mRNA pair experimentally validated (Torres-Quesada *et al.*, 2013). Both homologous sRNAs, AbcR1 and AbcR2, have been reported to be Hfq-dependent (Vosset *et al.*, 2009; Torres-Quesada *et al.*, 2010, 2013). Accordingly, these sRNAs were enriched in Hfq CoIP-RNA libraries (Torres-Quesada *et al.*, 2014). CopraRNA and the

periplasmic proteome evidenced a number of additional target mRNAs of AbcR2, mostly coding for ABC transport systems. We selected three of these mRNAs, namely ***oppA*** *(SMb21196),* ***prbA*** (*SMc01642*) and ***SMa0495***, to further assess their regulation by AbcR2 *in vivo* using a double-plasmid reporter system (Torres-Quesada *et al*., 2013) (Material and Methods). As *livK*, these mRNAs code for amino acid binding proteins, are Hfq-bound and up-regulated in an *hfq* mutant (Torres-Quesada *et al*., 2014). CopraRNA predictions identified *oppA*, *prbA* and *SMa0495* with lower significance values compared to other putative mRNA targets, but OppA and PrbA were found in the proteomic analysis (Figure 1.5). The double-plasmid reporter assay revealed down-regulation of *prbA* and *SMa0495* by AbcR2 (and also by AbcR1, Torres-Quesada *et al*., 2014), but failed to demonstrate targeting of *oppA* (Figure 1.6). Specific antisense interactions between AbcR2 and these mRNAs were further predicted with IntaRNA (http://rna.informatik.uni-freiburg.de/IntaRNA /Input.jsp) (Smith *et al*., 2010) (Material and Methods) (Figure 1.6).

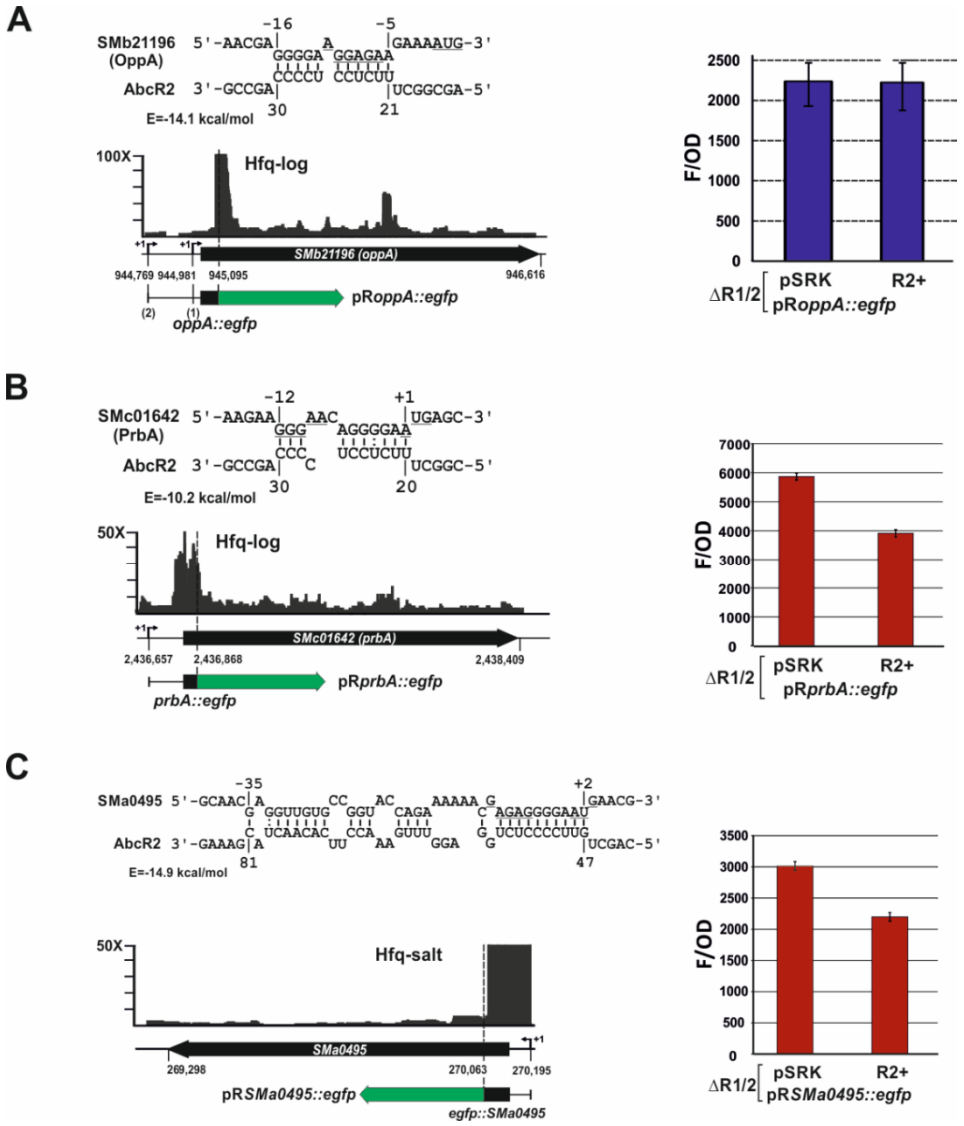

Figure 1.6. Targeting of the *oppA* (A), *prbA* (B) and *SMa0495* (C) mRNAs by the AbcR2 sRNA. IntaRNA predicted duplexes are shown, with the RBS and AUG start codons of the *oppA, prbA* and *SMa0495* mRNAs underlined. In these diagrams, numberings denote positions relative to the AUG start codon of the mRNA and the TSS of AbcR2. The predicted minimum hybridization energy (E) is indicated in each case. The IGB diagrams show the fold enrichment (vertical axis) of these mRNAs in the indicated Hfq CoIP–RNA library. Schematics of the reporter fusions cloned into plasmid pR_EGFP are shown below the genomic information of each mRNA. The histograms show the fluorescence of the reporter *S. meliloti* double deletion mutant (ΔR1/2) co-transformed with the target fusions and plasmids pSRK (empty control vector) and pSRK-R2 (R2+). Values reported are means and standard deviation of 48 fluorescence measurements, i.e., four determinations in three independent exponential cultures of four double transconjugants representing each plasmid combination. Fluorescence values were normalized to culture OD_{600} (F/OD). Background fluorescence from strains harboring pSRK and the empty pR_EGFP plasmid instead of the target fusions was subtracted from the fluorescence of target fusions.

As for *livK* (Torres-Quesada *et al*., 2013), AbcR2 was predicted to specifically target the ribosome binding site (RBS) and flanking nucleotides in the *prbA* and *SMa0495* mRNAs through aSD motifs (positions 50-59 or 21-31 within the sRNAs) (Figure 1.6, B and C). *prbA* and *SMa0495* were enriched in the Hfq-log and Hfq-salt libraries, respectively, with cDNA reads mostly clustering over their 5'-regions, including the respective 5'-untranslated region (5'-UTR) (Figure 1.6, B and C, left panels). Therefore, two genomic DNA fragments spanning from the native TSS to the 18th and 16th codons of *prbA* and *SMa0495* mRNAs, respectively, were translationally fused to *egfp* in plasmids pR*prbA::egfp* and pR*SMa0495::egfp*. These plasmids were used as reporters of AbcR2 activity in an *S. meliloti* Rm1021 AbcR1/2 double deletion mutant (ΔR1/2) (Torres-Quesada *et al*., 2013). Fluorescence of exponential cultures of the reporter strains harboring pR*prbA::egfp* and expressing AbcR2 from plasmid pSRK-R2 (R2+) was 33% less than that of control strain co-transformed with the same target fusion and the empty vector pSRK (Figure 1.6, B, right panel). Similarly, overexpression of AbcR2 resulted in reduction of the fluorescence signals conferred by presence of pR*SMa0495::egfp* by 25% (Figure 1.6, C, right panel).

Altogether these data validate *pbrA* and *SMa0495* as common targets of AbcR2 sRNA, thus rendering the the Hfq CoIP-RNA dataset as a valuable tool to search for mRNAs targeted by other *S. meliloti trans*-acting sRNAs in an Hfq-dependent manner.

1.3.2.4. AbcR2 loss-of-function alters expression of an array of salt-responsive genes

As a next step towards further characterization of the AbcR2-salt regulon, we explored the AbcR2-dependent molecular responses of *S. meliloti* upon osmotic upshift by profiling the transcriptomes of the Rm1021 strain and its deletion mutant derivative Rm1021Δ*abcR2* on Sm14kOLI microarrays (Figure 1.7).

Total RNA was obtained from bacteria grown to exponential growth phase (log RNA) in MM upon 1 h shock with 400 mM NaCl. This experiment identified 71 AbcR2-dependent mRNAs, i.e. displaying at least 2-fold changes in their abundance between the two strains, with 56 down-regulated and 15 up-regulated in Rm1021Δ*abcR2* (Table 1.2)

Table 1.2. AbcR2-dependent genes (mRNAs).

Gene ID	Protein product	M value
mRNAs positively regulated by AbcR2		
SMa1225	transcriptional regulator FixK	-2.41
SMb20434	hydrolasepeptidase	-2.19
SMc02051	hypothetical protein	-1.89
SMc01266	hypothetical protein	-1.88
SMb20435	hypothetical protein	-1.85
SMc03253	L-proline 3-hydroxylase	-1.79
SMa0762	transcriptional regulator FixK	-1.78
SMc03780	hypothetical protein	-1.65
SMc02655	hypothetical protein	-1.60
SMc01489	signal peptide protein	-1.56
SMb20704	GlgA glycogen synthase	-1.50
SMc02900	hypothetical protein	-1.44
SMb21446	GlgX2 glycosyl hydrolase	-1.43
SMb21674	hypothetical protein	-1.42
SMb20431	arylmalonate decarboxylase	-1.40
SMb20430	amino acid ABC transporter permease	-1.36
SMc00591	hypothetical protein	-1.35
SMc03999	hypothetical protein	-1.35
SMb20251	hypothetical protein	-1.33
SMb20960	ExoN UDP-glucose pyrophosphorylase	-1.32
SMb20346	efflux protein	-1.31
SMb21448	DNA polymerase related protein	-1.31
SMb20429	amino acid ABC transporter permease	-1.29
SMc01410	lipoprotein transmembrane	-1.26
SMc04215	CobS cobalamin synthase	-1.26
SMc02146	phosphate-binding periplasmic protein	-1.24
SMb21507	amino acid transporter, exporter protein	-1.23
SMc02396	outer membrane protein	-1.22
SMc02374	hypothetical protein	-1.22

SMc01875	LpxCUDP-3-O-3-hydroxymyristoylN-acetylglucosamine deacetylase	-1.19
SMb20994	hypothetical protein	-1.19
SMb21169	ArsR family transcriptional regulator	-1.18
SMc02373	hypothetical protein	-1.18
SMb20571	aliphatic sulfonate ABC transporter	-1.16
SMb20253	hypothetical protein	-1.15
SMa2359	hypothetical protein	-1.15
SMb20934	ExsF two-component response regulator protein	-1.13
SMc00122	Pbp penicillin-binding protein	-1.13
SMb20092	hypothetical protein	-1.12
SMc03900	cyclic beta-1,2-glucan ABC transporter	-1.12
SMc04190	signal peptide protein	-1.08
SMb21585	hypothetical protein	-1.08
SMc01370	response regulator PleD	-1.08
SMc04451	chloramphenicol phosphotransferase	-1.07
SMc01507	hypothetical protein	-1.07
SMb21133	sulfate ABC transporter substrate-binding protein	-1.07
SMc03844	hypothetical protein	-1.07
SMc01263	hypothetical protein	-1.06
SMa2357	CyaO guanylate cyclase	-1.05
SMc00895	hypothetical protein	-1.04
SMc04194	transmembrane protein	-1.03
SMa1223	FixM	-1.02
SMb21270	transcriptional regulator	-1.02
SMc02656	hypothetical protein	-1.01
SMa1077	Nex18 symbiotically induced protein	-1.01
SMb21691	nitrilotriacetate monooxygenase subunit A	-1.00
mRNAs negatively regulated by AbcR2		
SMb21234	transposase of insertion sequence ISRm1	1.01
SMc04318	Csp1 cold shock transcription regulator protein	1.04
SMc03770	RplU 50S ribosomal protein L21	1.04
SMa0445	TRm1a transposase	1.08
SMc01210	hypothetical protein	1.09
SMc01959	transposase ISRM1	1.10
SMc03898	transposase ISRM1	1.11
SMb20989	stomatin-like protein	1.13
SMb20918	transposase of insertion sequence ISRm1	1.14
SMa1615	TRm1a transposase	1.19
SMc02298	transposase ISRM1	1.27

SMc01563	SigA RNA polymerase sigma factor RpoD	1.30
SMc03295	transposase ISRM1	1.34
SMa0011	SelA selenocysteine synthase	1.36
SMc02845	antibiotic resistance protein	1.87

In red, mRNAs involved in the glycogen regulation.

According to the *S. meliloti* Rm1021 genome sequence annotation, functional clustering of these genes revealed that 26 are hypothetical proteins, 11 encode metabolic functions, 10 are related to transport and membrane processes, 9 represent transcription and signal transduction processes, and eight are enzymes involved in transposition mechanisms (Figure 1.7, left panel). Overall, the differential regulation of all identified mRNAs between Rm1021 and Rm1021Δ*abcR2* was rather weak, indicating that AbcR2 has only limited influence on the regulation of these mRNAs (Figure 1.7, right panel). Comparison of the AbcR2 regulon and the set of genes differentially expressed in the wild-type strain upon an osmotic upshift (virtual microarray hybridization) revealed downregulation in the mutant of a number of genes that are induced in response to high external osmolarity (Table 1.2 and Figure 1.7, right panel). This group includes *glgX2* and *glgA,* coding for proteins for the biosynthesis of the osmolyte glycogen, a couple of genes coding for proteins involved in membrane trafficking (*SMb20346* and *SMb21507*), and genes related to transcription and signal transduction (Table 1.2 and Figure 1.7, right panel) These mRNAs do not have complementarity to AbcR2 and therefore, they must be regarded as secondary molecular targets of this sRNA.

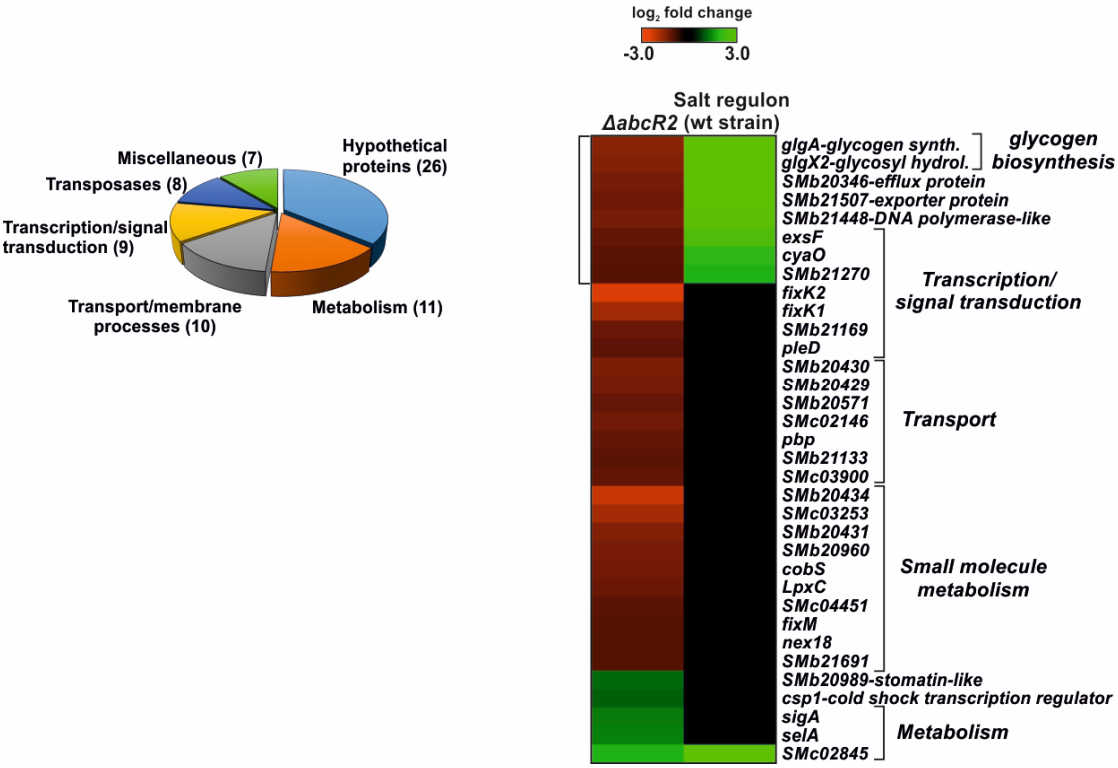

Figure 1.7. The AbcR2-dependent transcriptome. Number and functional categories of mRNAs differentially accumulated in the Rm1021Δ*abcR2* mutant with respect to the wild-type strain when both were grown in MM upon an osmotic upshift (1 h, 400 mM NaCl). Heatmap, expression of salt-responsive genes with predicted function. Plotted are M values of changes (log_2 fold) in mRNA abundance in the comparison wild-type strain Rm1021 vs. Rm1021Δ*abcR2* mutant when both were subjected to the osmotic upshift. In the color scale, red represents down-regulation and green up-regulation with respect to the reference (wild-type) in each comparison. The bracket denotes the subset of functional categories of mRNAs differentially accumulated in the mutant. Name and putative function of each gene are indicated to the right.

Taken together, these results suggest a discrete global effect of AbcR2 on the physiology of *S. meliloti* subjected to osmotic stress.

1.4. DISCUSSION

The repertoire of non-coding RNAs expressed by the legume endosymbiont *S. meliloti* is one of the best characterized among those of its α-proteobacterial counterparts (del Val *et al*., 2007, 2012; Schlüter *et al*., 2010, 2013; Sallet *et al*., 2013). However, current information about the function of these transcripts is still scarce. The first set of sRNAs identified in the reference strain *S. meliloti* Rm1021 included the AbcR2 transcript, encoded in tandem with its homologous AbcR1, both similar in sequence and structure, with genomic boundaries experimentally determined by independent approaches (del Val *et al*., 2007; Schlüter *et al*., 2010). Subsequent analysis revealed AbcR2 as an example of Hfq-dependent sRNA potentially important for *S. meliloti* stationary growth and adaptation to abiotic stresses (Torres-Quesada *et al*., 2013; Harfouche *et al*., 2014). Here, we provide further insights into the function of AbcR2 revealing that this sRNA integrates the regulon of RpoH1 (σ^{H1}factor) acting downstream in the post-transcriptional regulation of multiple transporter mRNAs coding for proteins involved in the uptake of amino acids and other nitrogen sources.

The alternative RNA polymerase σ factor RpoH1 is responsible of the Hfq-dependent AbcR2 sRNA transcription

One of the general characteristics of *trans*-sRNAs is their expression in response to specific environmental stimuli. The analysis of the AbcR1 and AbcR2 expression in *S. meliloti* Rm1021 showed divergent likely unrelated patterns, with AbcR2 accumulating upon entry of bacteria into stationary phase and under a number of abiotic stresses (Torres-Quesada *et al*., 2013). This information together with reported RNASeq data performed in this genome (Schlüter *et al*., 2010) suggest that the *abcR1/2* promoters are differentially regulated, responding to different environmental cues. The opposite expression patterns of AbcR1 and AbcR2 contrast with those of

their *A. tumefaciens* homologs, which are also encoded in tandem in the circular chromosome of this bacterium, but showed identical expression profiles (Wilms *et al.*, 2011). Here, we have extended the study of AbcR2 expression, performing an *in silico* analysis to search for conserved motifs in its promoter regions and Northern experiments to validate the predictions. Both approaches provided further support to the differential regulation of *abcR1/2* genes in *S. meliloti*.

The multiple alignment of the promoter sequences of the αr15 sRNAs grouped AbcR2 with its homologs in bacteria of the *Sinorhizobium*, *Rhizobium* and *Agrobacterium* species. The analysis of the conserved sequences of their promoter regions predicted binding sites for the σ factor RpoH1, one of the alternative σ subunits of the RNA polymerase holoenzyme that do not belong to the extracytoplasmic function (ECF) subfamily (Galibert *et al.*, 2001). Northern blot analysis of RNA from bacteria devoid of RpoH1, RpoH2 and of both proteins confirmed specific transcription of AbcR2 by RpoH1.

These results will contribute to decipher the possible role of the regulatory mechanisms underlying AbcR2 activity in the context of the RpoH regulon governing the general stress adaptation (stationary phase) and the hyperosmotic and heat shock pathways in *S. meliloti*.

Fine-tuning of nutrient uptake by the Hfq-dependent AbcR2 sRNA

The identity of the mRNAs directly regulated by the *trans*-sRNAs is, certainly, the most revealing data of their functions. The interaction of these riboregulators with their target mRNAs through mating, in most cases 6-7 nt (sRNA seed sequence) that are generally unpaired in any of the regions of the sRNA, is sufficient for effective mRNA regulation, even permitting a single sRNA for its interaction with multiple messages. This kind of interaction makes the prediction / identification of the target

mRNAs a complex task that constitutes a real challenge for progress in the functional characterization of *trans*-sRNAs. In fact, the *trans*-sRNAs with targets validated experimentally are in an absolute minority compared to the number known in other bacterial species (Storz *et al*., 2011). First approaches to this aim are bioinformatics predictions and transcriptomic and proteomic comparative analysis of the sRNAs mutants to allow identification of sRNAs-regulated genes in bacteria.

The technical proteomic approach selected is relevant because post-transcriptional riboregulatory mechanisms may result in changes in protein levels without being reflected in changes in mRNA levels and thus escape transcriptomic profiling, and because sometimes the control of gene expression can be subtle and then requires more sensitive techniques (Sobrero *et al*., 2012; Fan *et al*., 2013). The bioinformatics predictions are potentially applicable to any bacterium with sequenced genome, based on the thermodynamics of the sRNA-mRNA interaction and on the accessibility of the sequences supposedly involved in the mating of both molecules. As expected, this type of predictions generated in our work revealed a long list of 5' ends of mRNAs as probable AbcR2 targets. These programs do not account other factors presumably relevant to a productive riboregulation, such as the association sites of chaperone RNAs (e.g. Hfq) in the duplex sRNA-mRNA or other specific structural elements, which could substantially reduce the lists of potential targets for sRNAs (Tjaden *et al*., 2006). However, the functional relationship between candidate mRNAs is considered a first index of confidence in predictions (Tjaden *et al*., 2006). In this sense, among the multiple target mRNAs predicted by these approaches for sRNAs expressed by *E. coli* and *Salmonella* as GcvB, RybB and RybB, for predominately those that encode ABC transporters, outer membrane proteins (OMPs) and proteins that contain or accumulate iron, respectively, exist strong experimental evidence (Massé and Gottesman, 2002; Papenfort *et al*., 2006; Sharma *et al*., 2008, 2011). This functional relationship was evident in the candidate catalog for AbcR2, with vast majority of mRNAs specifying the structural components of ABC transporters, fundamentally the

periplasmic proteins involved in the recognition and binding of the specific substrates, and on them we focused our attention.

In certain cases, specific regions of the mRNAs, rather than the full length transcripts, appeared enriched in the CoIP-RNA libraries. These fragments, mostly derived from UTRs, could correspond to novel Hfq-dependent *trans*-sRNAs resulting from parallel transcriptional output or post-transcriptional processing of certain mRNAs (Vogel *et al*., 2003; Kawano *et al*., 2005; Loh *et al*., 2009; Chao *et al*., 2012) Alternatively, these enrichment patterns likely reveal primary high affinity binding sites for Hfq that remain protected during exoribonucleolytic degradation of the message upon base pairing with a *trans*-acting sRNA. Indeed, these cDNA clusters preferentially mapped to the 5' regions of the *S. meliloti* mRNAs, which are the most common binding sites of this type of riboregulators in bacteria (Storz *et al*., 2011). The classical model of RNA cycling on Hfq predicts that the hexameric Hfq ring uses its two faces to simultaneously bind the sRNA-mRNA regulatory pairs for antisense interaction (Gerhart and Wagner, 2013). In line with this notion, many targets of well-characterized *Salmonella* sRNAs have been identified within the pool of Hfq-associated transcripts (Sittka *et al*., 2008). Similarly, among the identified Hfq-bound mRNAs, those encoding the ABC transporters PrbA and SMa0495 were validated here as new negatively regulated targets of the *S. meliloti* Hfq-dependent AbcR2 sRNA.

The unambiguous validation of target sRNA-mRNA interactions always requires the accumulation of several experimental evidences (Urban and Vögel, 2007). In our case, as a direct experimental test of the specificity of the interactions, we used an *in vivo* genetic test based on the observations of visible changes in the fluorescence produced from translational fusions of the 5' mRNAs to *egfp*, when they were co-expressed with the AbcR2 sRNA. The use of GFP as a reporter protein that does not require additional chromogenic substrates will allow future studies about gene regulation at the single cell level, which will be of great use also in the research of ribo-regulation

in bacteria. Finally, this double-plasmid genetic assay is qualitatively designed but supports multi-level quantification, fluorescence determination by flow cytometry or Western blot titration of the translation rate of the fusion protein in each assay. It is predictable, therefore, that the introduction of improvements in the assay and its quantification may reveal interactions of mRNAs-sRNAs not evidenced in our studies.

The reported up-regulation of *prbA* and *SMa0495* in the absence of Hfq and the overrepresentation of their 5' regions in CoIP-RNA are consistent with a canonical mechanism for AbcR2 activity involving antisense interaction with complementary sequences within or in the vicinity of the RBS of their targets leading to blocking of translation and subsequent mRNA decay. Our findings further support the hypothesis that the AbcR family of α-proteobacterial sRNAs is involved in the fine-tuning of nutrient uptake by selective repression of multiple ABC transporters of nitrogen compounds, similar to GcvB sRNA in enterobacteria (Sharma *et al*., 2007, 2011; Pulvermacher *et al*., 2009; Wilms *et al*., 2011; Caswell *et al*., 2012; Torres-Quesada *et al*., 2013). This pervasive regulation of nutrient uptake would contribute to coordinate downstream responses to external hyperosmolarity such as transcription, biosynthesis of certain osmolytes and membrane trafficking.

APPENDIX 1

Table 1.3. Bacterial strains and plasmids used in this work.

***S. meliloti* strains**	Description	Reference/Source
Rm1021	Wild-type SU47 derivative, Sm^r	Meade *et al.*, 1982
1021ΔR1	AbcR1 deletion mutant; Err, Smr	Torres-Quesada *et al.*, 2013
1021ΔR2	AbcR2 deletion mutant; Er^r, Sm^r	Torres-Quesada *et al.*, 2013
1021ΔR1/2	AbcR1/2 deletion mutant; Er^r, Sm^r	Torres-Quesada *et al.*, 2013
VO3128 (*rpoH1*::*aadA*)	1021 carrying gene-disrupting constructs at the respective native genomic locations	Oke *et al.*, 2001
AB3 (*rpoH2*::*aacCI*)	1021 carrying gene-disrupting constructs at the respective native genomic locations	Bittner and Oke, 2006
AB9 (*rpoH1*::*aadA rpoH2*::*aacCI*)	1021 carrying gene-disrupting constructs at the respective native genomic locations	Bittner and Oke, 2006
Plasmids	**Description**	**Reference/Source**
pSRKKm	pBBR1MCS-2 derivative containing the *lac* promoter, *lacIq*, *lacZ*α^+, Km^r	Khan *et al*., 2008
pSRK-C	Engineered pSRKKm lacking the LacIQ operator; Km^r	Torres-Quesada *et al*., 2013
pSRK-R2	pSRK_C carrying the *abcR2* coding sequence	Torres-Quesada *et al*., 2013
pBB-*egfp*	pBBR1MCS-2 derivative for generation of promoter- *egfp* fusions; Km^r	This work
pBB*syn::egfp*	pBBR1MCS-2 derivative constitutively expressing *egfp* ; Km^r	This work
pR_EGFP	Vector for generating of target mRNA-*egfp* translational fusions	Torres-Quesada *et al*., 2013
pR*oppA::egfp*	pR_EGFP expressing the *oppA::egfp* translational fusion; Ap^r, Tc^r	This work
pR*prbA::egfp*	pR_EGFP expressing the *prbA::egfp* translational fusion; Ap^r, Tc^r	This work
pR*SMa0495::egfp*	pR_EGFP expressing the *SMa0495::egfp* translational fusion; Ap^r, Tc^r	This work

Table 1.4. Oligonucleotides used in this work.

Name	**Sequence (5'-3')**
AbcR1	ACTGGGAGGAGAACGGAGCAAAGAT
AbcR2	GAGGAGAAAGCCGCTAGATGCACCA
secSRK	TTCCATTCGCCATTCAGGCT
FwSRK	ACTAAAGGGATCCAAAGCTGGAGC
RvSRK	GCTCACAATTGGATCCAACATACGAG
SMc01642_F	GGATCCGAACAGCGCGGATAACGCGCAA
SMc01642_R	GCTAGCTTTGCCGAGCATGACCTGAC
SMa0495_F	GCTAGCCATTGCAACCGCCGACCCCA

SMa0495_R	GGATCCTAGAAGGCATCGAATTTCCA
SMb21196_F	GGATCCTGGTGCTTCCGTGCAAGCAG
SMb21196_R	GGATCCGTTCGTTTGGGCCTGATATC

CHAPTER 2

PRIMARY CHARACTERIZATION OF THE SALT-INDUCED NfeR1 sRNA

2.1. INTRODUCTION

2.1.1. IMPLICATIONS OF HIGH SALINITY FOR THE RHIZOBIA-LEGUME SYMBIOSIS

High salinity is one of the environmental factors that contribute most to the deterioration of soil structure and its fertility, thus limiting agricultural productivity. Nearly 40% of soils worldwide cause potential salinity problems. The general solution applied to overcome this problem in recent decades has been the extensive use of chemical fertilizers and salt-tolerant plants (Zahran, 1999). However, the production and application of fertilizers are costly practices, both economically and ecologically. Rhizobia are sensitive to salt or osmotic stresses in the free-living state and during the symbiotic process since these conditions may inhibit the initial steps of the symbiotic interaction like root colonization, nodule infection and nodule development, and also have a negative effect on nitrogen fixation (Zahran, 1991). In order to avoid a loss in nitrogen fixation capacity, it is important to understand the underlying mechanisms employed by the bacteria for osmoadaptation under high salinity conditions. This led to efforts to isolate salt-tolerant plants and rhizobial strains mediating efficient nodulation under high salinity conditions, although with limited success (Ibragimova *et al.,* 2006; Moawad and Beck, 1991).

2.1.2. IDENTIFICATION OF THE NfeR1 sRNA IN *Sinorhizobium meliloti*

In this work, the function of a new *S. meliloti trans*-sRNA has been approached, whose expression is induced in high salinity conditions. A previous study identified this sRNA, referred to as SmrC14, by a genome-wide computational analysis of intergenic regions conducted in the reference *S. meliloti* strain Rm1021 (del Val *et al*., 2007). Northern hybridization experiments confirmed that the predicted *smrC14* locus

expresses a single transcript of the expected size, which accumulated differentially in free-living and endosymbiotic bacteria. TAP-based 5'-RACE experiments mapped the transcription start site (TSS) of the full-length *smrC14* transcript to the 1,667,613 nt position in the *S. meliloti* Rm1021 genome, while the 3'-end was assumed to be located at the 1,667,491 nt position matching the last residue of the consecutive stretch of uridines of a bona fide Rho-independent terminator (del Val *et al*., 2007). Soon after this, an independent study confirmed the expression of this sRNA in the closely related strain Rm2011 (Valverde *et al,* 2008). Subsequent deep sequencing-based characterization of the small RNA fraction (50-350 nt) of *S. meliloti* Rm2011 further confirmed the expression of SmrC14 (here referred to as SmelC397), and mapped the 5'- and 3'-ends of the full-length transcript to positions similar to those determined by Val *et al.,* 2007 in the *S. meliloti* Rm1021 genome (Schlüter *et al.,* 2010). Homologs of this sRNA are commonly encoded in multiple copies per genome and can be found in a range of bacteria integrating the so-called αr14 (α-proteobacteria RNA 14) family of α-proteobacterial sRNAs (del Val *et al.,* 2012; Reinkensmeier and Giegerich, 2015). SmrC14 was renamed Smr14C2 when six putative αr14 loci were identified in the *S. meliloti* Rm1021 reference genome (del Val *et al.,* 2012).

Here, we show that Smr14C2 is strongly transcribed in response to salt stress and throughout the symbiotic interaction, influencing adaptation of free-living bacteria to high salinity as well as the symbiotic performance of *S. meliloti* on alfalfa roots. According to its associated symbiotic phenotypes, Smr14C2 has been renamed NfeR1 (Nodule Formation Efficiency RNA). Furthermore, we predicted the interaction of this sRNA to possible mRNA targets encoding the periplasmic protein of ABC transport system, and we revealed the up-regulation of the aerobic nitrogen metabolism pathway in a mutant strain lacking NfeR1.

2.1. EXPERIMENTAL SETUP

2.2.1. DISTRIBUTION AND STRUCTURE OF NfeR1

The analysis of the conservation and predictions of sRNA secondary structures were performed in collaboration with Dr. Coral del Val (University of Granada) as described in del Val *et al*. (2012). The nucleotide sequence of NfeR1 was initially used as query to search against the **Rfam** database (version 10.0) (http://www.sanger.ac.uk/Software/Rfam). This homology search rendered no matches to known bacterial sRNA in this database. The NfeR1 sequence was subjected to **BLAST** analysis with default parameters against all the available bacterial genomes in the time of the analysis (http://www.ncbi.nlm.nih.gov). The regions exhibiting significant homology to the query sequence (78-89% similarity) were used to generate automated **Infernal** alignment (version1.0) (covariance model; CM) (Nawrocki and Eddy, 2007; Nawrocki *et al*., 2009) for the αr14 family. This initial alignment (CM), in Stockholm format, was hand-curated and manually inspected to deduce a consensus secondary structure for the family, using **RNAalifold** (Bernhart *et al*., 2008). The prediction of secondary structure of the NfeR1 was determined experimentally using the algorithm **RNAfold** (Hofacker *et al*., 1994). Both, RNAalifold and RNAfold are part of the core programs of the **Vienna RNA package** (http://rna.tbi.univie.ac.at/) (Gruber *et al*., 2008). The consensus structure was also independently predicted with the program **locARNATE** (http://www.bioinf.uni-freiburg.de/Software/LocARNA/) (Will *et al*., 2007) in an automatic manner and differences reconciled giving priority to the structural conservation. Given the initial hand-curated structural alignment of close homologs, Infernal was used to interrogate the same set of bacterial genomes, searching for new members of the covariance models (CMs). The alignment process was repeated three times. The candidates obtained with the Infernal model were selected as members of a given family if their Infernal E-value was 10^{-03} or lower, or after manual inspection

for those with higher Infernal E-values. The hierarchical cluster-tree for αr14 family is derived by **WPGMA clustering** of the pairwise alignment distances and the optimal number of clusters was calculated from the tree using **RNAclust** (http://www.bioinf.uni-leipzig.de/~kristin/Software/RNAclust/).

In order to study the microsynteny of the αr14 family, we located and extracted the flanking genes of its members (del Val *et al*., 2012). Non-annotated ORFs were further annotated using **Blast2GO** (Conesa *et al*., 2005; Vinayagam *et al*., 2006) and the high-throughput pipelines **ProtSweep**, and **DomainSweep** (del Val *et al*., 2006). The obtained results were later manually inspected in order to annotate and predict a biological function for these ORFs. In the few cases where the predicted sRNAs overlapped ORFs, the same procedure as with the flanking genes was carried out. ORFs shorter than 30 aa, that neither showed similarity with any database entry, nor motif or signatures when searched against family and motif databases such as **Interpro** (Hunter *et al*., 2009), **PFAM** (Finn *et al*., 2010) or **Smart** (Letunic *et al*., 2009) were considered as miss-annotations and thus not registered in the genomic context graph of the αr14 family (del Val *et al*., 2012).

Sequence alignments of *nfeR1* copies were generated with BioEdit (http://www.mbio.ncsu.edu/BioEdit/bioedit.html).

2.2.2. CONSTRUCTION OF *S. meliloti* MUTANTS AND DERIVATIVE STRAINS

2.2.2.1. Construction of *S. meliloti* mutants altered in NfeR1 expression

2.2.2.1.1. Generation of the *S. meliloti* SmΔ*nfeR1* mutant

The first *S. meliloti nfeR1* deletion mutant was generated by del Val *et al*. (2007) in strain Rm1021 (Galibert *et al.*, 2001) by replacement of the chromosomal sRNA gene

by a 195 bp erythromycin resistance cassette (SSDUT1) following a previously described procedure in Torres-Quesada *et al*. (2013). First, a 2,062-bp DNA region encompassing the *nfeR1 locus* and flanking sequences of 950 bp (downstream) and 989 bp (upstream) was PCR amplified using Rm1021 genomic DNA as template and primers 5-C14 and 3-C14 (Table 2.5, Appendix 2), the latter adding a *Xba*I site to the 3'-end of the fragment. This amplification product was inserted into the pGEM-T Easy vector (Promega), yielding pGEMgNfeR1. This plasmid was amplified with the divergent primers flanking the NfeR1 *locus* 5-C14i and 3-C14i, both carrying an internal *Kpn*I recognition site (Table 2.5, Appendix 2). The resulting PCR product was restricted with *Kpn*I and self-ligated to generate pGEMΔNfeR1. The SSDUT1 cassette (Torres-Quesada *et al*., 2013) was then inserted into the unique *Kpn*I site of pGEMΔNfeR1 generating pGEM-EryΔNfeR1. The insert of this plasmid was retrieved by restriction with *Xba*I and *Sph*I (site immediately downstream of the 3-C14 annealing sequence) and ligated to the suicide vector pK18*mobsacB* (Schäfer *et al*., 1994) yielding pK18Δ*nfeR1*. This plasmid was first mobilized by biparental mating to *S. meliloti* Rm1021 to induce allelic replacement by double cross-over events (Torres-Quesada *et al*., 2010). In this work, the pK18Δ*nfeR1* construct was also mobilized by triparental mating (using pRK2013 as a helper plasmid) to *S. meliloti* Sm2B3001 (Bahlawane *et al*., 2008) (Figure 2.1, A). A series of colonies exhibiting a $Km^s/Suc^r/Er^r$ phenotype were finally checked for the targeted deletion by colony PCR with the primers SecC14i and SecC14 (Table 2.5, Appendix 2), and following *Kpn*I restriction of the PCR product (Figure 2.1, B).

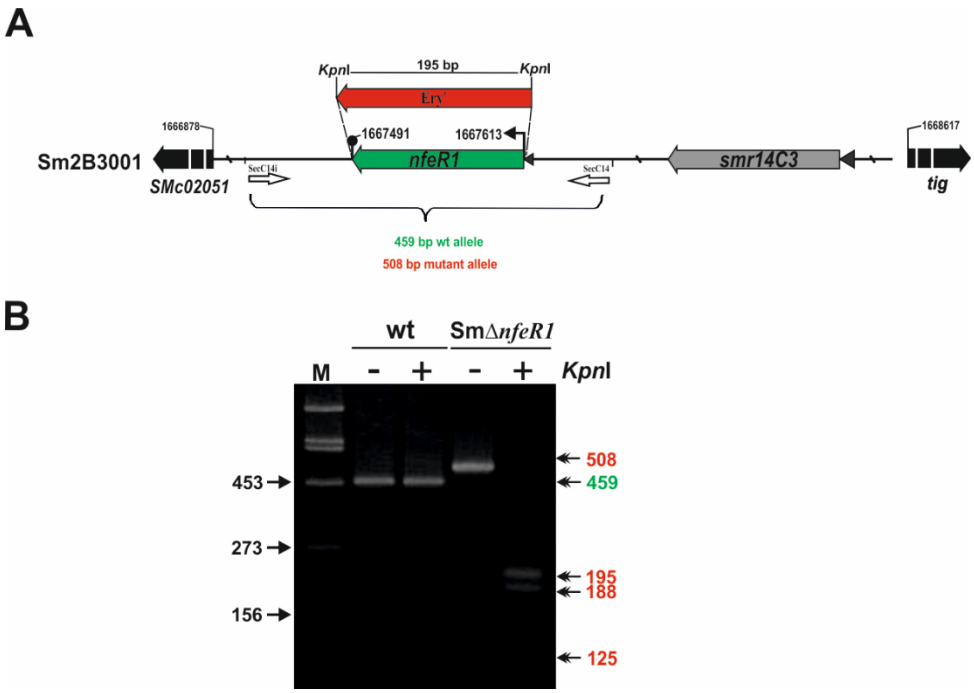

Figure 2.1. Generation of the *S. meliloti* SmΔ*nfeR1* mutant. (A) Depiction of the *nfeR1* genomic region in *S. meliloti* Sm2B3001 where the generated SmΔ*nfeR1* mutant is indicated with the position of the erythromycin resistance cassette (Ery^r) replacing the *nfeR1* gene. The sizes of the relevant fragments and key restriction sites are indicated. **(B)** The deletion was first verified by PCR with primers SecC14i and SecC14 (Table 2.5, Appendix 2) that generated 459 bp and 508 bp amplification products in the parental and mutant strains, respectively. The mutant allele could be restricted with *Kpn*I, releasing the 125 bp Ery^r cassette (hardly detectable in the stained gel)

2.2.2.1.2. Complementation of the *S. meliloti* SmΔ*nfeR1* mutant

The SmΔ*nfeR1* mutant was complemented with plasmids pSRKNfeR1 (expressing *nfeR1* constitutively) or pJBNfeR1 (expressing *nfeR1* from its own promoter) that were generated as follows.

To construct **pSRKNfeR1**, the full-length *nfeR1* coding sequence (i.e. from the transcription start site to the last residue of the Rho-independent terminator) was amplified by PCR using Sm2B3001 genomic DNA as template and the pair of primers SmrC14Fw/SmrC14Rv that incorporate *BamH*I and *Sac*I sites to the 5'- and 3'-end of

the amplicon, respectively (Table 2.5, Appendix 2). This PCR product was first cloned into the pGEM-T Easy vector, next retrieved as a *BamH*I-*Sac*I fragment and finally inserted into the medium-copy plasmid pSRK-C (Torres-Quesada *et al.*, 2013) to generate pSRKNfeR1 (Figure 2.2, A).

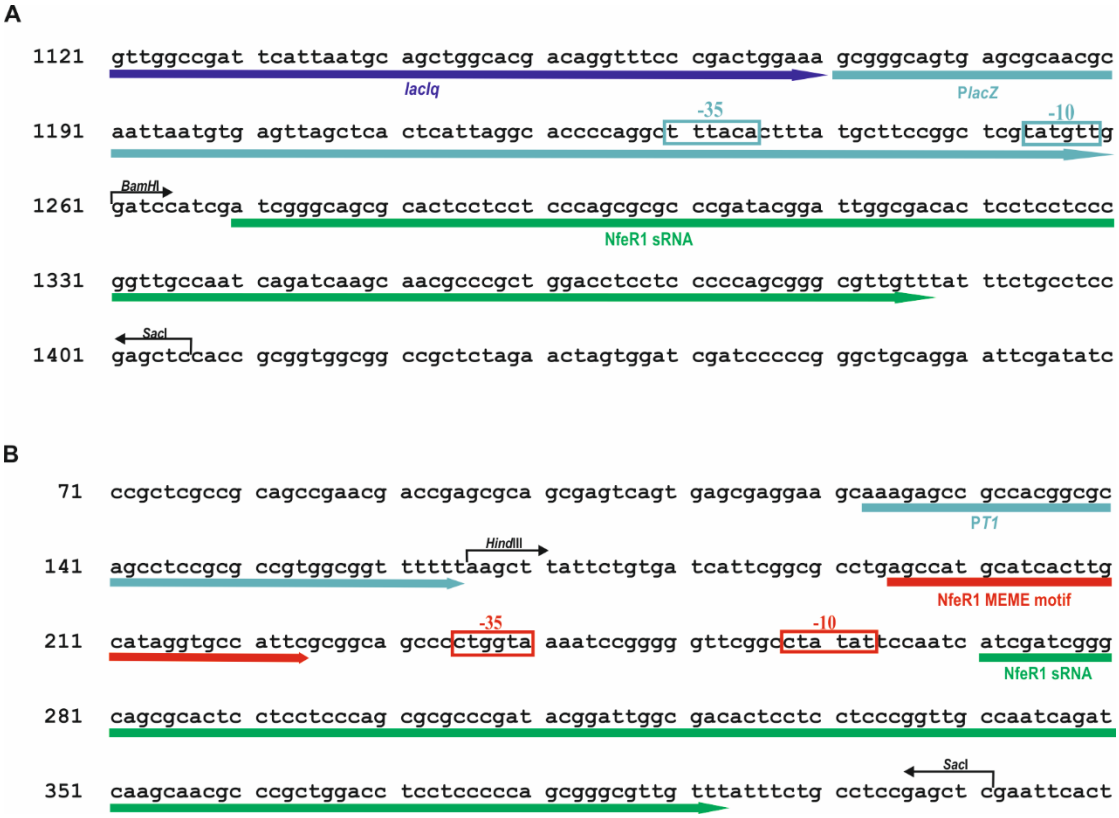

Figure 2.2. Design of constructs used for complementation of *S. meliloti* Sm*ΔnfeR1*. (A) Section of the nucleotide sequence of pSRKNfeR1. The sRNA was cloned as *BamH*I/*Sac*I fragment. In light blue, the *lacZ* promoter (P*lacZ*) is shown upstream of the NfeR1 sRNA (in green), and the -35 and -10 boxes are indicated. **(B)** Section of the nucleotide sequence of pJBNfeR1. The sRNA was cloned as *Hind*III/*Sac*I. P*lac* of the pJB3Tc19 vector was replaced by a Rho-independent transcriptional terminator (T1, in blue). Consensus sequences of the σ^{70} promoter signature with the -35 and -10 boxes, and the conserved MEME motif (in red), will be further analysed in this chapter, are indicated.

To construct **pJBNfeR1**, the *lac* promoter of the low-copy broad host-range vector pJB3Tc19 (Blatny *et al.*, 1997) was first replaced by a Rho-independent transcriptional terminator (T1) generated by annealing of the oligonucleotides T1_F

and T1_Rv that generate *Sap*I and *Hind*III compatible overhangs (Table 2.5, Appendix 2). Insertion of T1 between these two sites in pJB3Tc19 yielded pJB-T1. The *nfeR1* locus along with its promoter region was then PCR-amplified from Sm2B3001 genomic DNA with primers P14C2_H and SmrC14Rv (Table 2.5, Appendix 2). This amplicon was first cloned into pGEM-T Easy, then recovered as a *Hind*III and *Sac*I fragment and finally inserted between these sites in pJB-T1 to generate pJBNfeR1 (Figure 2.2, B).

Plasmids pSRKNfeR1 and pJBNfeR1 were conjugated into Sm*ΔnfeR1* by biparental mating involving *E. coli* S17-1. All PCR amplifications required for cloning were performed with the proof-reading Phusion™ High-Fidelity DNA polymerase (Thermo Scientific). Plasmid inserts were always checked by sequencing to confirm the absence of PCR-introduced mutations.

2.2.2.2. Reporter transcriptional fusions with *egfp*

The **transcriptional fusions** reporting promoter activity were generated in the vector pBB-*egfp* carrying promoterless *egfp*. This plasmid was constructed by insertion of a *Sac*I-*Kpn*I fragment from pHC125, consisting of the *egfp* coding sequence and a plasmid stability region flanking a multiple cloning site, into pBBR1MCS-2 (Kovach *et al*., 1994). The full-length *nfeR1* promoter region was PCR-amplified with the primer pair P14C2Fw/P14C2Rv (Table 2.5, Appendix 2) and the resulting product was inserted into pBB-*egfp* as a *Spe*I-*Xba*I fragment, yielding p*nfeR1*::*egfp*. A deleted version of the *nfeR1* promoter was generated by annealing oligonucleotides P14C2_54 and P14C2_54i (Table 2.5, Appendix 2) to create a *Spe*I-*Xba*I fragment that was inserted into pBB-*egfp*, yielding p*nfeR1-40*::*egfpf*. Plasmid pBB*syn::egfp*, which expresses the reporter gene constitutively, was generated by cloning the *syn* promoter-*egfp* cassette (Torres-Quesada *et al*., 2013) as a *Kpn*I-*BamH*I fragment into pBBR1MCS-2.

To determine promoter activity, **fluorescence reporter assays** were employed using *S. meliloti* harboring p*nfeR1*::*egpf*, p*nfeR1-40*::*egpf*, pBB*syn::egfp* or pBB-*egfp* and grown in TY or MM medium to late exponential phase. Cultures of the reporter strains (100 µl) were transferred to 96-well microtiter plates followed by determination of OD_{600} and eGFP-mediated fluorescence with an Infinite M200 Pro microplate reader (Tecan). Fluorescence values were normalized to the culture OD_{600} and F/OD_{600} values obtained for strains harboring the control plasmids (pBB-*egfp* or pR-*egfp*) were subtracted from those obtained for the promoter-*egfp* fusions.

2.2.2.3. Construction of *S. meliloti hfq* mutant and derivative strains

In this work, an *S. meliloti* ***hfq*** **knockout** mutant was used. The *hfq* gene corresponds to ORF *SMc01048* (formerly denoted as *nrfA*) of the *S. meliloti* genome project (http://iant.toulouse.
inra.fr/bacteria/annotation/cgi/rhime.cgi) for reference strain Rm1021 (Galibert *et al.,* 2001). Deletion of *hfq* and verification of the mutation was performed using strain Rm1021 (Torres-Quesada *et al.,* 2010). For the use of an *hfq* deletion mutant in the framework of this study, pK18Δ*hfq* was mobilized to Sm2B3001 through triparental mating using pRK2013 as a helper plasmid. Obtained colonies were finally checked for the targeted deletion by colony PCR with the primers 5HfqMut/3HfqMut (Table 2.5, Appendix 2), and subsequent *Hind*III restriction of the PCR product (Figure 2.3).

Tagging of *hfq* with the FLAG epitope (Sigma-Aldrich) to generate the **Sm*hfq*^{FLAG}** strain was done in Sm2B3001, and correct insertion was verified by PCR and Western blot analysis by Torres-Quesada *et al.,* 2010.

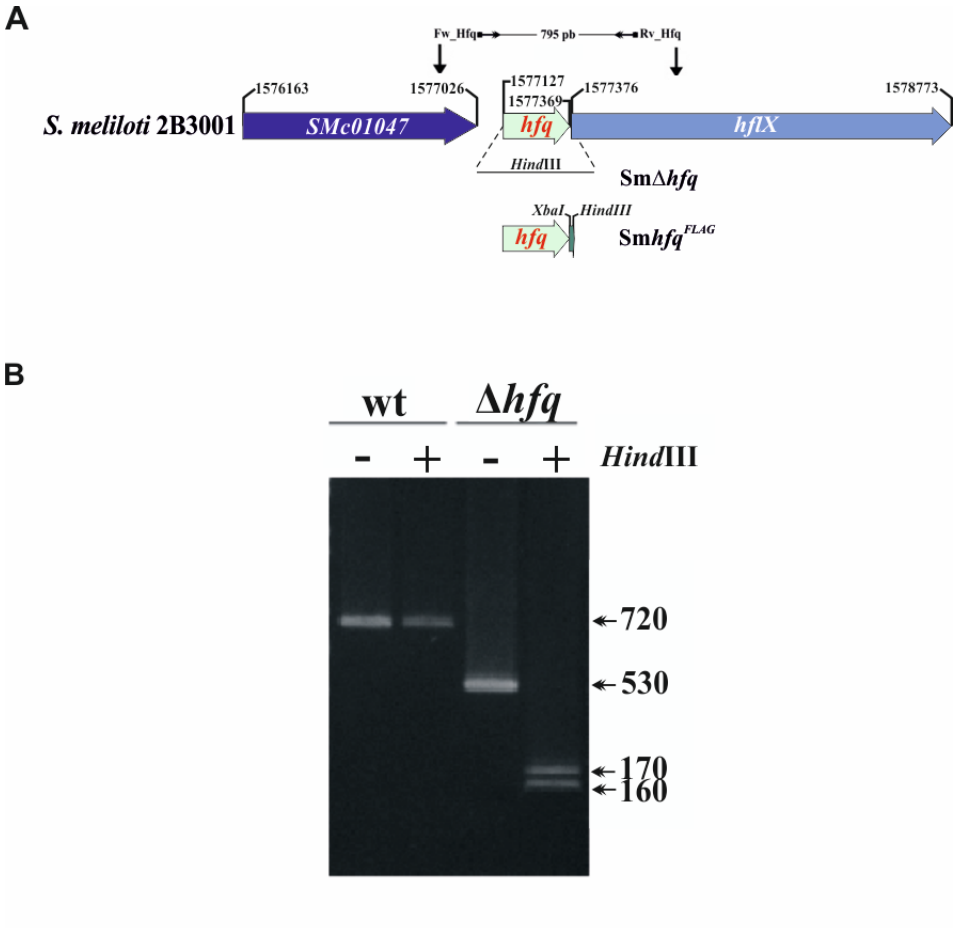

Figure 2.3. Construction and verification of the *S. meliloti hfq* deletion mutant. (A) Arrangement of the genomic *hfq* region in strain Sm2B3001. Numbering denotes the gene coordinates in the *S. meliloti* genome database (https://iant.toulouse.inra.fr/bacteria/annotation/cgi/rhime2011/rhime2011.cgi). The full-length *hfq* ORF was replaced by a *Hind*III site. The genetic map is not drawn to scale. **(B)** The deletion was first verified by PCR with primers 5HfqMut and 3HfqMut (Table 2.5, Appendix 2) that generated 720 bp and 530 bp amplification products in the parental and mutant strains, respectively. The mutant allele could be restricted with *Hind*III.

2.2.3. ANALYSIS OF THE NfeR1 PROMOTER

The NfeR1 promoter analysis was performed in collaboration with Dr. Coral del Val (University of Granada) and as described in https://en.wikipedia.org/wiki/%CE%91r14_RNA#Promoter_Analysis. To identify binding sites for known transcription factors we used the fasta sequences provided by **RegPredict** (Novichkov *et al.*, 2010) (http://regpredict.lbl.gov/regpredict/help.html),

and used those position weight matrices (PSWM) provided by **RegulonDB** (Gama-Castro *et al.*, 2011) (http://regulondb.ccg.unam.mx). We built PSWM for each transcription factor from the RegPredict sequences using the **Consensus/Patser program**, choosing the best final matrix for motif lengths between 14-30 bp with a threshold average E-value < 10^{-10} for each matrix that was established. Moreover, we searched for conserved unknown motifs using **MEME algorithm** (Bailey and Elkan, 1994) (http://meme-suite.org/) and used relaxed regular expressions (i.e. pattern matching) over all NfeR1 homologs promoters. The logo of the consensus sequence of this motif was generated at http://weblogo.berkeley.edu/logo.cgi. Promoter sequence alignments were generated with **ClustalW** implemented in **BioEdit** (http://www.mbio.ncsu.edu/
BioEdit/bioedit.html).

2.2.4. PHENOTYPIC ANALYSIS OF THE *nfeR1* DELETION MUTANT

2.2.4.1.Plant assays

For **nodulation assays**, *Medicago sativa* L. 'Aragón' seeds were surface sterilized, germinated and finally transferred to test tubes containing nitrogen-free nutrient solution (Rigaud and Puppo, 1975) as explained in Material and Methods. Seedlings were inoculated with 1 ml of a 10^6 bacterial suspension of either the wild-type Sm2B3001, mutant Sm∆*nfeR1* or complemented Sm∆*nfeR1*(pJBNfeR1) strain. The number of nodulated plants and the number of nodules per plant were recorded at defined days after inoculation. For **competition assays**, plants were inoculated with 1:1 mixtures of a GUS tagged wild-type strain and either of its mutant derivatives (Material and Methods).

2.2.4.2.Microscopy

Representative nodules from plants inoculated with either the wild-type or Sm∆*nfeR1* strain were excised and examined by microscopy. For assaying promoter activities, alfalfa plants were inoculated with Sm2B3001 carrying the reporter plasmids p*nfeR1*::*egpf* or p*nfeR1-40*::*egpf*. The empty vector pBB-*egfp* was used for measuring background fluorescence. GFP-mediated fluorescence during bacterial root hair colonization and infection thread formation was observed 6 and 9 dpi, respectively.

In this work, nodules emerging on plants roots were harvested 28 dpi for nodulation assays, and 14 dpi for nodule development assays. Bacteroids were isolated essentially as described (Material and Methods) and placed on 1% agarose pads.

2.2.5. ANALYSIS OF THE NfeR1 REGULON

2.2.5.1. Microarrays-based transcriptomics

Total RNA was obtained with the RNeasy Mini Kit (Qiagen) which was applied for Sm2B3001 and Sm∆*nfeR1* bacteria cultured to exponential growth phase (OD_{600} 0.5) in MM and following an osmotic upshift (four independent cultures per strain and growth condition) (Material and Methods). cDNA synthesis, Cy3- and Cy5 labeling, competitive hybridization of wild-type and mutant RNA to Sm14kOLI microarrays (ArrayExpress Accession No. A-MEXP-1760), image acquisition and data analysis were performed as described in Material and Methods. Signal intensities obtained for the wild-type strain grown exponentially in MM and subjected to the salt shock in the same medium were also compiled and compared in an *in silico* microarray experiment. Normalization and t-statistics were carried out using the EMMA 2.8.2 microarray data analysis software (Dondrup *et al.*, 2009). Genes with P-value ≤0.05 and M ≥1.0 or ≤-1.0 were included in the analysis. The M value represents the $\log_2$

ratio between both channels. Data are available at ArrayExpress (Accession No. E-MTAB-5588). Functional categories of the differentially expressed genes were established according to the *S. meliloti* Rm1021 genome sequence annotation (Galibert *et al*., 2001).

2.2.5.2. Quantitative proteomics

Strains Sm2B3001 and Sm*ΔnfeR1* were cultured to exponential growth phase (OD_{600} 0.5) in MM followed by an osmotic upshift in medium containing either $^{15}C_5H_9NO_4$ (isotope-labelled L-glutamic acid) or $^{14}C_5H_9NO_4$ (unlabeled L-glutamic acid), as the sole nitrogen source. Three independent cultures of each strain were processed as biological replicates. During cell harvesting, both strains were mixed and cell pellets were collected for proteomic comparison. The complete extraction procedure and analysis of the periplasmic fraction of proteins are described in Material and Methods.

2.2. RESULTS

2.3.1. CHARACTERIZATION OF THE αr14 sRNA FAMILY

αr14 is a family of bacterial small non-coding RNAs with representatives in a broad group of α-proteobacteria (del Val *et al*., 2012). NfeR1 was the first member of this family found in a *S. meliloti* Rm1021 locus located in the chromosome (Smr14C2) (del Val *et al*., 2007). In this work, we have performed *in silico* structural comparative analysis to further characterize this sRNA and assess its conservation across α-proteobacteria.

2.3.1.1. Conservation and distribution of the sRNA NfeR1 in α-proteobacteria

As described in the experimental setups of this chapter, the nucleotide sequence of the full-length *nfeR1* transcript was first used to query the Rfam database and next subjected to BLAST analysis using default parameters and including all available bacterial genomes. The regions with at least 78% similarity were collected to generate initial alignments to construct a covariance model (CM) for the sRNA. The identification of homologous sequences within the αr14 family revealed a conservation limited to species from the order Rhizobiales within the α-subgroup of proteobacteria (Figure 2.4, A). The CM was used to infer a consensus secondary structure for the family (Figure 2.4, B). The RNA family presented a typical sRNA arrangement in sub-structural domains with three main hairpin loops. This structure is supported by a variable degree of nucleotide covariance that was particularly high in the three stem-loops of the αr14 members. In the members of this family, the 3' domain consists of a GC rich hairpin followed by tails of uridine residues, thus matching the main structural feature of the Rho-independent terminators of transcription. The αr14 family mostly includes putative *trans*-encoded transcripts, which are expected to influence translation of target mRNAs through short base-pairing interactions that usually occlude the ribosome-binding site (RBS). Interestingly, the loop anti Shine-Dalgarno sequence "CUCCUCCC" was found to be conserved in all the three hairpin loops of these family members (Figure 2.4, B) (del Val *et al*., 2012; Reinkensmeier and Giegerich, 2015; Rivas *et al*., 2017). Besides, the terminal uridines of the Rho-independent terminators are all paired in NfeR1 and could not be easily available for Hfq binding, as described for other bacterial sRNAs. The mRNAs targets and Hfq-independence of the *S. meliloti* NfeR1 sRNA will be further analysed in this chapter.

The occurrence of the αr14 family in sequenced bacterial species of the order Rhizobiales was further assessed using the Infernal model (CM) generated in this work. The results of this comparative analysis are summarized in Figure 2.5, A. The CM identified αr14 family members distributed in the three taxonomic families of the

order that include the bacterial species most closely related to *S. meliloti* i.e., Rhizobiaceae, Brucellaceae and Phyllobacteriaceae. The manual inspection of the sequences found with the CM using Infernal allowed finding of 101 homologous sequences. The rhizobial species encoding the 36 closer homologs to NfeR1 were: *S. medicae* and *S. fredii*, two *R. leguminosarum* bv. trifolii strains (WSM2304 and WSM1325), two *R. etli* strains CFN42 and CIAT652, the reference *R. leguminosarum* bv. viciae 3841 strain, and the *Agrobacterium* species *A. vitis*, *A. tumefaciens*, *A. radiobacter* and *A. H13*. All these sequences showed significant Infernal E-values ($5.63*10^{-29}$-$8.16*10^{-18}$) and bit-scores. The rest of the sequences found with the model were identified with high E-values ($1.33*10^{-17}$-$8.79*10^{-3}$) but lower bit-scores and are encoded by *Brucella* species (*B. ovis*, *B. canis*, *B. abortus*, *B. microtis*, and several biobars of *B. melitensis*), *Ochrobactrum anthropic* and the *Phyllobacteriacea* species (*Mesorhizobium* species *M. loti*, *M. cicero* and *M.* BNC) (Figure 2.5, A). sRNA genes of the αr14 family exist in highly variable copy numbers in the individual genomes; many of them located on extrachromosomal replicons i.e., large accessory plasmids in Rhizobiaceae/Phyllobacteriaceae representatives (Figure 2.5, A). The αr14 RNA family showed then a complex distribution pattern in the Rhizobiales. Similarly to the situation of AbcR2/AbcR1 (Smr15C1/Smr15C2) (Chapter 1), genes arranged in tandem in the same *S. meliloti* Rm1021 intergenic region (IGR) encode Smr14C2 (NfeR1) and Smr14C3. The same results were obtained for *S. medicae* (Smedr14C1, Smedr14C2), *S. fredii* (Sfr14C1, Sfr14C2), *M. loti* (Mlr14C1, Mlr14C2) and *M. cicero* (Mcr14C1, Mcr14C2). However, in *O. anthropi* ATCC49188 and in species of the genera *Agrobacterium* and *Brucella*, the second chromosomal prediced copies of this gene are located in positions far from the first (i.e. it does not occur in such a syntenic context). A variable number of additional αr14 copies (up to six more in the genome of *R. leguminosarum* bv. trifolii WSM1325) were identified in the main chromosome and accessory plasmids of most of the bacterial species belonging to the Rhizobiaceae and Phyllobacteriaceae families (Figure 2.5, A). The αr14 family members are mostly encoded in IGRs with a few exceptions of genes predicted within annotated ORFs. However, these ORFs are frequently small, seemingly coding for hypothetical proteins

and/or absent from syntenic positions in bacterial genomes, thus representing probable mis-annotations as protein coding regions (https://en.wikipedia.org/wiki/%CE%91r14_RNA#Genomic_Context). In general, tandemly-arranged αr14 genes occur in complete or partial microsynteny with the flanking genes in genomes of Rhizobiaceae and Phyllobacteriaceae as do their homologs on the main chromosome of *O. anthropic* ATCC49188 and *Brucella* species. However, microsynteny is much more fragmented or even absent for many of the remaining chromosomal and plasmidic copies of the αr14. Altogether, these observations suggest that αr14 constitute a family of paralogous sRNA gene copies in the Rhizobiales probably originated from duplication events of their respective ancestral chromosomal genes over evolutionary time scales. Nonetheless, horizontal DNA transfer events could certainly contribute to the current distribution patterns of some αr14 gene copies, particularly of those occurring without signs of microsynteny in the accessory plasmids of plant-interacting bacteria.

The structural clustering generated from sRNA sequences in the αr14 family correlates with the phylogeny of the inferred order of classical taxonomic markers (i.e. 16S RNA), and groups them into four main branches (Figure 2.5, B), being five of the six putative αr14 loci in *S. meliloti* enclosed within the same branch. The proposed nomenclature for these sRNAs is the initial genus and species (adding the strain code if there were more than one), followed by the letter "r" (in allusion to RNA), the family identification number (14) and the genomic location (e.g. ReCIATr14C3: *Rhizobium etli* CIAT, RNA, family 14, chromosomal copy 3). The complete list of members of the αr14 family and their coordinates in the respective genomes can be found on the website https://en.wikipedia.org/wiki/%CE%91r14_RNA.

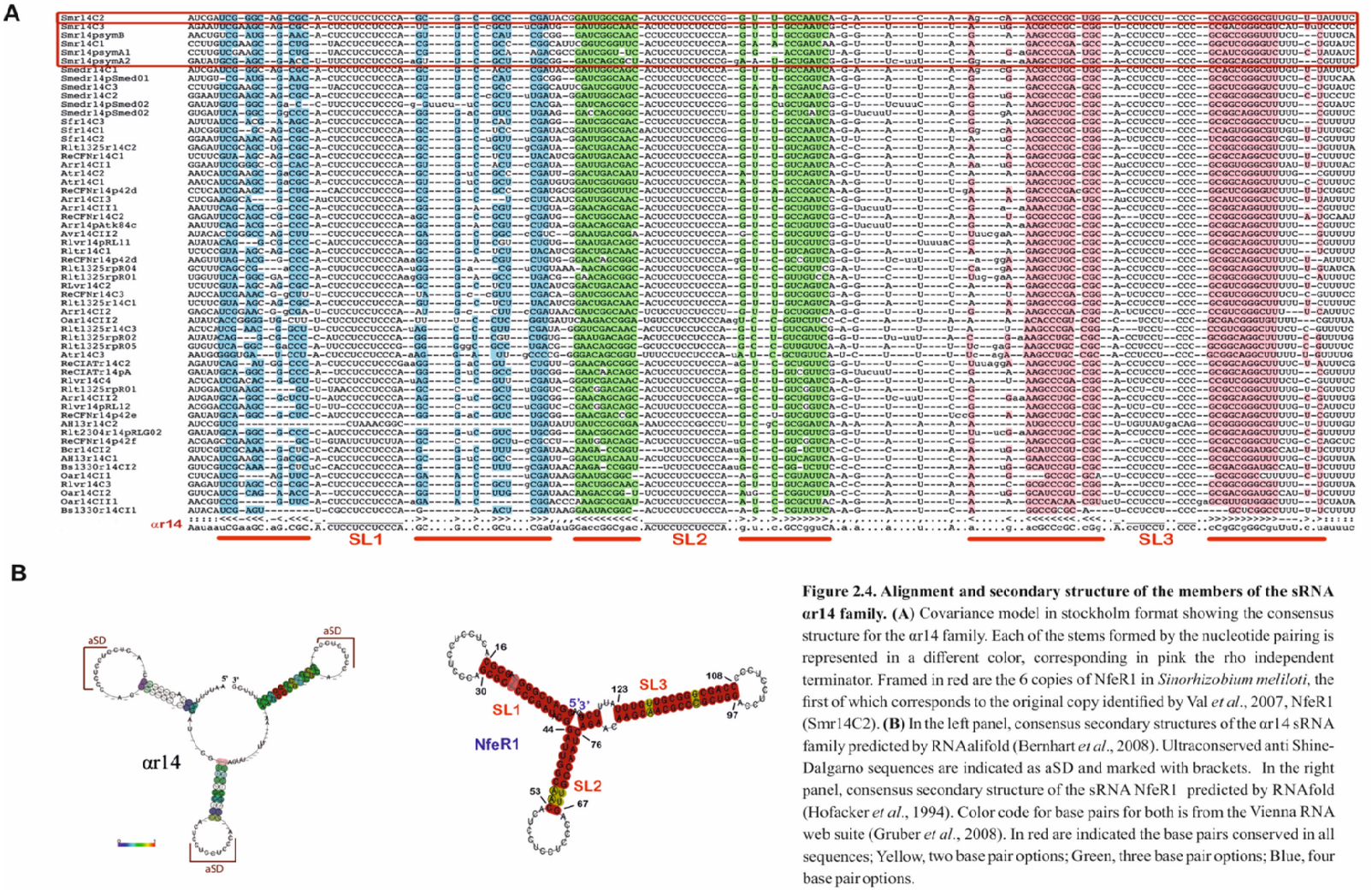

Figure 2.4. Alignment and secondary structure of the members of the sRNA *αr*14 family. (A) Covariance model in stockholm format showing the consensus structure for the αr14 family. Each of the stems formed by the nucleotide pairing is represented in a different color, corresponding in pink the rho independent terminator. Framed in red are the 6 copies of NfeR1 in *Sinorhizobium meliloti*, the first of which corresponds to the original copy identified by Val *et al*., 2007, NfeR1 (Smr14C2). **(B)** In the left panel, consensus secondary structures of the αr14 sRNA family predicted by RNAalifold (Bernhart *et al*., 2008). Ultraconserved anti Shine-Dalgarno sequences are indicated as aSD and marked with brackets. In the right panel, consensus secondary structure of the sRNA NfeR1 predicted by RNAfold (Hofacker *et al*., 1994). Color code for base pairs for both is from the Vienna RNA web suite (Gruber *et al*., 2008). In red are indicated the base pairs conserved in all sequences; Yellow, two base pair options; Green, three base pair options; Blue, four base pair options.

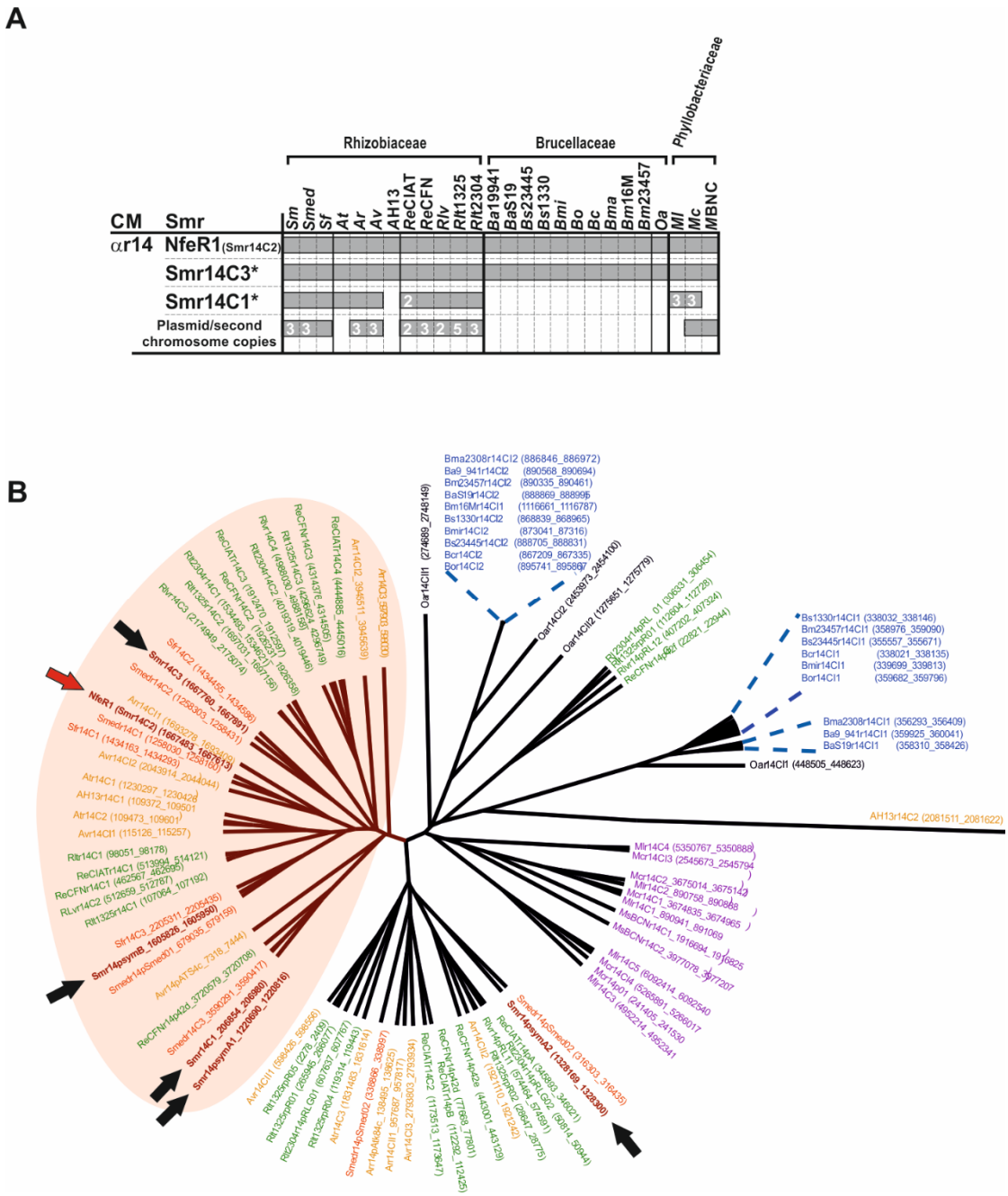

Figure 2.5. Structural clustering and occurrence of known and predicted components of the sRNAs αr14 family. (A) CMs (covarianze models) were generated along with the name of the query *S. meliloti* sRNA sequences listed on the left. The newly predicted chromosomal copies in *S. meliloti* are indicated with an asterisk. All bacterial species with representatives of the αr14 family are indicated on top of the panel grouped by taxonomic families i.e., Rhizobiaceae, Brucellaceae and Phyllobacteriaceae, as follows; *Sm*, *S. meliloti* Rm1021; *Smed*, *S. medicae* WSM419; *Sf*, *S. fredii* NGR234; *At*, *Agrobacterium tumefaciens* C58; *Ar*, *A. radiobacter* K84; *Av*, *A. vitis* S4; *A*H13, *A.* sp H13–3; *Re*CIAT, *Rhizobium etli* CIAT652; *Re*CFN, *R. etli* CFN42; *Rlv*, *R. leguminosarum* bv.

viceae 3841; *Rlt*1325, *R. leguminosarum* bv. trifolii WSM1325; *Rlt*2304, *R. leguminosarum* bv. trifolii WSM2304; *Ba*19941, *Brucella abortus* bv. One 9–941; *Ba*S19, *B. abortus* S19; *Bs*23445, *B. suis* ATCC23445; *Bs*1330, *B. suis* 1330; *Bmi*, *B. microti* CCM4915; *Bo*, *B. ovis* ATCC25840; *Bc*, *B. canis* ATCC 23365; *Bma*, *B. melitensis* bv. *abortus* 2308; *Bm*16M, *B. melitensis* bv. 1 16M; *Bm*23457, *B. melitensis* ATCC23457; *Oa*, *Ochrobactrum anthropi* ATCC49188; *Ml*, *Mesorhizobium loti* MAFF303099; *Mc*, *M. ciceri* bv. *biserrulae* WSM1271; *M*BNC, *M*. sp BNC1. Grey bars indicate distribution of each sRNA in these bacterial species. If more than one, the number of chromosomal and extrachromosomal copies of each sRNA gene in *S. meliloti* is also indicated. **(B)** Gene numbers are based on computational analysis using the program Infernal. Shadowed in pink are the closer homologs (i.e. with lower Infernal E-values and bit scores). The red arrow indicates the first of all putative αr14 loci identified in *S. meliloti*. The remaining five loci are indicated with black arrows.

2.3.1.2. Expression of αr14 representatives predicted in *S. meliloti*

The αr14 CM also identified up to five additional predicted copies of the query **NfeR1** (**Smr14C2**)-encoding gene in the *S. meliloti* Rm1021 genome, with a length between 115-125 nt: two of them chromosomally located (**Smr14C1** and **Smr14C3**), two in pSymA (**Smr14A1** and **Smr14A2**) and the remaining one in pSymB (**Smr14B**) (Figures 2.5 and 2.6). Ascending numbers were assigned to these genes according to their genomic location. Evidences of the expression of four out of the five predicted copies have been obtained in independent studies relying on Northern hybridization or RNaseq profiling of the transcriptome: Smr14C3, referred to as sra38 (Ulvé *et al.*, 2007), Sm7 (Valverde *et al.,* 2009) or SmelC398 (Schlüter *et al.*, 2010); Smr14A1, referred to as sma8 (Valverde *et al.,* 2009) or SmelA075 (Schlüter *et al.*, 2010); Smr14A2, referred to as SmelA099 (Schlüter *et al.*, 2010); Smr14B, referred to as SmelB161 (Schlüter *et al.*, 2010). Smr14C1 is the only predicted copy with no experimental evidences so far.

We further assessed the expression of all putative αr14 loci in *S. meliloti* Sm2B3001 strain by Northern blot probing. We started by aligning the gene sequences of all the sequences in order to design 25mer oligonucleotide probes that were highly specific for each transcript (Figure 2.6, A and Table 2.5 Appendix 2). All new putative αr14 loci predicted in the *S. meliloti* genome are encoded in IGRs, with the exception of Smr14B, which is encoded antisense to the *SMb20591* gene (Figure 2.6, B). These probes were labeled at the 5'-end with γ32-ATP, as explained in the section Material

and Methods and used to hybridize total RNA from the wild-type strain Sm2B3001 grown under a range of stress conditions, i.e. acidic and alkaline pH, osmotic upshift and oxidative stress, or in complete TY and minimal MM media (Figure 2.6, C). Strikingly, this set of Northern hybridizations revealed the 123 nt long NfeR1 transcript (coordinates 1,667,613 to 1,667,491 in the Rm1021 chromosome) as the only copy detected with this method, whereas no transcript for any of the other five genes was detected (Smr14C1, Smr14C3, Smr14A1, Smr14A2 and Smr14B) (Figure 2.6, C). Multiple nucleotide sequence alignments of the promoter regions of all the

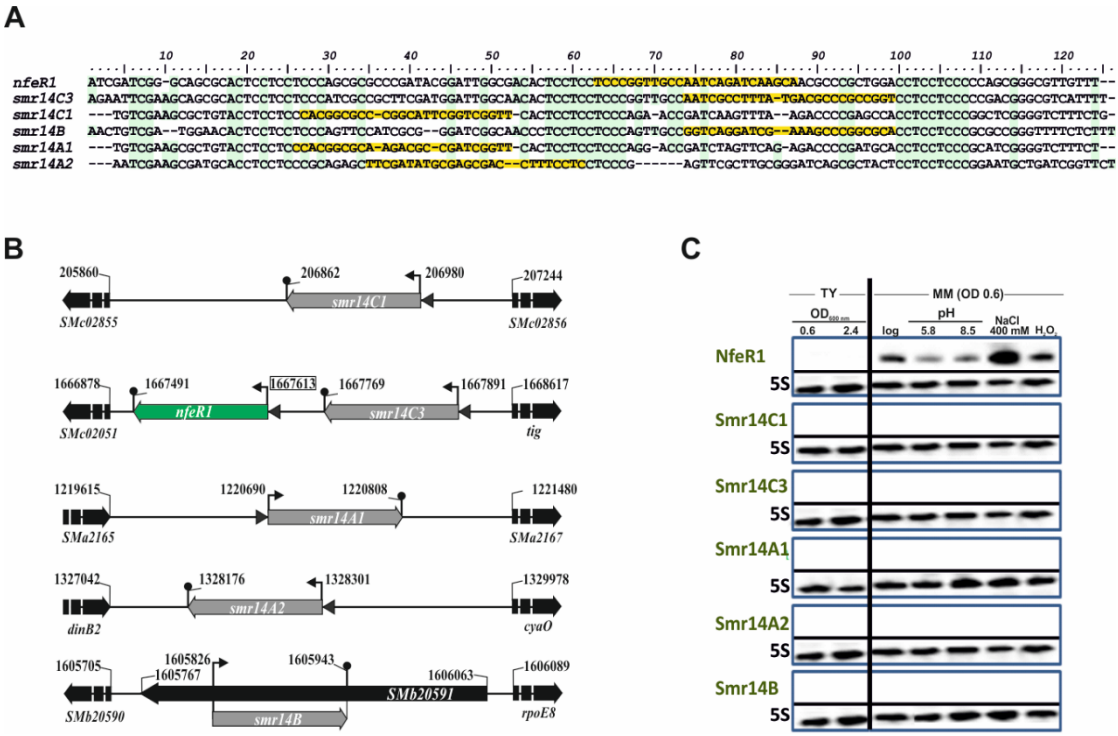

Figure 2.6. Analysis of the *smr14* sRNAs expression in *S. meliloti*. (A) Alignment of the gene sequences corresponding to all putative αr14 loci in *S. meliloti*. The probes used for Northern blot hybridization were oligonucleotides of sequences complementary to those labeled in yellow in each sRNA. The most conserved nucleotides are shadowed in green. **(B)** Maps of the genomic regions (not drawn to scale) of all the genes predicted by the αr14 CMs in *S. meliloti* Rm1021. Numbers denote coordinates of the genes in the genome. **(C)** Northern blot probing of total RNA from *S. meliloti* strain Sm2B3001 obtained at OD_{600} and in growth conditions indicated on top of the panels. 5S rRNA was probed as RNA loading control. Log, logarithmic growth phase (OD_{600} 0.4-0.5); TY, complex tryptone-yeast medium; MM, minimal medium.

genes encoding αr14 members in species of the Rhizobiales identified diverse conserved motifs that could contribute to the differential expression of these genes in specific biological conditions (https://en.wikipedia.org/wiki/%CE%91r14_RNA#Promoter_Analysis). Supporting this prediction, RNASeq of the *S. meliloti* sRNAs expressed in a number of stress conditions has rendered variable number of reads for the *S.meliloti* αr14-transcripts, possibly correlating with a diversity of accumulation profiles (Schlüter *et al*., 2010). Further analysis of the *nfeR1* promoter activity is presented below.

2.3.2. DIFFERENTIAL EXPRESSION OF THE sRNA NfeR1 IN *S. meliloti*

A common feature of bacterial regulatory sRNAs is their differential accumulation in response to specific external cues. Therefore, we first profiled NfeR1 expression by Northern analysis of total RNA from the wild-type strain Sm2B3001 as described above (Figure 2.6, C). Membranes were probed with the specific NfeR1 probe (PbNfeR1) targeting a central stretch of the NfeR1 transcript, which is variable in nucleotide sequence among the predicted homologs in the Sm2B3001 genome (Figure 2.6, A). Specificity of PbNfeR1 for detection of NfeR1 was verified by probing total RNA from the *nfeR1* deletion mutant (SmΔ*nfeR1*) and its derivative constitutively expressing NfeR1 from plasmid pSRK-NfeR1 (Figure 2.7, A).NfeR1 was reliably detected in bacteria growing exponentially in MM (OD_{600} 0.6) (Figure 2.7, B, upper panel). Oxidative stress did not alter accumulation of this sRNA whereas shifts in the extracellular pH resulted in a decrease of transcript levels. Of note, NfeR1 was strongly up-regulated upon a sudden increase in external osmolarity elicited by addition of NaCl (400 mM). In contrast, NfeR1 was not detected in exponential (OD_{600} 0.6) and stationary (OD_{600} 2.4) Sm2B3001 TY cultures. The transcript was particularly abundant in nodules, confirming previous analysis made with the reference *S. meliloti* strain Rm1021 (Figure 2.7, B, upper panel) (del Val *et al*., 2007,

Roux *et al*., 2014). In a second series of experiments, we analyzed the expression kinetics of NfeR1 during growth in unmodified TY and MM media and under moderate salinity (50 mM NaCl) (Fig. 2.7, B, middle and bottom panels). The strongest hybridization signals were detected in RNA from bacteria reaching late exponential growth phase (OD_{600} 0.8) in both media, whereas entering into stationary phase rendered NfeR1 undetectable. The presence of salt accelerated NfeR1 expression in both TY and MM cultures (i.e. the transcript was now detectable in early exponential TY cultures, OD_{600} 0.4) and resulted also in an increase of sRNA levels during exponential growth. Expression of NfeR1 in these culture conditions was also switched off during stationary growth in both media.

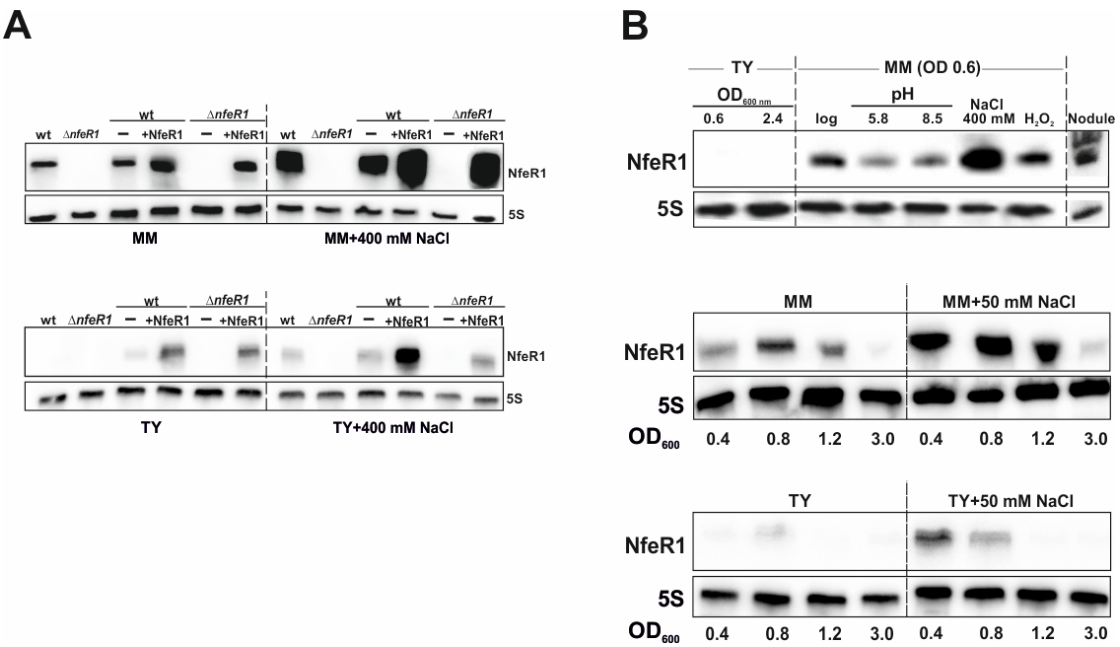

Figure 2.7. NfeR1 accumulation profile and specificity of the oligonucleotide probe for NfeR1 detection. (A) Northern blot probing of total RNA from the wild-type strain and its Δ*nfeR1* mutant derivative, both harbouring either, no vector, empty pSRK-C vector (-) or pSRK-NfeR1 overexpressing NfeR1 from a constitutive variant of the *lac* promoter (+NfeR1), as indicated on top of the panels. Bacteria were grown in unmodified MM or TY to logarithmic phase, or were subjected to an osmotic upshift (400 mM NaCl) in the same media. The 5S rRNA was probed as RNA loading control. **(B)** Northern blot probing of total RNA from *S. meliloti* strain Sm2B3001 obtained at the OD_{600} and in growth conditions indicated either on top or below the panels. 5S rRNA was probed as RNA loading control. Log, logarithmic phase (OD_{600} 0.4-0.5); TY, complex tryptone-yeast medium; MM, minimal medium.

We therefore conclude that *S. meliloti* NfeR1 is preferentially accumulated in response to salt as environmental abiotic stress factor, and in nodules as the plant-derived organs.

2.3.2.1. Transcriptional regulation of NfeR1 in free-living bacteria

Transcriptional regulation of the NfeR1 sRNA was further investigated by closer inspection of the *nfeR1* promoter region in order to better understand the differential expression (Figure 2.8, A). DNA sequence stretches were collected, which reached up to 100 nt upstream of the predicted transcription start site (TSS) of αr14 loci, and exhibited the closest homology to NfeR1. This group of sequences included NfeR1 homologs encoded by diverse *Sinorhizobium*, *Rhizobium*, *Agrobacterium* and *Brucella* species (del Val *et al*., 2012). The alignment of a subset of eight representative sequences evidenced a recognizable RpoD (σ^{70})-dependent -35/-10 promoter signature (CTTAGAC-N_{17}-CTATAT) widely conserved in the α-subgroup of proteobacteria (MacLellan *et al*., 2006). A MEME search for putative additional unknown motifs within the eight aligned promoter regions revealed a conserved 29-nt stretch between positions -47 and -77 (Figure 2.8, A). However, the consensus sequence of this motif did not match that of any binding site of known *S. meliloti* transcriptional regulators.

To determine whether this motif is required for *nfeR1* transcription, we fused the 100 bp DNA fragment upstream of the NfeR1 TSS and a 40 bp variant truncated at the 5'-end (i.e. lacking the MEME motif) to *egfp* (p*nfeR1*::*egfp* and p*nfeR1-40*::*egfp* constructs, respectively). These reporter transcriptional fusions were independently mobilized to the wild-type *S. meliloti* Sm2B3001 strain. As indicative of the promoter activity, fluorescence of the reporter strains was determined in culture conditions that favored accumulation of the NfeR1 transcript, i.e. late exponential growth in TY and MM, and osmotic upshift. However, fluorescence values from bacteria subjected to salt shock were not reliable, likely because of eGFP misfolding in this condition, and therefore, were not considered further in the quantitative analysis (not shown).

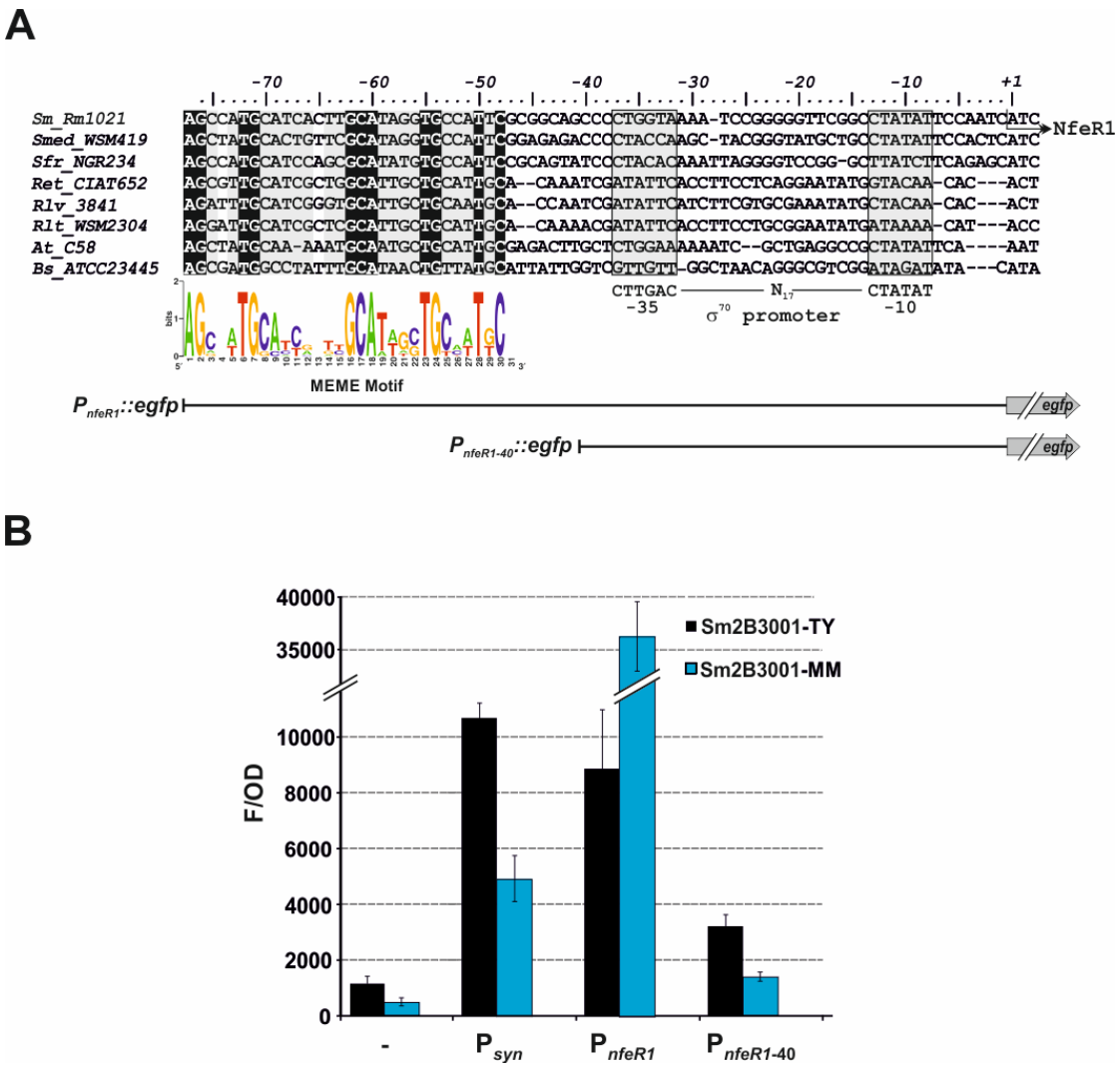

Figure 2.8. Transcriptional regulation of *nfeR1* in free-living bacteria. **(A)** Sequence alignment of the promoter regions of *S. meliloti* (*Sm*) NfeR1 homologs in type strains of *S. medicae* (*Smed*), *S. fredii* (*Sfr*), *R. etli* (*Ret*), *R. leguminosarum* bv. *viciae* (*Rlv*), *R. leguminosarum* bv. trifolii (*Rlt*), *A. tumefaciens* (*At*) and *B. suis* (*Bs*). Numbers on top stand for nucleotide positions with respect to the experimentally determined NfeR1 transcription start site (+1). Consensus sequences of the σ^{70} promoter signature and the conserved MEME motif (logo) are indicated at the bottom of the alignment. Promoter stretches transcriptionally fused to *egfp* are also indicated. **(B)** Fluorescence histogram of reporter *S. meliloti* Sm2B3001 strain transformed with plasmids pBB*syn::egfp*, p*nfeR1::egfp*, p*nfeR1-40::egfp* or a promoterless pBB-*egfp* (-). Values reported are means and standard deviations of nine fluorescence measurements normalized to the OD_{600} of the cultures in TY or MM media as indicated, i.e. three determinations of three independent cultures of each reporter strain.

Regarding the quantitative assay (Figure 2.8, B), fluorescence signals measured in the reporter Sm2B3001 strain expressing *egfp* from the full-length *nfeR1* promoter was

more than 3-fold higher in bacteria grown in MM than in TY cultures, thus paralleling the accumulation profile of NfeR1 determined by Northern hybridization (Fig. 2.7, B and C). Remarkably, in MM the P_{nfeR1}-derived fluorescence increased by ~7-fold with respect to that from the synthetic constitutive P_{syn}, commonly used for high-level gene expression in Gram-negative soil bacteria (Giacomini *et al*., 1994), whose activity was similar to P_{nfeR1} in TY broth. Trimming of P_{nfeR1} decreased promoter activity by ~3- and ~35-fold in TY and MM, respectively (Figure 2.8, B).

Altogether, these data indicate that the conserved MEME motif confers differential regulation and a particular strong transcriptional activity to the *nfeR1* promoter in *S. meliloti*.

2.3.2.2. Transcriptional regulation of NfeR1 during symbiosis

Previous Northern hybridization experiments anticipated high expression of *nfeR1* in mature nodules induced by the reference *S. meliloti* Rm1021 strain on alfalfa roots (del Val *et al*., 2007). Probing of RNA from Sm2B3001-elicited nodules also detected high levels of the NfeR1 transcript (Figure 2.7, B, upper panel).

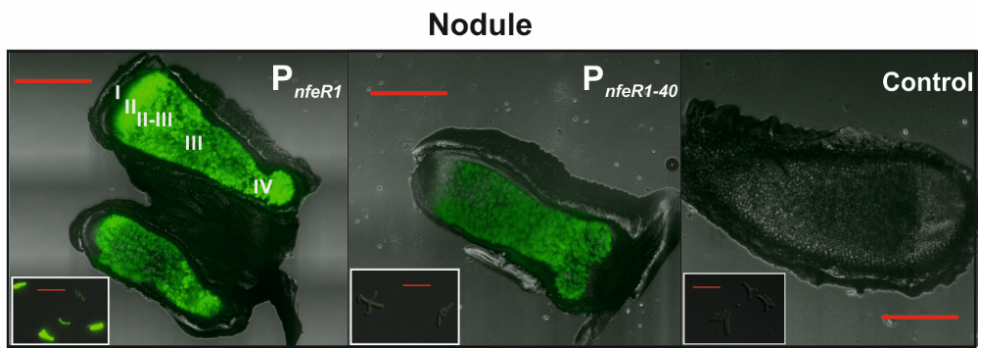

Figure 2.9. Transcriptional regulation of NfeR1 during symbiosis. Fluorescence microscopy images of mature alfalfa nodules harvested 28 dpi with bacteria expressing *egfp* from the full-length (P_{nfeR1}) or truncated ($P_{nfeR1\text{-}40}$) *nfeR1* promoter. The zones of typical indeterminate nodules are indicated. The insets show fluorescence of fully differentiated bacteroids. Control nodules were elicited by bacteria expressing a promoterless *egfp*. Red bars in nodules, 500 µm; red bars in bacteroids, 10 µm.

Therefore, we used P_{nfeR1}::*egfp* as reporter fusion to track NfeR1 transcription during the symbiotic interaction of Sm2B3001 with alfalfa plants grown hydroponically in nitrogen-free mineral solution (Figure 2.9). Fluorescence microscopy of longitudinal sections of 28 days-old root nodules evidenced high fluorescence signals derived from invading reporter bacteria in all the histologically defined zones of indeterminate nodule tissues that host symbiotic rhizobia, i.e. invasion zone II, interzone II-III, nitrogen-fixation zone III, and senescence zone IV (Figure 2.9, left image). Fluorescence was more intense in undifferentiated bacteria within zone II, but isolated fully differentiated nitrogen-fixing bacteroids occupying zone III also displayed strong transcription of *egfp* from P_{nfeR1} (inset). The activity of the shorter version of the promoter lacking the conserved MEME motif ($P_{nfeR1\text{-}40}$) was markedly lower than that of the full-length promoter throughout the nodule, particularly in the invasion zone II and in bacteroids (Figure 2.9, central image). Background fluorescence in nodules was set with bacteria carrying the promoterless control plasmid pBB-*egfp* (Figure 2.9, right image).

This fluorescence profile revealed very active expression of *nfeR1* during symbiosis, which is largely conferred by the conserved MEME motif within its promoter region.

2.3.3. PHENOTYPIC ANALYSIS OF THE *S. meliloti* Sm∆*nfeR1* MUTANT

We next investigated the phenotypes associated with the NfeR1 loss-of-function in free-living and symbiotic bacteria.

2.3.3.1. Phenotype of free-living bacteria

Growth kinetics of the wild-type Sm2B3001 strain and its deletion mutant derivative Sm∆*nfeR1* were similar in TY broth and MM (Figure 2.10, A). Given that NfeR1 accumulates specifically under saline stress, we compared the growth curves of both

strains in MM supplemented with 50 mM NaCl (Figure 2.10, B). Growth of the mutant was slightly impaired in MM containing 50 mM NaCl. Subsequent to this work, growth curves of both strains were recorded using MM supplemented with 250 mM instead of 50 mM NaCl, and plate counting of viable bacteria was performed in early stationary phase. This experiment confirmed that lack of NfeR1 compromised tolerance to salinity (Robledo *et al.*, 2017). In another series of experiments, 10-fold serial dilutions of cells (10^{-3}-10^{-6} cells/ml) were spotted onto MM agar supplemented with increasing concentrations of NaCl (0-250 mM) (Figure 2.10, C). Sm2B3001 and Sm∆*nfeR1* colonies were similar and addition of 100 mM NaCl to plates negatively influenced growth of both strains in a similar manner. Higher salt concentrations progressively impaired growth of the wild-type strain, which nonetheless still proliferated at 250 mM NaCl. However, sensitivity of Sm∆*nfeR1* to high salinity was markedly higher, i.e. growth of the mutant in 200 mM NaCl was barely detectable. The use of sucrose (700 mM) in these assays as alternative known elicitor of osmotic stress did not reveal differences in growth between the parental and mutant strain (data not shown), suggesting that osmostress sensitivity of the mutant is a salt-specific effect.

Together, these results show that NfeR1 is required for adaptation of free-living *S. meliloti* cells to osmotic stress elicited by high salt concentrations.

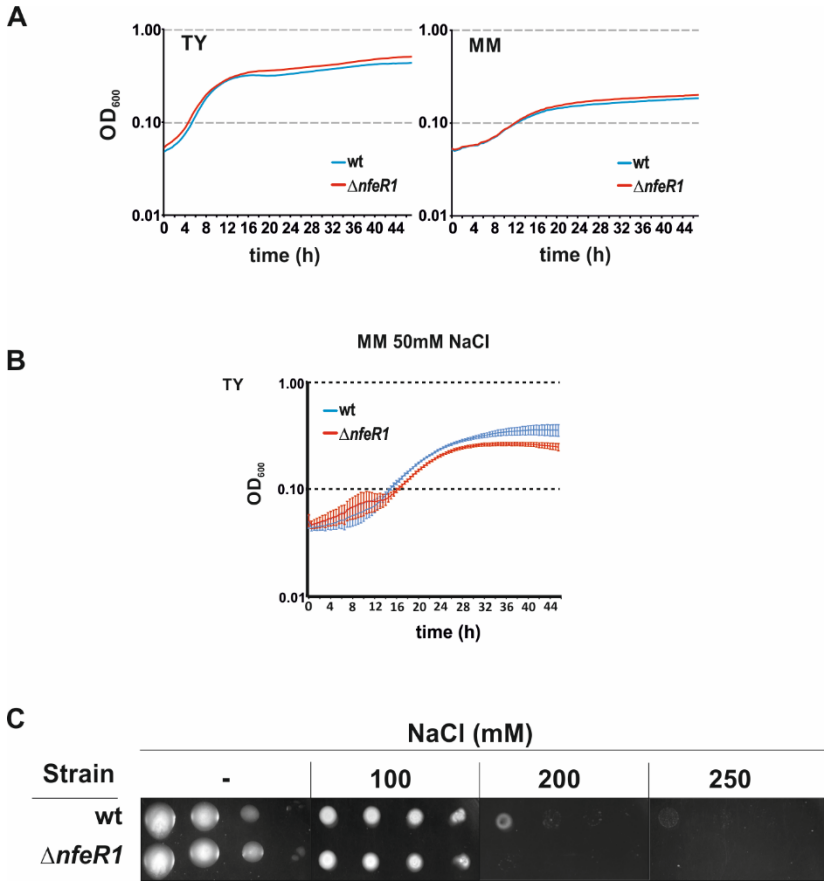

Figure 2.10. Free-living phenotype of the Sm*ΔnfeR1* mutant. (A) Growth curves of the wild-type and mutant strains in TY broth and MM. **(B)** Growth of wild-type and mutant strains in MM supplemented with 50 mM NaCl. Plotted values are means and standard deviations of two determinations in two independent cultures of both strains. The experiment was repeated twice (i.e. two series of two independent cultures) with similar results. **(C)** Growth of wild-type and mutant strains in unmodified MM agar (-) or in the presence of 100, 200 or 250 mM NaCl in the same medium. Series of drops (10 µl) containing approximately 10^6, 10^5, 10^4 and 10^3 bacteria were spotted. Pictures were taken after 4 days incubation at 30 °C. Representative colonies from two independent experiments are shown.

2.3.3.2. Phenotype of symbiotic bacteria

Since NfeR1 is highly expressed during nodulation, we also assessed the overall symbiotic performance of the Sm*ΔnfeR1* mutant (Figure 2.11). First, we performed competition experiments in which sets of 24 alfalfa (*M. sativa*) plants individually grown hydroponically in nitrogen-free mineral solution were co-inoculated with

mixtures containing equivalent numbers of two bacterial strains (i.e. 10^6 bacteria/ml each): a GUS-tagged wild-type Sm2B3001 strain (marker strain), the non-tagged Sm2B3001 (wild-type control), or Sm∆*nfeR1* bacteria. Nodule occupancy of each strain was then inferred from the number of blue (GUS-tagged bacteria) and white (tested strain) nodules formed on roots 28 dpi with each combination of strains (Figure 2.11, A). As expected, the wild-type Sm2B3001 strain occupied ~50% of nodules when co-inoculated with its GUS-tagged derivative. In contrast, the Sm∆*nfeR1* mutant scarcely reached 30% of nodule occupancy in the presence of the marker strain. This phenotype was largely, but not fully, complemented with pJB-NfeR1 (strain Sm*nfeR1**), which was most likely the consequence of a negative effect in wild-type competitiveness of the empty pJB-T plasmid itself (Figure 2.11, A). These results revealed that NfeR1 positively influences competitiveness of *S. meliloti* for nodulation of alfalfa roots.

In a second series of experiments, sets of 24 alfalfa plants grown as above mentioned were independently inoculated with the wild-type or Sm∆*nfeR1* strain and the respective nodulation kinetics were determined by recording the percentage of nodulated plants at daily intervals after inoculation (Figure 2.11, B).This assay revealed a marked delay of Sm∆*nfeR1* in plant nodulation. Plasmid pJB-NfeR1 partially restored the wild-type phenotype. At the end of the assay (28 dpi), the mutant hardly nodulated 65% of the inoculated plants compared to the 85% and 95% nodulated by the complemented Sm*nfeR1** and wild-type strains, respectively. These findings indicate that the *nfeR1* deletion compromises nodule formation efficiency of *S. meliloti* on alfalfa roots. Motility in the plant rhizosphere or adhesion to and colonization of plant roots have been reported to contribute to nodule formation efficiency of rhizobia on legume roots (Ames and Bergman, 1981; Gulash *et al*., 1984; Armitage *et al*., 1992).

As a follow-up experiment to this work, the absorption, colonization and motility phenotypes of the Sm∆*nfeR1* mutant, as well as the onset of nodule organogenesis have been also examined (Robledo *et al*., 2017). These experiments revealed that the

wild-type and mutant strains were equally proficient in attaching to and proliferating on the root surface, the swarming and swimming competence of Sm∆*nfeR1* did not present significant differences in the expansion areas with respect to the wild-type strain, and that the mutant is able to trigger early nodulation signaling (Robledo *et al*., 2017).

A closer inspection of nodule appearance throughout the symbiotic tests also revealed a delay in the kinetics of emergence of typical nitrogen-fixing nodules (Fix^+), i.e. of elongated morphology and marked pink color due to plant leghemoglobin expression, in plants inoculated with the NfeR1 mutant, compared to those treated with the wild-type or Sm*nfeR1** strains (Figure 2.11, C, left plot). At the end of the assays, the percentages of plants that developed at least one mature Fix^+-like nodule were ~85%, ~65%, ~30%, when they were inoculated with the wild-type, complemented or mutant strain, respectively. In the harvested plants, the average number of Fix^+ nodules elicited by the mutant per individual plant was scarcely 1.5 in contrast to ~2.5 induced by the wild-type or Sm*nfeR1** strains, respectively (Figure 2.11, C, right plot). Consistent with these observations, Sm∆*nfeR1*-inoculated plants exhibited a shoot length significantly shorter (average of 5-6 cm) than those treated with the wild-type or complemented strains (8-9 cm), which are symptoms of nitrogen deficiency (Figure 2.11, D).

Nodules elicited by the Sm∆*nfeR1* mutant were deformed and mostly smaller than the typical indeterminate nodules, which were predominant in plants inoculated with the wild-type strain (Figure 2.12, A and B). Nonetheless, microscopy of longitudinal sections of 14 days-old wild-type and mutant nodules revealed that both were similarly infected, with plant cells evenly invaded by differentiated bacteroids (Figure 2.12, C and D). However, no defined infection zone II or interzone II-III were manifest in the mutant nodules, where all endosymbiotic bacteria were likely confined to the nitrogen-fixation zone III. Further analysis of the development of the nodule by polarized light microscopy has shown the homogeneous crystallinity of plant cell walls in wild-type nodules, except at the actively growing meristem. In contrast, plant

cells in mutant nodules did not evidence noncrystaline regions, which further suggest a developmental arrest (Robledo *et al*., 2017).

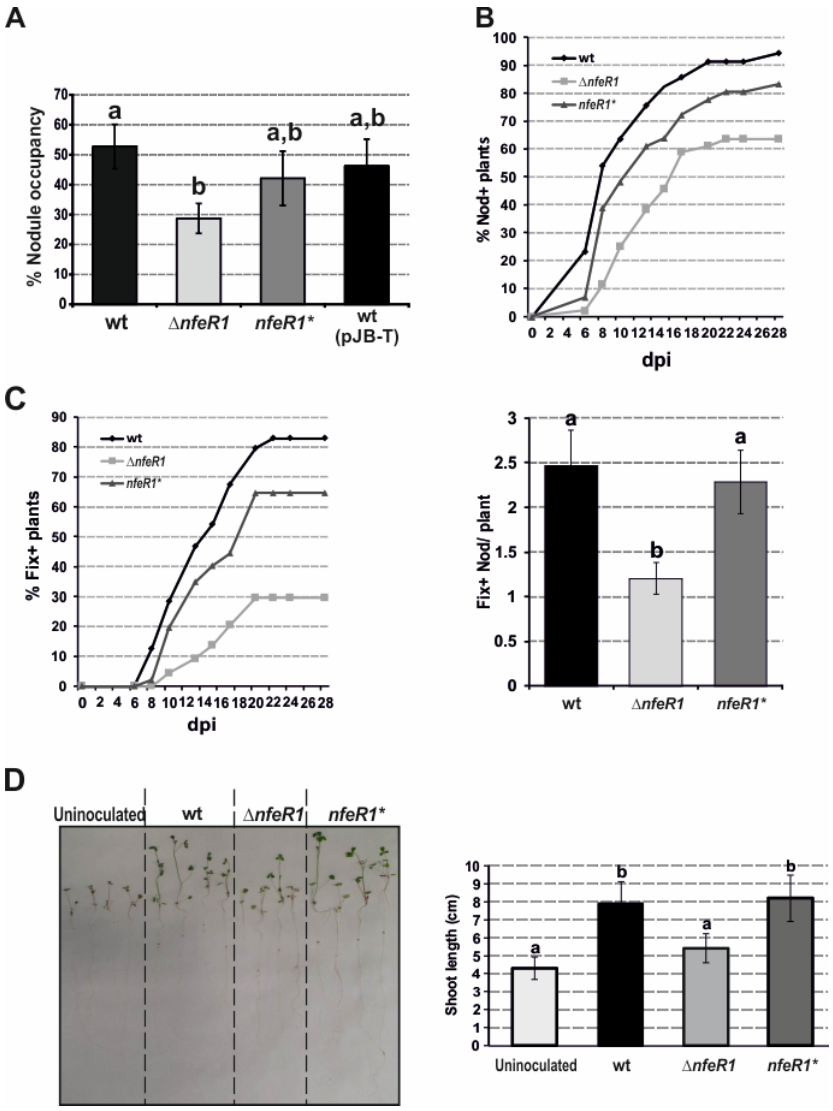

Figure 2.11. Symbiotic phenotype of the SmΔ*nfeR1* mutant. (A) Nodule competition assays. Data represent the percentage of white nodules occupied by the wild-type strain Sm2B3001, its *nfeR1* deletion mutant derivative (SmΔ*nfeR1*), the complemented strain (Sm*nfeR1**) or the wild-type strain transformed with the empty complementing vector pJB-T, 30 days after plants inoculation with 1:1 mixtures of each strain with Sm2B3001(pGUS3). Values are means and SD of three independent experiments. Letters indicate significant differences at $P<0.05$. **(B)** Nodulation kinetics of alfalfa plants inoculated with the wild-type, SmΔ*nfeR1* or Sm*nfeR1** strain. Values plotted represent the percentage of inoculated plants nodulated by each strain at different days post inoculation (dpi). A representative example of three independent experiments is shown. **(C)** Left plot, kinetics of appearance of Fix^+ (i.e. with evident pink color) nodules. Values represent the percentage of pink nodules induced on alfalfa roots by the aforementioned strains. Shown is a representative example of three

independent experiments. Right plot, number of Fix^+ nodules per inoculated plant induced by each strain at the end of the experiment (28 dpi). Values are means and standard deviation of three independent experiments. The statistical significance was set at $P<0.05$ and indicated by letters. **(D)** Whole-plant phenotype. Left, image of representative subsets of mock treated or inoculated alfalfa plants at the end of the assay (28 dpi). Bacterial strains were as in C. Right, shoot length of the plants. Plotted are mean and standard deviation of measurements in a total of 48 plants in two independent experiments. Statistical significance at $P<0.05$ is indicated by letters.

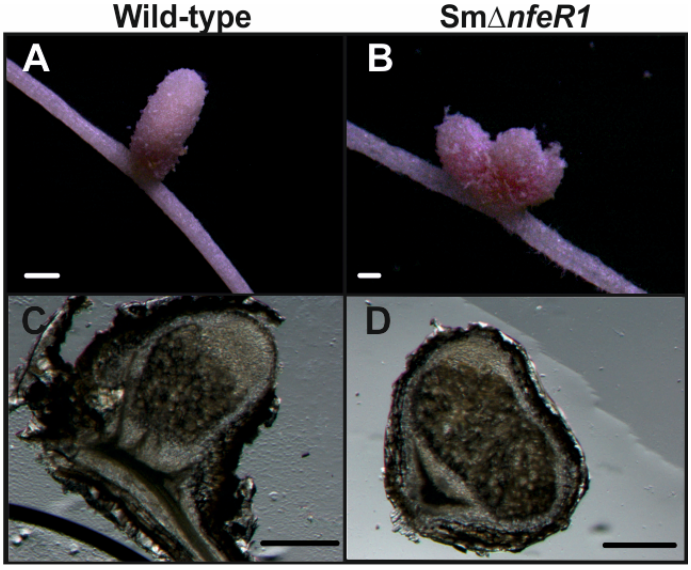

Figure 2.12. Endosymbiotic phenotype of the Sm*ΔnfeR1* mutant. (**A, B**) Whole wild-type and mutant nodules (14 dpi) observed under the binocular (bars, 500 µm). (**C, D**) Longitudinal sections of 14-days old nodules observed with microscopy (bars, 500 µm).

Collectively, the symbiotic phenotypes described above indicate that NfeR1 contributes to symbiotic nitrogen-fixation and nodule formation efficiency.

2.3.4. CHARACTERIZATION OF THE NfeR1 REGULON

2.3.4.1. Impact of *nfeR1* deletion on the *S. meliloti* transcriptome

For the characterization of the NfeR1 regulon, we first explored the influence of NfeR1 on the physiology of *S. meliloti* bacteria, probing the transcriptomes of the wild-type Sm2B3001 strain and its Sm*ΔnfeR1* mutant in microarray experiments (Figure 2.13). Total RNA was extracted from exponentially growing bacteria in MM

and upon 1 h shock with 400 mM NaCl in the same medium (MM+NaCl). In non-stressed bacteria (MM), only 32 NfeR1-dependent mRNAs were identified, i.e. displaying at least 2-fold changes in their abundance between the two strains, with 12 upregulated and 20 downregulated in Sm∆*nfeR1* (Table S1 in supplementary data of Robledo *et al*., 2017. http://onlinelibrary.wiley.com/doi/10.1111/1462-2920.13757/abstract). Upon the osmotic upshift (MM+NaCl), we found 86 genes exhibiting differential expression in the Sm∆*nfeR1* mutant with respect to the wild-type strain (42 induced and 44 repressed; Table S2 in supplementary data of Robledo *et al*., 2017. http://onlinelibrary.wiley.com/doi/10.1111/1462-2920.13757/abstract), with 11 common to those identified in unmodified MM, i.e. the expression of this set of common genes likely depends on NfeR1 regardless the growth condition (Venn diagram). These findings anticipate a broader impact of NfeR1 activity in stressed than in non-stressed bacteria. A comparison of the transcriptome of the wild-type strain, grown exponentially in MM until the salt shock treatment, identified 704 mRNAs (salt regulon) whose abundance was significantly altered (i.e. at least 2-fold) in response to the sudden increase of external NaCl concentration (Table S3 in supplementary data of Robledo *et al*., 2017. http://onlinelibrary.wiley.com/doi/10.1111/1462-2920.13757/abstract). Interestingly, this catalog of genes recapitulated 168 of those reported previously as differentially expressed upon an osmotic upshift in the closely related strain Rm1021 (Domínguez-Ferreras *et al*., 2006), including genes specifying translation, transcription, metabolic, membrane and stress-related functions. A comparison of all these three data sets (MM, MM+NaCl and salt regulon) identified 75 genes whose expression was influenced by NfeR1 upon an osmotic upshift but not in unstressed bacteria (Figure 2.13; Venn diagram, red outline).

According to the *S. meliloti* Rm1021 genome sequence annotation, 52 out of these 75 genes encode proteins with predictable function and, based on their NfeR1-dependent expression, can be clustered into two major categories (Figure 2.13; heatmap). **Cluster I** includes genes expressed at higher levels in the NfeR1 mutant than in the parental strain when both were subjected to the osmotic upshift. A large subset of

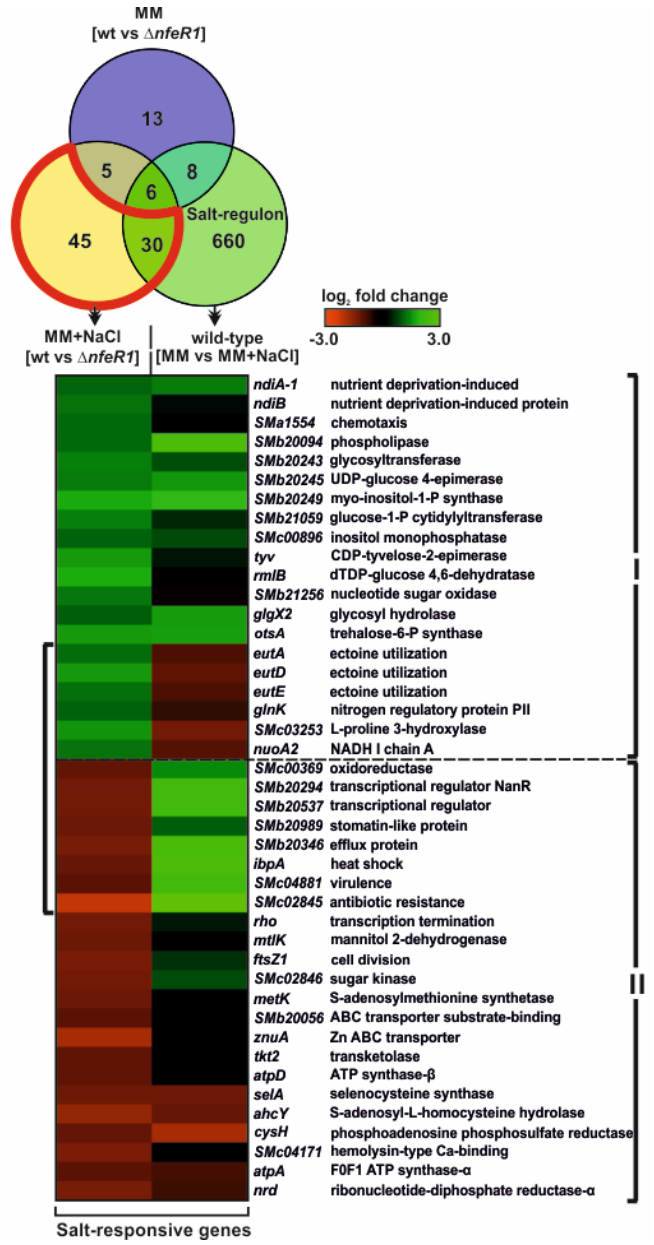

Figure 2.13. The NfeR1-dependent transcriptome. Venn-diagram, number of differentially expressed genes in the Sm*ΔnfeR1* mutant with respect to the wild-type strain [wt vs *ΔnfeR1*] when both were grown in unmodified MM or after 1 h shock with 400 mM NaCl (MM+NaCl), and genes with altered expression levels in the wild-type strain upon the osmotic upshift (MM vs MM+NaCl; salt-regulon). The red outline indicates the genes sensitive to both the *nfeR1* deletion and the osmotic upshift. Heatmap, expression of salt-responsive genes (red outline) with predicted function. Plotted are M values of changes (log_2 fold) in mRNA abundance in the following comparisons: wild-type strain vs Sm*ΔnfeR1* mutant when both were subjected to the osmotic upshift and non-stressed vs stressed wild-type bacteria. In the color scale, red represents downregulation and green upregulation with respect to the reference (wild-type or non-stressed bacteria) in each comparison. The bracket denotes the subset of genes with markedly opposite expression in both conditions. Name and putative function of each gene are indicated to the right.

these genes are related to small molecule metabolism and salt itself had none or subtle positive effects on their expression in the wild-type strain, thus suggesting that NfeR1 had only a discrete influence on the accumulation of the corresponding mRNAs. Of note, cluster I also included a number of genes specifying pathways that were downregulated in the wild-type strain by salinity stress, such as *eutAED*, coding for proteins for the catabolism of the osmoprotectant ectoine, which has been shown to

accumulate in response to osmotic stress in some bacteria (Jebbar *et al*., 2005; Schulz *et al*., 2017). **Cluster II** gathers the genes positively influenced by NfeR1 (i.e. downregulated in Sm∆*nfeR1* with respect to the parental strain) in response to the saline shock. The expression of more than half of these genes, mostly coding for metabolic proteins, was barely affected by salt in the wild-type genetic background. However, it is noteworthy that a subset of genes within this cluster II encode functions activated to counteract the osmotic shock in the wild-type strain, which therefore are presumably compromised in the NfeR1 mutant, e.g. transcriptional regulation (*SMb20294*, *SMb20537*), response to heat shock (*ibpA*), virulence (*SMc04881*), antibiotic resistance (*SMc02845*) and membrane trafficking (*SMb20989* and *SMb20346*).

2.3.4.2. Computational comparative prediction of NfeR1 mRNA targets

To predict targets, interaction domains and consequently the regulatory networks of this bacterial sRNA, we used the CopraRNA algorithm (comparative prediction algorithm for small RNA targets) (Wright *et al*., 2013) (Material and Methods), as we realized with the AbcR2 sRNA (Chapter 1).

The nucleotide sequence of NfeR1 and its closest homologs in *S. medicae*, *S. fredii*, *A. tumefaciens*, *R. leguminosarum* bv. viceae, *R. leguminosarum* bv. trifolii and *R. etli* were used as queries in these predictions. NfeR1 is expected to interact at the ribosome binding site (RBS) of its mRNA targets through the ultraconserved anti Shine-Dalgarno (aSD) sequence motifs that remain single-stranded within each stem loop (SLa-c) of the αr14 family members (Figure 2.4. B, left panel) (del Val *et al*., 2012; Reinkensmeier and Giegerich, 2015; Rivas *et al*., 2017), thereby downregulating translation. Therefore, sequence stretches extending 200 nt upstream and 100 nt downstream of the annotated start codons in each gene were set as the potential target regions (Figure 2.14). CopraRNA returned a large list of putative

targets for NfeR1 and its homologs, but revealed a functional enrichment for mRNAs derived from ABC transporter genes (Figure 2.14).

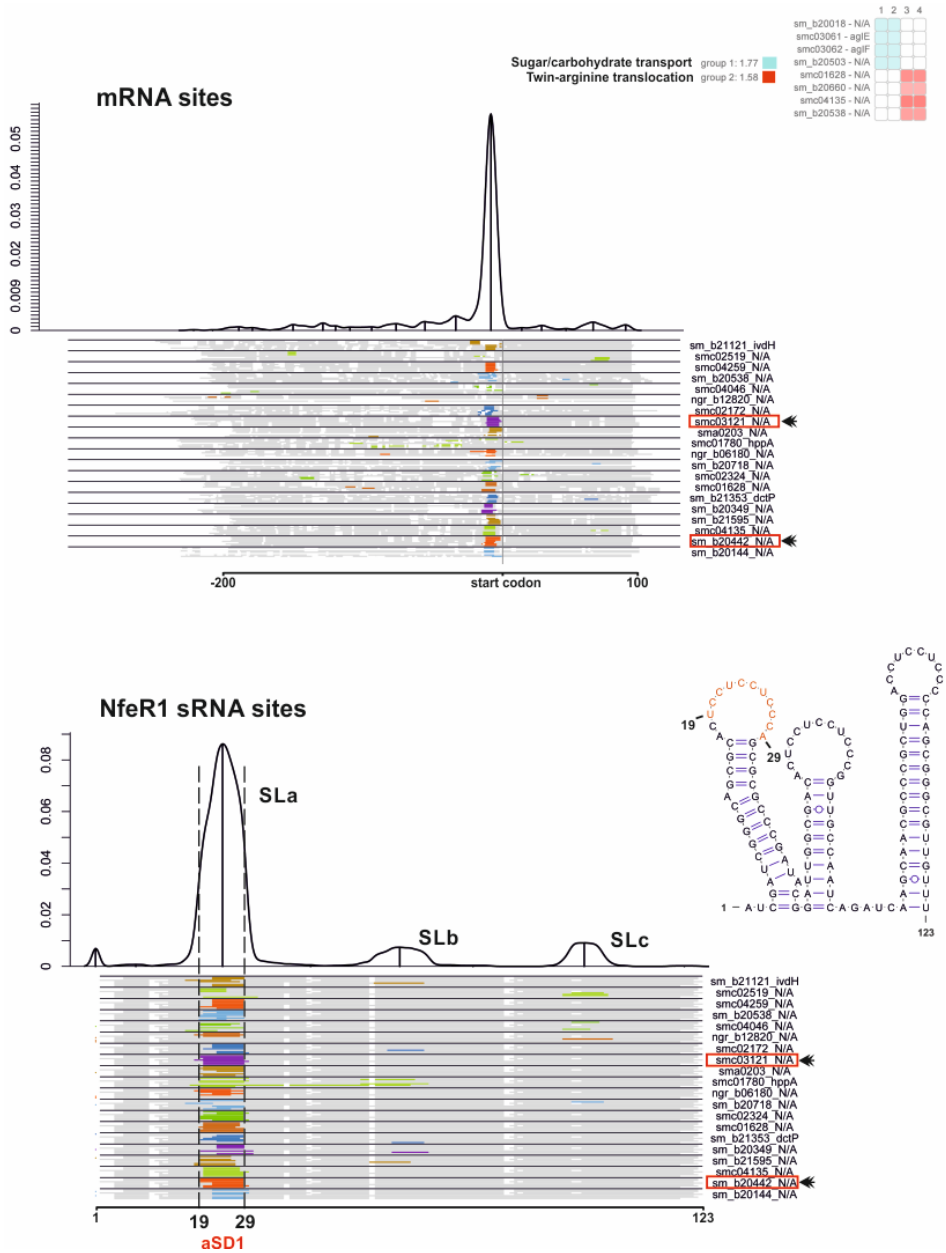

Figure 2.14. CopraRNA prediction of NfeR1 mRNA targets. Depiction of the putative interaction domains in the predicted mRNA targets of NfeR1 (mRNA sites) and in the sRNA (NfeR1 sRNA sites). The plots combine all predictions with a *P* value ≤ 0.01 in all included homologs. Local maxima indicate distinct interaction domains and are marked with upright lines. The schematic alignment of homologous sRNAs and targets at the bottom show the predicted interaction domains. The aligned regions are displayed in gray, gaps in white, and predicted interaction regions in color (color differences are for contrast only). The locus tag and gene name (if available) of a representative cluster member are given on the right. Framed in red are two mRNAs experimentally verified after this work (Robledo *et al*., 2017).

2.3.4.3. The NfeR1-dependent periplasmic proteome

Transcriptome analysis did not identify any of the targets predicted by CopraRNA. Therefore, we performed the comparative and quantitative proteomic approach for the periplasmic fraction of Sm2B3001 and Sm*ΔnfeR1* bacteria grown in MM and subjected to osmotic upshift. The experimental strategy followed in this work is based on the technique known as SILAC (Stable Isotopic Labeling of Amino acids in Cell Culture) (Ong, *et al*., 2002), described in more detail in Material and Methods.

The peptides of our samples were quantified and standardized in the two experimental conditions tested. The obtained M = 0 centered distribution indicatess that only a small fraction of the periplasmic proteome responds to NfeR1 activity (Figure 2.15).

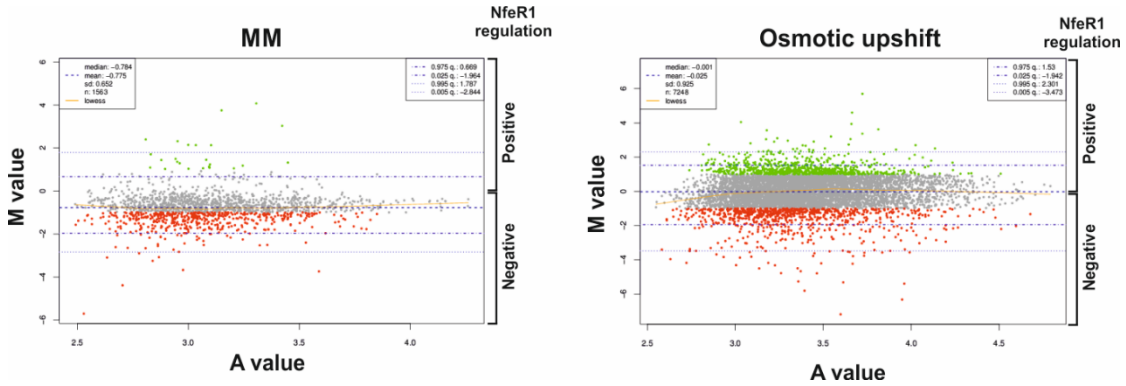

Figure 2.15. Quality control analysis of the proteomic approach. Representations of the M-A values of each peptide in MM (left panel), and osmotic upshift (right panel). In gray color are represented the peptides with values $-1 < M < 1$. The repressed ($M \geq 1$) or accumulated ($M \leq -1$) peptides in the mutant strain are indicated in green and red, respectively.

The results evidenced that of the total proteins identified, 33 ABC transporters were upregulated in the Sm2B3001Δ*nfeR1* mutant, and therefore represent putative NfeR1 mRNA targets. Of those, only two were predicted by Copra (*supA* and *SMc02514*), although with lower significance values compared to other putative mRNA targets. Table 2.1 and 2.2 show the periplasmic proteins whose differential abundance in the mutant strain has been statistically validated in the condition tested during growth. Table 2.1 highlights the AmtB ammonium transporter in blue, which is discussed

below. In Table 2.2 mRNA targets predicted by CopraRNA are coloured red.

Table 2.1. Periplasmic transport proteins with altered abundance in the NfeR1 mutant during growth in MM.

ID	Protein product	M[a]
Proteins negatively regulated by NfeR1		
SMc02737	ChoX choline ABC transporter	-2,66
SMa0252	TRAP-type periplasmic solute-binding protein	-2,23
SMb20320	TRAP-type large permease component	-2,22
SMc02259	periplasmic binding ABC transporter protein	-2,15
SMc02417	peptide-binding periplasmic ABC transporter protein	-2,06
SMb21273	PotD spermidine-putrescine ABC transporter substrate-binding protein precursor	-1,77
SMa1755	ABC transporter substrate-binding protein	-1,73
SMb20568	amino acid ABC transporter substrate-binding protein	-1,70
SMb21097	amino acid ABC transporter substrate-binding protein	-1,69
SMc02774	ABC transporter substrate-binding protein	-1,69
SMb21572	amino acid uptake ABC transporter substrate-binding protein precursor	-1,66
SMc02356	branched chain amino acid ABC transporter periplasmic protein	-1,63
SMc03807	AmtB ammonium transporter	-1,60
SMc02121	AapP general L-amino acid transport ATP-binding ABC transporter protein	-1,59
SMa0392	ABC transporter substrate-binding protein	-1,51
SMc04244	ZnuC high-affinity zinc uptake system ATP-binding ABC transporter protein	-1,50
SMc01605	periplasmic binding ABC transporter protein	-1,43
SMb21151	sugar uptake ABC transporter substrate-binding protein	-1,42
SMc02873	periplasmic binding ABC transporter protein	-1,35
SMc03864	amino acid-binding periplasmic ABC transporter protein	-1,33
SMc00789	DppD1 peptide ABC transporter ATP-binding protein	-1,33
SMb20428	EhuB amino acid ABC transporter substrate-binding protein	-1,31
SMc00770	PotF putrescine-binding periplasmic protein	-1,30
SMc01946	LivK leucine-specific binding protein	-1,29
SMc01828	transport transmembrane protein	-1,27
SMa0082	ABC transporter substrate-binding protein	-1,25
SMb21196	OppA oligopeptide uptake ABC transporter substrate-binding protein precursor	-1,22
SMc02378	periplasmic binding transmembrane protein	-1,18
SMb20056	CbtJ Cobalt transporter, periplasmic solute-binding protein	-1,08
SMc00265	periplasmic binding protein	-1,06
SMc02509	SitA iron-binding periplasmic ABC transporter protein	-1,01

[a]M describes the relative abundance of the corresponding proteins in the *S. meliloti* SmΔ*nfeR1*mutant relative to the wt strain as described in Materials and Methods.

In blue, mRNAs further analysed below.

Table 2.2. Periplasmic transport proteins with altered abundance in the NfeR1 mutant upon an osmotic upshift.

ID	Protein product	M[a]
Proteins negatively regulated by NfeR1		
SMc02346	ABC transporter ATP-binding protein	-5,27
SMb21432	iron uptake ABC transporter substrate-binding protein precursor	-3,17
SMb20570	aliphatic sulfonate ABC transporter substrate-binding protein	-2,95
SMc02344	periplasmic binding protein	-2,36
SMb20976	amino acid uptake ABC transporter substrate-binding protein precursor	-2,14
SMc03131	amino acid-binding periplasmic ABC transporter protein	-2,13
SMb21133	sulfate ABC transporter substrate-binding protein	-1,85
SMc03108	calcium-binding protein	-1,69
SMb20484	SupA sugar ABC transporter substrate-binding protein	-1,48
SMc02514	periplasmic binding ABC transporter protein	-1,41
SMc04171	hemolysin-type calcium-binding protein	-1,29
SMc03891	amino acid-binding periplasmic ABC transporter protein	-1,23
SMc01376	LolD ABC transporter ATP-binding protein	-1,14
SMc02873	periplasmic binding ABC transporter protein	-1,12
SMc00265	periplasmic binding protein	-1,09
SMc02774	ABC transporter substrate-binding protein	-1,08
SMb21097	amino acid ABC transporter substrate-binding protein	-1,05

[a]M describes the relative abundance of the corresponding proteins in the *S. meliloti* SmΔ*nfeR1*mutant relative to the wt strain as described in Materials and Methods.
In red, mRNAs predicted also by CopraRNA.

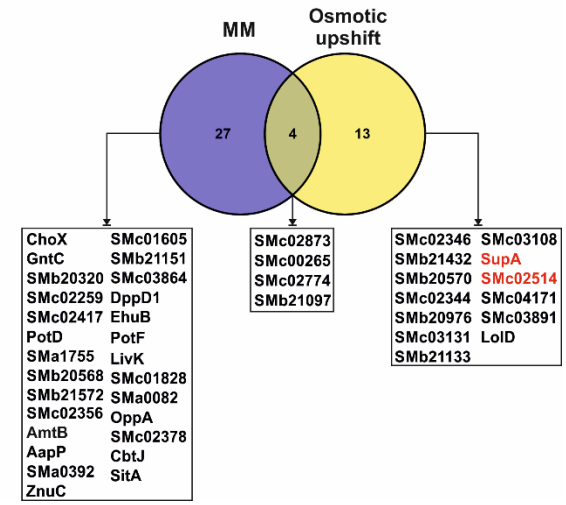

Figure 2.16. Transporter proteins likely regulated by NfeR1. Venn diagram showing the relations between the MM and osmotic upshift sets, and their overlap. In red, mRNAs predicted by CopraRNA.

2.3.4.4. *In silico* interactions between NfeR1-mRNAs

As a next step, we generated a multi-fasta file with the sequences -250 / + 100 respect to the ATG of those messengers encoding periplasmic transport proteins likely regulated by NfeR1. The file was interrogated with IntaRNA (http://rna.informatik.uni-freiburg.de/IntaRNA/Input.jsp) (Smith *et al*., 2010) (Material and Methods) to predict putative interaction sites between NfeR1 and these mRNAs. The results revealed interactions with low energy (-10 kcal/mol or lower) for some of them (Table 2.3), which probably were overlooked by CopraRNA due to a lack of conservation in other α–proteobacterial species.

Table 2.3. IntaRNA predicted interactions between NfeR1 and mRNAs encoding differentially accumulated periplasmic transporters.

Target[a]	Putative substrate[b]	Position[c]	Query[d]	Position[e]	Energy[f]
SMb20568	Amino acids	278-287	NfeR1	19-28	-19,38
SMc02356	Branched chain amino acids	148-165	NfeR1	55-66	-14,46
SMb20976	Amino acids	124-137	NfeR1	13-27	-13,54
SMc03131	Amino acids	239-246	NfeR1	19-26	-13,31
SMc00265	Oxobutyric acid	234-241	NfeR1	18-25	-13,23
supA	Sugars	240-246	NfeR1	21-27	-12,93
oppA	Oligopeptides	89-95	NfeR1	22-28	-12,38
SMb20976	Amino acids	238-244	NfeR1	21-27	-11,47
SMc02514	Glycerol	240-246	NfeR1	22-28	-11,26
SMc02417	Peptides	88-98	NfeR1	20-29	-11,06
SMb21572	Amino acids	160-166	NfeR1	19-25	-10,93

[a]Gene identity / [b]annotation according to https://iant.toulouse.inra.fr/bacteria/annotation/cgi/rhime2011/rhime2011.cgi
[c]Nucleotide positions of the mRNA involved in the predicted interaction with NfeR1
[d]sRNA query
[e]Nucleotide positions of the NfeR1 involved in the predicted interaction with the mRNA
[f]Predicted energy (kcal/mol) of the NfeR1 / target mRNA interaction
In red, mRNA predicted also by CopraRNA.

It has been reported that the *S. meliloti* response to osmostress involves changes in nutrient uptake (Rüberg *et al*., 2003; Domínguez-Ferreras *et al*., 2006). Therefore, we selected for further analysis one out of the 20 top-ranked mRNA targets predicted by CopraRNA (*SMc03121*) (Figure 2.14, framed in red), and one of those obtained in the Quantitative Proteomic Analysis and predicted by IntaRNA (*supA*) (Figure 2.16 and Tables 2.2, 2.3). These mRNAs putatively encode the periplasmic solute-binding

proteins of yet uncharacterized ABC transport systems. SMb20484 (SupA) likely binds monosaccharides, whereas SMc03121 is a predicted amino acid binding protein. Next, we performed individual IntaRNA analysis to further assert the interactions between the sRNA NfeR1 and these two mRNAs (Figure 2.17). It has been described that a particular sRNA may be able to regulate the translation/stability of more than one mRNA through interactions that differ in terms of the regions of the sRNA and mRNA involved (Papenfort *et al.*, 2010, Sharma *et al.*, 2011, Overloper *et al.*, 2014). In general terms, the presence of unpaired nucleotides, or G:U pairs, appear to be compensated for an increase in the length of the interaction. The IntaRNA graphics revealed similar possibilities for the interaction between the three aSD motifs of NfeR1 and the translation initiation site of both mRNAs (Figure 2.17 A, B and D).

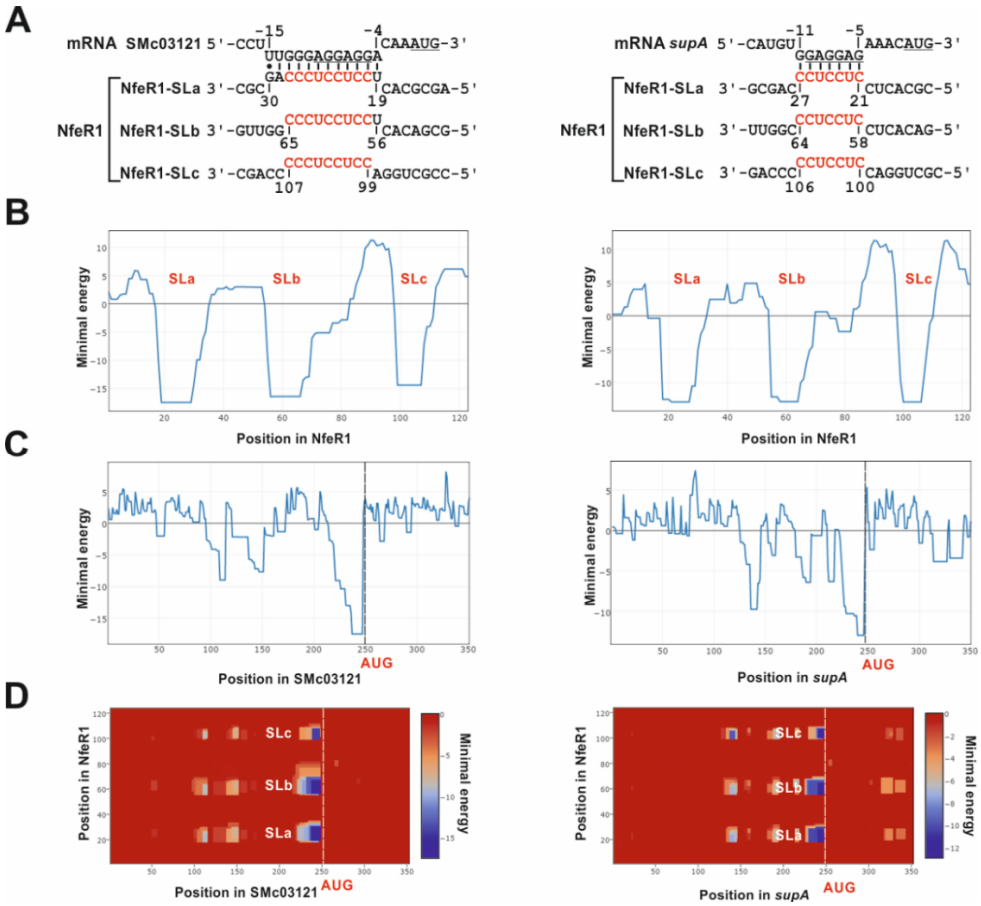

Figure 2.17. Prediction of interactions between NfeR1 and mRNAs involved in nutrient uptake. (A) Prediction of NfeR1-*Smc03121* mRNA interaction, and NfeR1-*supA* mRNA interaction. In red is indicated the

mRNA ribosome-binding site (RBS). Underlined, translational start, AUG. The numbers denote nucleotide positions in sRNA and mRNA (in this case with respect to the start of translation). **(B)** Energy diagram/position in the stem loops of the sRNA, and **(C)** in the mRNA. **(D)** Energy of the interaction in each stem loop of NfeR1.

Subsequent to this work, the hypothesis of *SMc03121* mRNA being a target of NfeR1 has been experimentally verified with an IPTG-inducible expression system (Robledo *et al*., 2017). NfeR1 down-regulates *SMc03121* transcript, as well as another mRNA with a function related to nutrient uptake, *SMb20442* (Figure 2.14, framed in red). In addition, redundant regulatory functions of the three ultraconserved aSD motifs (the three stem loops, SLa-c) have been uncovered, which probably act as interaction seeds for targeting (Robledo *et al*., 2017).

2.3.4.5. NfeR1 is an Hfq-independent sRNA

The RNA chaperone Hfq has been shown to assist most of the sRNA-mRNA target interactions documented to date. Internal single-stranded A/U-rich regions as well as a free 3′-hydroxyl end of an oligo-U stretch (e.g., of Rho-independent terminators) have been proposed as preferential sRNA interaction sites for Hfq (Schumacher *et al*., 2002; Sauer and Weichenrieder, 2011). However, the terminal uridines of the Rho-independent terminators predicted for αr14 family members are mostly base-paired to upstream sequences and hence could not be easily available for Hfq binding. In good correlation with these observations, the *S. meliloti* NfeR1 sRNA was not detected in the sub-population of transcripts co-immunoprecipitated with a chromosomally-encoded epitope-tagged Hfq protein in lysates of free-living bacteria, as described by Torres-Quesada *et al*. (2010) in the reference strain Rm1021, and later confirmed in this work in the strain Sm2B3001 (Figure 2.18, left panel). Since stability of the bacterial *trans*-sRNAs has also been assumed to be largely dependent on Hfq, we have assessed Hfq-dependent stability of NfeR1 upon inhibition of transcription with rifampicin (Figure 2.18, right panel). RNA samples were extracted from logarithmic MM cultures (OD_{600} 0.6) of Sm2B3001 and its *hfq* deletion mutant derivative Sm*Δhfq*

before or at 5, 15, 30 and 45 min following addition of the antibiotic. Half-life of NfeR1 was similar in the parent and mutant backgrounds, which suggests that Hfq does not influence the turnover of this sRNA. However, it is noteworthy that absolute NfeR1 levels in the mutant were lower than in the wild-type strain before rifampicin addition, which most likely reflects a pleiotropic indirect effect of *hfq* deletion, as already reported for other Hfq-independent sRNAs (Torres-Quesada *et al*., 2014). These results further support the notion that NfeR1 is an Hfq-independent sRNA.

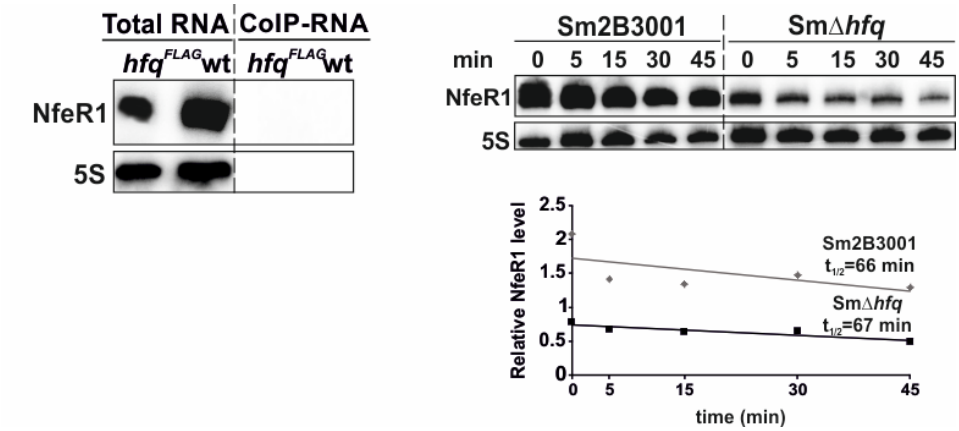

Figure 2.18. NfeR1 is an Hfq-independent sRNA. In the left panel, binding of *S. meliloti* NfeR1 sRNAs to a FLAG-epitope tagged Hfq protein and stability of the transcript in the absence of *hfq*. Northern analysis of CoIP RNA from the Sm2B3001*hfq*FLAG and wild-type strains for the detection of the NfeR1 sRNAs. Lane 1 and 2 show the expression pattern of the corresponding sRNAs in the total RNA as a control of the method. In the right panel, northern analysis of NfeR1 decay in the wild-type Sm2B3001 strain and its SmΔhfq derivative upon transcription arrest with rifampicin. RNA was extracted form bacteria withdrawn prior to or at the indicated time-points (in min) after antibiotic addition. The half-life of NfeR1 in both genetic backgrounds was calculated using the hybridization signal intensities normalized to those of the 5S rRNA (plot).

2.3.4.6. NfeR1 contributes to the silencing of nitrate/ammonia assimilation in non-stressed bacteria

The deletion of *nfeR1* resulted in reduced symbiotic efficiency of *S. meliloti* on alfalfa roots (Figure 2.11) and, at the molecular level, accumulation of mRNAs encoding proteins involved in the nitrate aerobic assimilation pathway (Table S1 in supplementary data of Robledo *et al*., 2017. http://onlinelibrary.wiley.com/doi/10.1111/1462-2920.13757/abstract) (Figure 2.19). The transcriptomic analysis

consistently revealed that the probale nitrate transporter SMb20436 (**NarK**), **NarB** nitrate reductase, **NirB**/**NirD** nitrite reductase, and main glutamine synthetase (**GlnII**) in *S. meliloti* were upregulated in the Sm*ΔnfeR1* mutant. In addition, among the accumulated proteins in fee-living Sm*ΔnfeR1* during growth in MM (a condition favoring high NfeR1 expression in the wild-type), the ammonium transporter **AmtB** was identified (Table 2.1, in blue). In nodules, the ammonia assimilation pathway is silenced to favor transfer of all fixed nitrogen to the plant. Therefore, the late symbiotic defects of the Sm*ΔnfeR1* mutant could be explained as these genes remained upregulated in nodule bacteroids. We were not able to predict thermodynamically favorable antisense interactions between NfeR1 and the mRNAs encoding these key proteins within the nitrogen assimilation pathway. Therefore, the NfeR1-dependent silencing mechanism of the nitrate/ammonia assimilation during symbiosis must be addressed in future work.

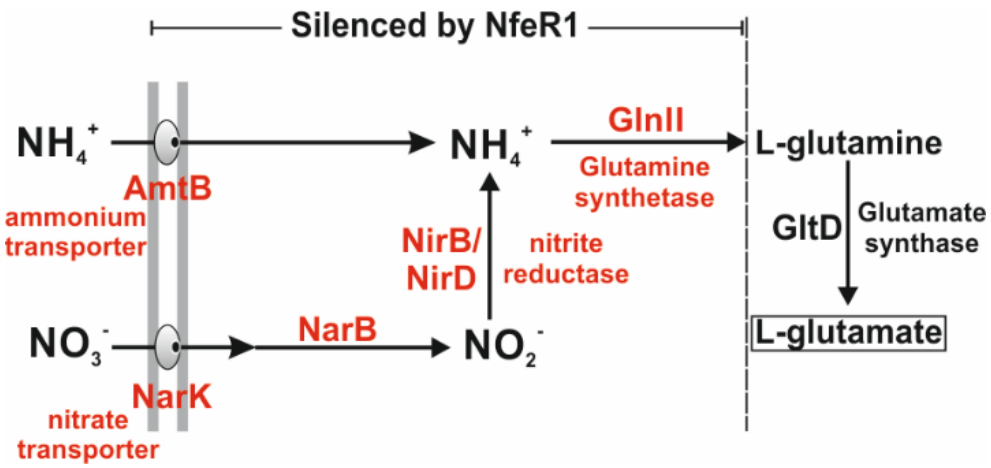

Figure 2.19. NfeR1 contributes to the silencing of nitrate/ammonia assimilation in free-living rhizobia. In red are key mRNAs/proteins within this pathway upregulated in the Sm*ΔnfeR1* mutant.

2.3. DISCUSSION

Salinity is one of the most important environmental factors limiting the agricultural productivity as a consequence of the deterioration of soil structure and its fertility. (Shrivastava and Kumar, 2015). It has been reported that highly saline soils may have adverse effects on the establishment and efficiency of the rhizobia-legume symbiosis (Zahran, 1999). High salinity affects growth and survival of soil-dwelling rhizobia as well as the infection process resulting in formation of newly developed plant organs where fixation of atmospheric nitrogen by the bacteria takes place. In order to avoid a loss in nitrogen fixation capacity, it is important to understand the underlying mechanisms employed by rhizobia for osmoadaptation under high salinity conditions. Bacterial sRNAs are now viewed as hubs of post-transcriptional regulons coupling perception of biotic and abiotic stress factors to the adequate cellular adaptive responses.

In this chapter, we have genetically approached the function of a yet uncharacterized *S. meliloti trans*-sRNA, NfeR1, which is widespread in phylogenetically related α-proteobacteria interacting with eukaryotic hosts (del Val *et al*., 2012; Reinkensmeier and Giegerich, 2015). Its expression profile and associated phenotypes placed NfeR1 as a novel regulator of a salt stress response influencing both osmoadaptation and the overall symbiotic performance of *S. meliloti* on alfalfa roots. To our knowledge, this is the first report describing a symbiotic phenotype linked to a loss-of-function of a RNA regulator in rhizobia. It has been previously shown that disruption of the *Bradyrhizobium japonicum* chromosomal *sra* locus coding for tmRNA severely impaired symbiosis of this bacterium with soybean plants (Ebeling *et al*., 1991). However, tmRNA is an almost ubiquitous sRNA that exerts housekeeping rather than regulatory functions related to translational surveillance in eubacteria and some eukaryotic organelles (Moore and Sauer, 2007).

Structural characteristics and distribution of the αr14 sRNA family

The first set of sRNAs identified in the reference strain *S. meliloti* Rm1021 included eight transcripts with genomic boundaries experimentally determined by independent approaches (del Val *et al*., 2007; Schlüter *et al*., 2010). Here, we have performed a comprehensive computational comparative analysis of the NfeR1 sRNA sequence to identify conserved structural motifs putatively relevant to its function as well as to assess its conservation patterns in bacterial genomes. The transcript represents a structurally and functionally novel prokaryotic sRNA and was collected into an Infernal model. This CM was used to accurately identify new members of the family in available sequenced bacterial genomes. This search revealed conservation of the sRNAs in bacterial species belonging to the order Rhizobiales within the α-subgroup of proteobacteria and, hence this RNA family was accordingly termed αr.

Single-copy genes hardly represent 58% of the total gene content of the *S. meliloti* genome (Galibert *et al*., 2001). The genomes of plant-interacting bacteria usually evidence high levels of paralogy suggesting that their expansion through gene duplications has been little constrained during the evolution, facilitating the acquisition of new adaptive functions for life in the soil and within plant cells (Batut *et al*., 2004; Galibert *et al*., 2001). The αr14 family members occur in multiple copies in the individual genomes. Multiple sRNA copies are not unusual in bacteria, although the physiological/ecological advantages of these reiterations have been only investigated in a subset of cases (Waters and Storz, 2009). Seemingly homologous sRNAs could act either redundantly, serving as backups in critical pathways, additively sensing different stimuli to integrate diverse environmental signals, independently, regulating different set of genes, or hierarchically upon each other (Lenz *et al*., 2004; Urban and Vogel, 2008). In this work, we have investigated the expression of the NfeR1 gene copies identified by the covariance model in *S. meliloti* Rm1021. Northern experiments, promoter predictions and reported RNASeq data (Schlüter *et al*., 2010) provide evidences for the differential regulation of the *nfeR1* gene. On the other hand, the undetectable expression of the transcripts

grouped within the αr14 sRNA family anticipates that they could be only expressed under not tested specific biological conditions to fulfill different adaptive functions in *S. meliloti*.

Collectivelly, these findings provide a basis for the forthcoming investigation of the functional plasticity and evolution of this sRNA family in *S. meliloti* and related plant-interacting bacteria.

The NfeR1 sRNA is induced by salt in *S. meliloti*

In *S. meliloti*, gene expression patterns during the osmoadaptive response differ between cells subjected to a sudden increase in the external salt concentration and those that underwent prolonged growth in the presence of the stressor, with the largest transcriptome alteration recorded 1 h upon an osmotic upshift (Rüberg *et al*., 2003; Domínguez-Ferreras *et al*., 2006). Interestingly, NfeR1 was upregulated in both conditions, whereas other stress factors such as pH oscillations or nutrient depletion at the onset of stationary phase drastically decreased or fully prevented the intracellular accumulation of this sRNA. Other stress-induced *S. meliloti trans*-sRNAs have been shown to accumulate under high salinity, but did not display such specific salt-responsive expression profiles (Torres-Quesada *et al*., 2013; Baumgardt *et al*., 2015). Transcription of *nfeR1* is driven by an RpoD-dependent promoter with a particularly strong activity, as revealed by fusions to the reporter *egfp*. NfeR1 is among the transcripts with the highest coverage scores in previous RNAseq experiments (Schlüter *et al*., 2010, 2013; Sallet *et al*., 2013; Roux *et al*., 2014). This high transcription rate is likely conferred by a conserved motif identifiable in the promoter regions of αr14 homologs from plant symbionts and intracellular mammal pathogens phylogenetically close to *S. meliloti*. This finding hints at a similar transcriptional regulation of NfeR1 homologs in α-proteobacteria within a common osmotic stress transduction cascade. NfeR1 has been previously shown to accumulate in

endosymbiotic bacteria occupying all the histologically defined zones of indeterminate nodules induced by *S. meliloti* on alfalfa roots (Roux *et al*., 2014). Accordingly, the P_{nfeR1}*::egfp* fusion revealed high *nfeR1* transcription throughout the symbiotic interaction, including early colonization of the rhizoplane by bacteria. From a practical point of view, the strength of the *nfeR1* promoter envisages the reporter P_{nfeR1}*::egfp* cassette as a suitable tool for monitoring the behavior of a wide range of rhizobial symbionts *in planta*.

NfeR1 contributes to *S. meliloti* osmoadaptation

Gene knockout usually fails to unveil sRNA functions when this is assessed by end-point phenotypes that, in many bacteria, rely on diverse and redundant pathways (Waters and Storz, 2009; Robledo *et al*., 2015). We have shown that lack of NfeR1 compromised *S. meliloti* survival to prolonged exposure to high external salt concentrations (i.e. >200 mM NaCl). Tolerance to salt stress comprises a complex process involving additive and simultaneous activation of several mechanisms, which in bacteria mostly, but not exclusively, converge in the intracellular accumulation of osmoprotectant compounds (i.e. osmolytes) such as glycine betaine or ectoine (Sleator and Hill, 2002; Flechard *et al*., 2010). Conversely, the presence of glycine betaine in the growth medium has been shown to prevent transcription of genes involved in osmostress adaptation, e.g. genes for the synthesis and uptake of various types of compatible solutes (Hoffmann *et al*., 2013). It is well known that yeast extract used in the formulation of rich media contains glycine betaine. Therefore, this could explain downregulation of the salt-responsive NfeR1 expression during growth in complete TY medium. Our transcriptomics analysis revealed that, upon bacterial growth in MM (i.e. condition promoting highest NfeR1 expression), lack of NfeR1 led to the altered expression of a set of genes involved in signal transduction, small molecule metabolism, protein folding (i.e. chaperones), catabolism of the compatible solute ectoine, or solute efflux. Proteins encoded by these genes have been functionally linked to cell processes that in rhizobia and other bacteria underlie an adequate

response to high salinity (Sleator and Hill, 2002; Rüberg *et al*. 2003; Domínguez-Ferreras *et al*. 2006). Therefore, it is plausible that the simultaneous misregulation of such diverse functions may adversely affect *S. meliloti* salt tolerance.

NfeR1 influences symbiotic performance of *S. meliloti* on alfalfa roots

In further correlation with its accumulation profile, we found that *nfeR1* deletion was also detrimental for symbiosis. Firstly, we observed that this mutation compromised nodule formation efficiency of *S. meliloti* on alfalfa roots, as revealed by both nodulation competitiveness (i.e. co-inoculation) and kinetics (i.e. single inoculation) assays. This phenotype could be explained by an impairment of the NfeR1 mutant to colonize the plant rhizosphere/rhizoplane, elicit nodulation signaling or further progress through the infection thread into the nodule primordia upon root hair infection. Complementary plant assays did support the rhizospheric competence of the mutant, therefore suggesting that NfeR1 is likely required for wild-type early host infection. Secondly, we noticed that the mutant strain mostly elicited developmentally arrested nodules with altered zonation. Their morphology resembled that of the so-called *dnf4/3* (defective nitrogen fixation) nodules developed in roots of *M. truncatula* mutants used as genetic markers of intermediate and late symbiotic stages (Starker *et al*., 2006; Domonkos *et al*., 2013; Lang and Long, 2015). The reduced shoot length of the plants developing these nodules further indicated that NfeR1 contributes to symbiotic effectiveness.

Osmotolerance and competence to establish chronic intracellular residences within eukaryotic host cells have been noticed as concurrent phenotypic traits both in mammal pathogens and nitrogen-fixing plant endosymbionts (Ohwada *et al*., 1998; Nogales *et al*., 2002; Sleator and Hill, 2002; Wemekamp-Kamphuis *et al*., 2002; Domínguez-Ferreras *et al*., 2009). Several salt-sensitive mutants of different rhizobial species (e.g. *R. tropici*, *S. meliloti* or *S. fredii*) have shown alterations in early (infectivity) and/or late (effectiveness) stages of the symbiosis with their respective

legume host under mild external conditions. It is also well known that invading bacteria perceive the plant host cell as a stressful environment, but hyperosmolarity has not been recognized yet as a physiological determinant of the endosymbiotic compartments (Soto *et al*., 2009). Therefore, our work provides another piece of genetic evidence supporting this possibility.

On top of that, transcriptomics and proteomics data suggested that NfeR1 contributes to the silencing of nitrate/ammonia assimilation in free-living bacteria by a yet unknown mechanism. In the nodule, this ammonia assimilation pathway is totally repressed in favor of symbiotic nitrogen fixation whose products are only used by the plant (Patriarca *et al*., 2002). This repression is essential for an efficient symbiosis to take place, but the underlying mechanisms are largely unknown. It is therefore tempting to speculate that NfeR1 might contribute to the silencing of ammonium assimilation in bacteroids.

NfeR1 might regulate sugar and amino acid transporter mRNAs in an Hfq-independent manner

Metabolic reprogramming also underlies bacterial adaptation to osmostress (Weber and Jung, 2002; Rüberg *et al*., 2003; Domínguez-Ferreras *et al*., 2006; Weber *et al*., 2006). In rhizobia, central carbon metabolism, energy production pathways, amino acid biosynthesis, and uptake of diverse carbon (e.g. sorbitol and mannitol) and nitrogen compounds (e.g. branched-chain amino acids) are all downregulated processes following an osmotic upshift (Rüberg *et al*., 2003; Domínguez-Ferreras *et al*., 2006). Of note, comparative biocomputational predictions anticipated a large array of mRNAs encoding ABC transporters, with putative diverse substrate preference, as the most probable primary targets of NfeR1. Transcriptome profiling upon pulse (over)expression of *trans*-acting sRNAs typically uncovers direct effects of RNA-mediated regulation at a genomic scale (Sharma and Vogel, 2009). Strikingly, none of the mRNAs whose steady-state levels were altered by salt shock-induced transcription

of NfeR1 were among the top-ranked predicted targets of this sRNA. Indeed, nutrient uptake was an underrepresented function in our transcriptomics data set. These seemingly contradictory observations could be reconciled if, *i*) most of the primary NfeR1 target mRNAs were not transcribed under our assay conditions, or *ii*) NfeR1-mediated regulation mostly proceeds without concomitant mRNA decay, as occurs, for example, in the post-transcriptional control of the *gal* operon by the well characterized *E. coli* Spot42 sRNA (Moller *et al*., 2002). Consequently, the differentially accumulated mRNAs identified in our study should be regarded as secondary molecular targets, indirectly regulated by NfeR1.

The CopraRNA algorithm predicted a large list of putative targets for NfeR1 and its homologs, revealing a functional enrichment for mRNAs derived from ABC transporter genes. Quantitative proteomics and subsequent IntaRNA analysis added mRNAs putatively encoding the periplasmic solute-binding proteins of these transport systems to the list of putative NfeR1 targets. The transcripts *SMc03121* and *supA*, encoding periplasmic amino acid- and monosaccharide-binding proteins, gave optimal energy of hybridization with NfeR1, involving its three stem loops. Subsequent to this work, *SMc03121* and *SMb20442* (also predicted by the CopraRNA algorithm) have been confirmed to be down-regulated by NfeR1 (Robledo *et al*., 2017), through a GFP-based reporter assay successfully used in previous studies to validate sRNA-mRNA base-pairing in bacteria (Chapter 1) (Urban and Vogel, 2007; Torres-Quesada *et al*., 2013, 2014; Robledo *et al*., 2015). Besides, the three single-stranded aSD sequences of the sRNA were shown to be indistinctly suited for targeting the translation initiation region of these mRNAs. Bacterial *trans*-sRNAs characterized to date typically use single or several motifs to address multiple target mRNAs (Balbontin *et al*., 2010; Papenfort *et al*., 2010; Sharma *et al*., 2011; Shao *et al*., 2013; Overlöper *et al*., 2014). In the latter case, particular sequence differences among the interaction regions determine target specificities. Thus, our data support an unprecedented redundant, rather than discriminatory, regulatory role of the three NfeR1 sites suitable for base pairing with its mRNA partners. In this scenario, the

third NfeR1 stem-loop would serve a novel dual function as both rho-independent transcriptional terminator and targeting domain. This redundancy may increase the accessibility of the regulatory motifs for base pairing with the target mRNAs, thus rendering NfeR1-mediated translational inhibition more feasible and independent of Hfq.

NfeR1 is not the only sRNA involved in post-transcriptional regulation of nutrient uptake in α–proteobacteria. In *S. meliloti*, *A. tumefaciens* and *B. abortus*, homologs of the Hfq-dependent AbcR1 and AbcR2 sRNAs have been shown to use aSD motifs for the control of large sets of transporter gene mRNAs (Chapter 1) (Wilms *et al*., 2011; Caswell *et al*., 2012; Torres-Quesada *et al*., 2013, 2014; Overlöper *et al*., 2014). Nutrient uptake has pivotal roles in rhizobial free-living growth, nodule colonization and nodule functioning (Lodwig *et al*., 2003; Mauchline *et al*., 2006; Prell and Poole, 2006; Prell *et al*., 2010), which further explains both the NfeR1-associated phenotypes and the prevalence of transport mRNAs as sRNA targets in rhizobia.

Collectively, our data anticipate that NfeR1 contributes to the metabolic remodeling demanded by the symbiotic transitions through the post-transcriptional rewiring of nutrient uptake in response to external hyperosmolarity.

APPENDIX 2

Table 2.4. Bacterial strains and plasmids used in this work.

***S. meliloti* strains**	Description	Reference/Source
Rm1021	Wild-type SU47 derivative, Sm^r	Meade *et al.*, 1982
Sm2B3001	*expR*$^+$ Sm2011 derivative; Nal^r Sm^r	Bahlawane *et al.*, 2008
Sm*ΔnfeR1*	Sm2B3001 Δ*nfeR1* derivative	This work
Sm*hfq*^{FLAG}	Sm2B3001 derivative expressing 3 x FLAG-tagged Hfq; Sm^r	This work
Sm*Δhfq*	Sm2B3001 *hfq* mutant strain; Sm^r	This work
Plasmids	**Description**	**Reference/Source**
pK18mobsacB	Suicide plasmid in S. meliloti, sacB, oriV, Kmr	Schafer *et al.*, 1994
pK18ΔnfeR1	Suicide plasmid for nfeR1 deletion; Err, Kmr	This work
pJB-T1	pJB3Tc19 derivative in which Plac was replaced by the Rho-independent terminator T1; Apr, Tcr	This work
pJBNfeR1	pJB-T1 derivative expressing NfeR1 from its native promoter; Ap^r, Tc^r	This work
pSRKKm	pBBR1MCS-2 derivative with a P_{lac} promoter, *lacIq, lacZa*$^+$, Km^r	Khan *et al.*, 2008
pSRK-C	Engineered pSRKKm lacking the LacIQ operator; Km^r	Torres-Quesada *et al.*, 2013
pSRKNfeR1	pSRK derivative constitutively expressing NfeR1; Km^r	This work
pBB*syn::egfp*	pBBR1MCS-2 derivative expressing eGFP constitutively; Km^r	This work
pBB-*egfp*	pBBR1MCS-2 derivative for generation of promoter eGFP fusions; Km^r	This work
p*nfeR1::egfp*	pBBR1MCS-2 derivative expressing a transcriptional fusion of the *nfeR1* promoter to *egfp*; Km^r	This work
p*nfeR1-40::egfp*	pBBR1MCS-2 derivative expressing a transcriptional fusion of a truncated *nfeR1* promoter (40 nt) to *egfp*; Km^r	This work
pGUS	Vector carrying the GUS reporter gene	García-Rodríguez and Toro, 2000
pR-*egfp*	Vector for generation of target mRNA-*egfp* translational fusions	Torres-Quesada *et al.*, 2013

Table 2.5. Oligonucleotides used in this work.

Name	**Sequence (5'-3')**	**Target sequence[a]**
Smr14C1	AATCGCCTTTATGACGCCCGCCGGT	206,954–206,930
NfeR1 (Smr14C2)	TCCCGGTTGCCAATCAGATCAAGCA	1,667,552–1,667,528
Smr14C3	CACGGCGCCCGGCATTCGGTCGGTT	1,667,818–1,667,794
Smr14A1	CCACGGCGCAAGACGCCGATCGGTT	1,220,715–1,220,739
Smr14A2	TTCGATATGCGACGCACCTTTCCTC	1,328,303–1,328,279
Smr14B	GGTCAGGATCGAAAGCCCGGCGCAC	1,605,895–1,605,919

5-C14	TTGTGCAGTGCATCGATCAT
3-C14	TCTAGATTCACGTTGACAGTGCTCTT
5-C14i	GAGGTACCTCGCAGTGAAACCGAGAA
3-C14i	GAGGTACCAACCCCCGGATTTTACCA
SmrC14Fw	GGATCCATCGATCGGGCAGCGCAC
SmrC14Rv	GAGCTCGGAGGCAGAAATAAACAA
T1_F	AGCAAAGAGCCGCCACGGCGCAGCCTCCGCG CCGTGGCGGTTTTTTA
T1_Rv	AGCTTAAAAAACCGCCACGGCGCGGAGGCTG CGCCGTGGCGGCTCTTT
P14C2_H	AAGCTTATTCTGTGATCATTCGGCGC
P14C2Fw	ACTAGTATTCTGTGATCATTCGGCGC
P14C2Rv	TCTAGAGCTGCCCGATCGATGATTGG
P14C2_54	CTAGTGCCCCTGGTAAAATCCGGGGGTTCG GCCTATATTCCAATCATCGATCGGGCAGCT
P14C2_54i	CTAGAGCTGCCCGATCGATGATTGGAATAT AGGCCGAACCCCCGGATTTTACCAGGGGCA
SecC14	AAACAAGCCGCCCCGGGTAT
SecC14i	GAGGAGTGTTGCCAATCCAT
5HfqMut	TCTTCATCACCGCTGCTACC
3HfqMut	AACGATCATGCCGTGAACGA

Restriction sites are underlined

Additional supplementary data can be found following this link: http://onlinelibrary.wiley.com/doi/10.1111/1462-2920.13757/abstract

CHAPTER 3

APPROACH TO THE FUNCTION OF THE RNase YbeY IN RIBOREGULATION

3.1. INTRODUCTION

3.1.1. DISTRIBUTION AND STRUCTURAL FEATURES OF YbeY IN BACTERIA

The *ybeY* gene encodes a member of a highly conserved protein family, is found in every bacterium whose genome has been sequenced and is part of the postulated minimal bacterial gene set, encompassing 206 genes. This would suggest that the *ybeY*-encoded protein either has numerous independent biological roles or affects an essential physiological function that involves an extensive range of processes (Gil *et al*., 2004; Davies *et al.,* 2008; Finn *et al*., 2014; Saramago *et al.,* 2017). Accordingly, the phenotype of *ybeY* deletion mutants is pleiotropic in some bacterial species, as it has been documented in E. coli, exhibiting a strong sensitivity to a wide range of physiological stresses. In others, removal of the gene even had toxic or lethal effects (Akerley *et al*., 2002; Kobayashi *et al*., 2003; Davies *et al.,* 2008, 2010; Leskinen *et al.,* 2015; Vercruysse., 2014). Bioinformatics analysis of the sequences of YbeY orthologs and structure-based homology analysis from the hyperthermophilic bacterium *Aquifex aeolicus* uncovered a conserved three histidine $H(X)_3H(X)_4DH$ motif, strongly resembling the active site of metal-dependent hydrolases. Besides, the analysis of the *S. meliloti* ortholog revealed global similarity of YbeY to the MID domain of eukaryotic AGO proteins involved in RNA-directed gene silencing (Oganesyan *et al.,* 2003; Davies *et al.,* 2008; Pandey *et al.,* 2011; Tatusov *et al.,* 2003; Zhan *et al.,* 2005). These features suggest that YbeY could serve catalytic and/or RNA-binding/chaperone functions in bacteria (Figure 3.1).

E. coli YbeY (from here on *Eco*YbeY) is a heat-shock protein involved in translation and, together with RNase R, has been proposed to be involved in ribosome quality control and 16S rRNA maturation of the 3' terminus of the 16S rRNA. In addition, *Eco*YbeY participates in Hfq-dependent and independent sRNA pathways (Figure 3.2) (Jacob *et al.,* 2013; Davies *et al.,* 2010; Pandey *et al.,* 2011; Rasouly *et al.,* 2009, 2010). The severe defects of *EcoybeY* deletion mutants in ribosome biogenesis have

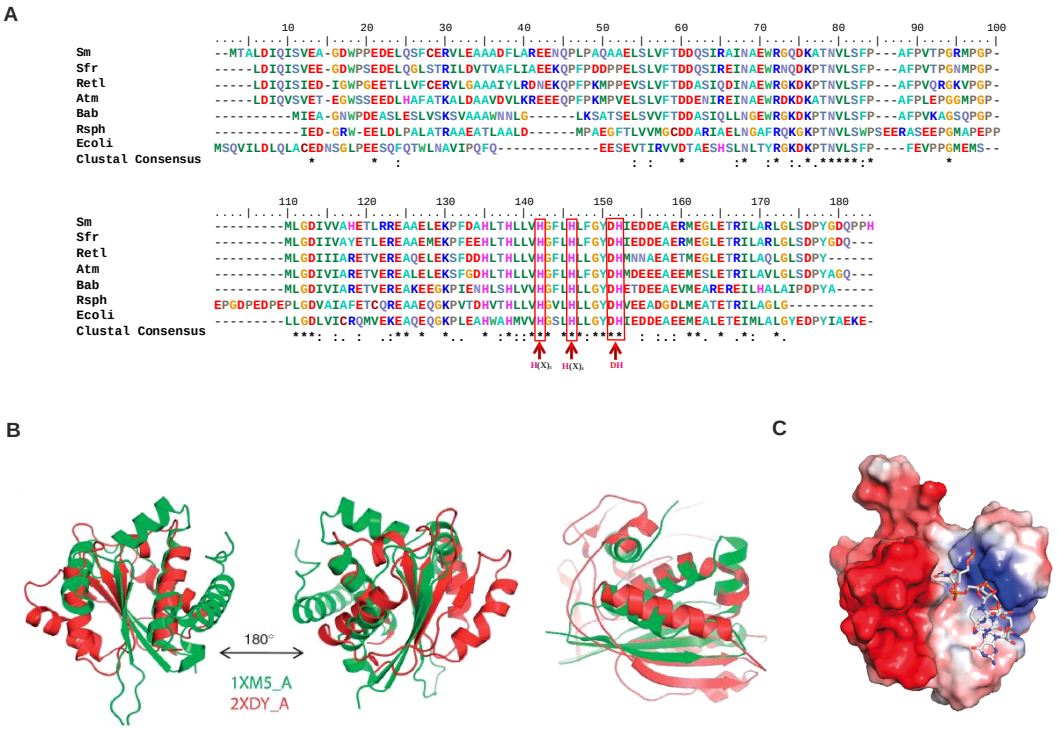

Figure 3.1. SMc01113/YbeY contains a conserved three histidine motif, shows structural similarities to the MID domain of the AGO protein and contains a probable RNA binding site. (A) Alignment of the protein sequences (generated with ClustalW) of YbeY orthologs from the representative α-proteobacterial species *S. meliloti* (Sm), *S. fredii* (Sfr), *Rhizobium etli* (Retl), *Agrobacterium tumefaciens* (Atm), *Brucella abortus* (Bab), *Rhodobacter sphaeroides* (Rsph), and the γ-proteobacterium *E. coli* (Ecoli). The conserved three histidine $H(X)_3H(X)_4DH$ motif is shared by metallo-hydrolases (framed in red) (adapted from Saramago *et al.*, 2017). **(B)** Structural alignment of Ago-Mid domain (red) and YbeY (green) from *Neurospora crassa* (left) and *Aquifex aeolicus* AGO (2NUB) proteins (right) (taken from Pandey *et al.*, 2011). **(C)** Docking of a 4mer RNA suggests a probable RNA binding site in the YbeY protein (taken from Pandey *et al.*, 2011).

been mainly attributed to a failure in *rrn* transcription antitermination and abnormal maturation of the rRNA precursors, which together result in accumulation of misprocessed 16S, 23S and 5S rRNA species (Figure 3.2, B) (Condon *et al.*, 1995; Davies *et al.*, 2010; Jacob *et al.*, 2013; Grinwald and Ron, 2013). Similar effects on rRNA maturation have been also described for YbeY orthologs of the human pathogens *Yersinia enterocolitica* and *Vibrio cholerae* (Leskinen *et al.*, 2015; Vercruysse *et al.*, 2014). Supporting these observations, it has been also reported that the purified *Eco*YbeY acts as a single strand-specific metallo-endoribonuclease with

the ability to efficiently cleave at several sites a short synthetic substrate imitating the sequence of the unprocessed 3' terminus of the 16S rRNA (Figure 3.2, C-F). Mutation in YbeY of the highly conserved histidine triad allowed the identification of two residues (H114, R59) that were found to have a significant effect for *Eco*YbeY activity *in vivo* and *in vitro,* specifically for survival under heat shock (Figure 3.2, G). Furthermore, there are strong genetic interactions between YbeY and additional RNases, including RNase III, RNase R, RNase E, RNase G and PNPase, further suggesting a role for YbeY in rRNA maturation (Figure 3.2, A) (Davies *et al.*, 2010; Zhan *et al.*, 2005; Jacob *et al.*, 2013; Sulthana *et al.*, 2016).

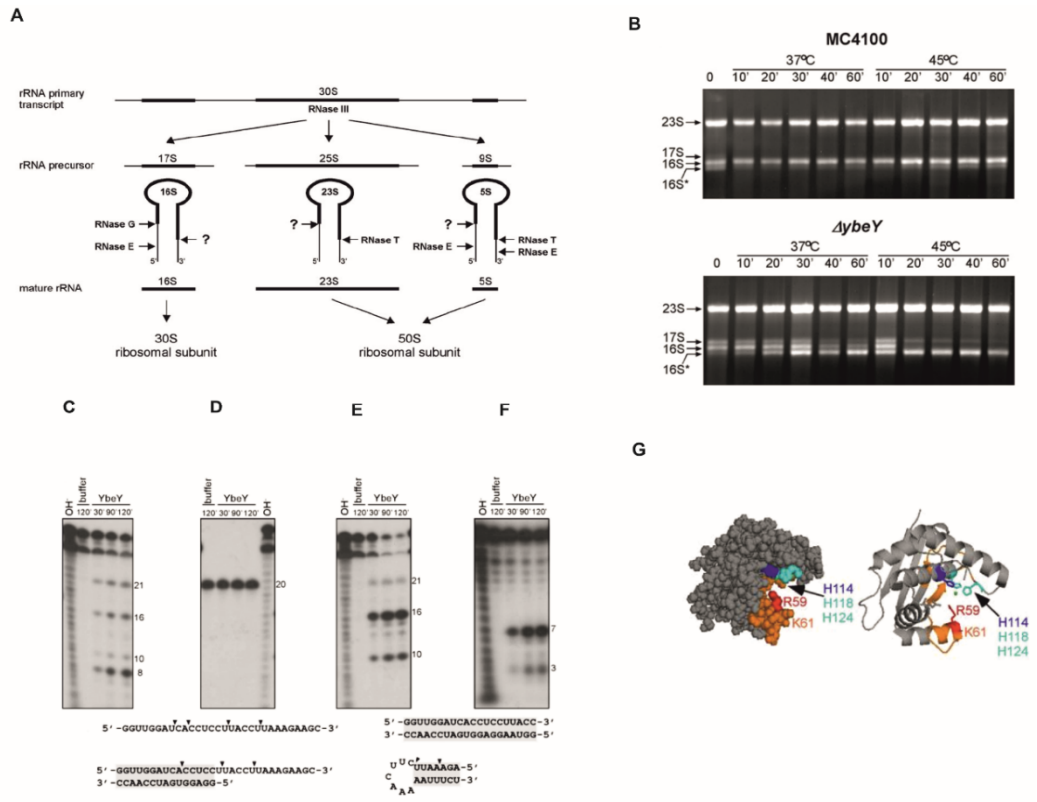

Figure 3.2. *Eco*YbeY is a heat-shock protein involved in translation, 70S ribosome quality control and 16S rRNA maturation/Roles in Hfq-dependent and independent sRNA pathway. (A). Processing of rRNA in *E. coli* (taken from Davies *et al.*, 2010). **(B)** *Eco*YbeY mutant is deficient in processing of the 16S rRNA 3´ terminus. **(C-F)** *Eco*YbeY is a single strand-specific endoribonuclease involved in 16S rRNA maturation. **(G)** Crystal structure of YbeY showing the positions of conserved residues R59, K61, H114, H118 and H124 (taken from Jacob *et al.*, 2013).

Structural modelling of the YbeY protein encoded by the *S. meliloti* genome evidenced a positively charged cavity resembling the AGO MID domain, which anchors si/miRNAs to the RNA-induced silencing complex (RISC). In bacteria, a large fraction of the known sRNAs regulates translation and/or stability of *trans*-encoded target mRNAs (Davies *et al.,* 2010; Zhan *et al.,* 2005; Jacob *et al.,* 2013; Storz *et al.,* 2011). Although the chaperone activity assisting *trans*-sRNA function has been almost exclusively attributed to Hfq (Vogel *et al.,* 2011; Sobrero and Valverde, 2012), nearly half of the sequenced bacterial genomes do not encode a recognizable Hfq homolog and several well-characterized *trans*-sRNAs have been shown to be Hfq-independent (Romby and Charpentier, 2010; Bardill and Hammer, 2012). For example, in *S. meliloti* only 14% of the annotated *trans*-acting sRNAs bind to Hfq (Torres-Quesada *et al*., 2014). These observations suggest that other proteins may assist sRNA-mediated post-transcriptional control of gene expression. In the last years, YbeY has also been proposed to fulfil this important role since: *i*) lack or depletion of YbeY results in differential accumulation of subsets of sRNAs and their predicted mRNA targets in *E. coli*, *S. meliloti* and *V. cholerae* (Vercruysse *et al.,* 2014; Pandey *et al.,* 2011; Pandey *et al.,* 2014) and *ii*) YbeY and Hfq similarly influence *S. meliloti* sensitivity to a number of stress agents and environmental cues (e.g. paraquat, SDS, ethanol, NaCl or heat), as well as the symbiotic interaction with its legume host, alfalfa (Davies *et al.,* 2007, 2008; Pandey *et al.,* 2011). However, these studies did not provide data supporting an active role of YbeY in the sRNA-mRNA interplay.

1.1.2. BIOCHEMICAL CHARACTERIZATION OF SmYbeY

The catalytic activity of YbeY from *S. meliloti* (from here on *Sm*YbeY) had remained elusive but *in vitro* assays that complement the results presented in this chapter provided insights into its biochemical properties. Incubation of purified *Sm*YbeY with a series of labeled RNA substrates revealed endoribonuclease activity of this protein

on both single-stranded (ss) and double-stranded (ds) RNA molecules. This catalytic versatility is unique among bacterial endoribonucleases, including *Eco*YbeY (Saramago *et al.,* 2017). Of note, *Sm*YbeY behaved as the prototypical double-strand endoribonuclease RNaseIII in cleaving particular structured RNA substrates (Figure 3.3).

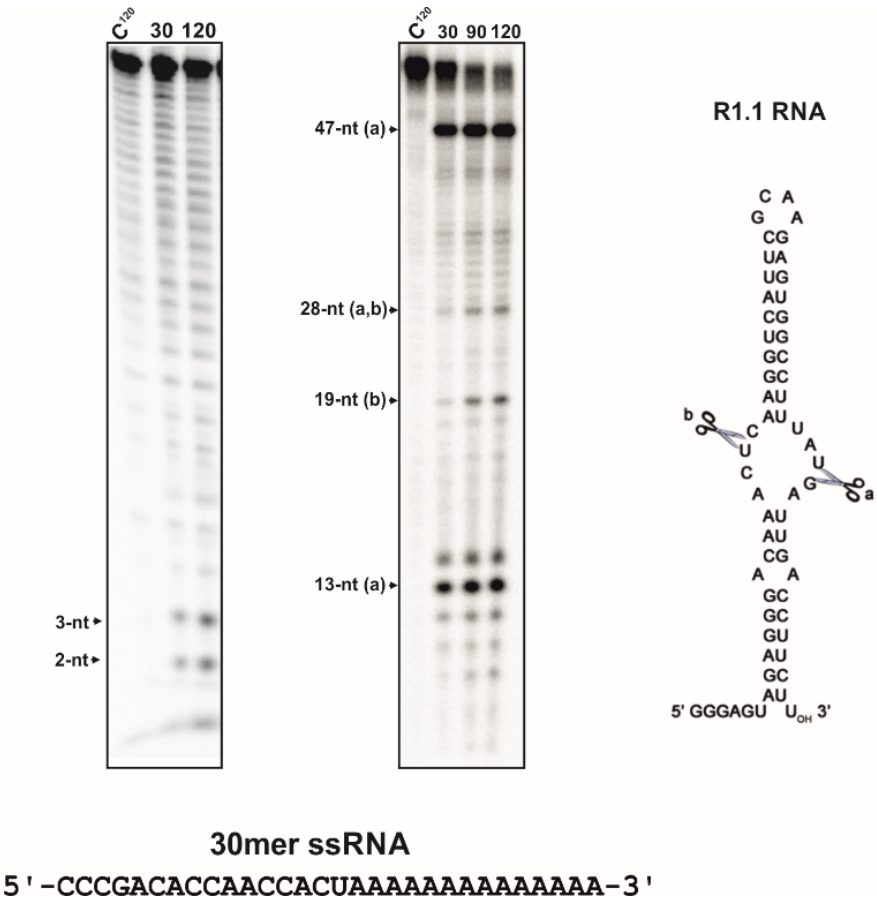

Figure 3.3. Activity of *Sm*YbeY on single-stranded RNA (ssRNA) and on structured RNA substrates. **The left panel** shows *in vitro* reactivity of purified wild-type *Sm*YbeY on a 5'-labeled 30mer ssRNA oligonucleotide, whose sequence is indicated on the left. **The right panel** shows reactivity patterns of *Sm*YbeY on the R1.1 RNA, the canonical substrate for RNase III. Numbers with arrowheads indicate sizes of the major reaction products. Known cleavage sites (a and b) of RNase III on R1.1 are indicated. Reaction times are indicated on top of the panels. C, control reactions (modified after Saramago *et al.,* 2017).

In this chapter, we further approached the function of *Sm*YbeY genetically, regarding its RNA-binding properties, associated phenotypes and influence in the *S. meliloti* transcriptome.

3.2. EXPERIMENTAL SETUP

3.2.1. CONSTRUCTION OF *S. meliloti* MUTANTS AND DERIVATIVE STRAINS

*Sm*YbeY is encoded by the gene SMc01113 located on the chromosome of reference strain Rm1021 (Galibert *et al.,* 2001). *S. meliloti* Sm∆*ybeY* and Sm*ybeY*^{FLAG} derivatives were both generated in the ExpR^{+} Rm2011 derivative strain Sm2B3001 (Bahlawane *et al.,* 2008).

3.2.1.1. Generation of the S. meliloti Sm∆*ybeY* mutant

To create an in-frame deletion of *ybeY*, 938-bp and 921-bp DNA fragments flanking the SMc01113 ORF were PCR amplified from genomic DNA with primer pairs YbeY_F1/YbeY_iR and YbeY_iF/YbeY_R1. Primers YbeY_iR and YbeY_iF carry *Hind*III sites at their 5'-ends (Table 3.2, Appendix 3). The resulting PCR products were thus restricted with *Hind*III and ligated to each other. The ligation reaction was used as a template for a second PCR with primers YbeY_FS and YbeY_RE that yielded a 1,871-bp DNA fragment flanked by *Eco*RI sites and containing the *ybeY* deletion with the junction sequence ATGACGGCGAAGCTTTAA. This fragment was digested with *Eco*RI and inserted into the suicide vector pK18*mobsacB* (Schäfer *et al.,* 1994) to yield pK18∆*ybeY*, which was mobilized to the Sm2B3001 strain by a biparental mating involving *E. coli* S17-1 (Simon *et al.,* 1983). Recombinants which underwent single and double crossover events were subsequently isolated by Kanamycin resistance and counterselection in 10% sucrose as described in (Torres-Quesada *et al.,* 2010). Deletion of *ybeY* in the selected double recombinant was confirmed by PCR with primers YbeY_MutF and YbeY_MutR (Table 3.2, Appendix 3) followed by *Hind*III restriction of the PCR product as well as by full re-sequencing of parent and mutant strain genomes on an Illumina MiSeq System applying a TruSeq,

V2 Chip (2 x 250 bp) (Figure 3.4). This analysis confirmed the in-frame deletion of *ybeY* and the absence of second site suppressor mutations in SmΔ*ybeY* (Figure 3.4).

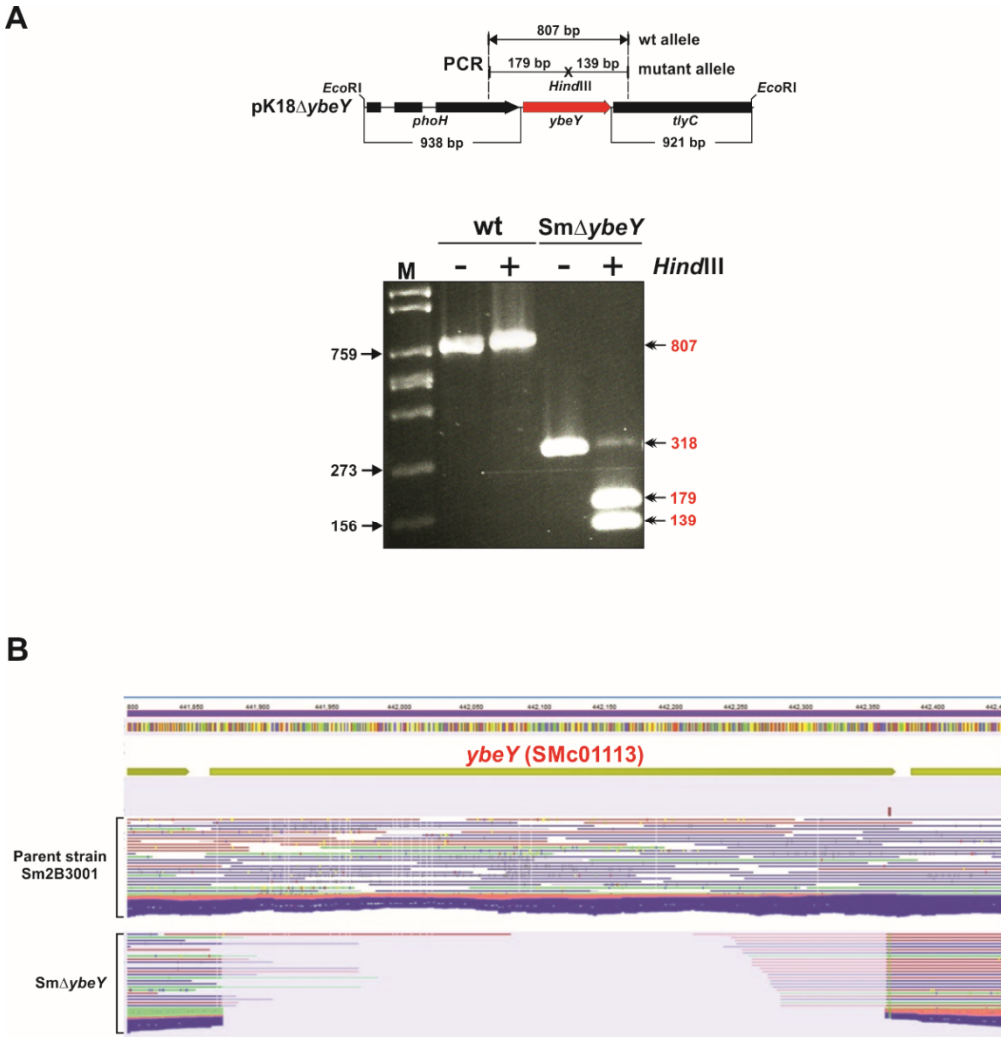

Figure 3.4. Generation of the *S. meliloti* SmΔ*ybeY* mutant. (A) Diagram of the *ybeY* genomic region cloned in pK18Δ*ybeY* to delete the gene by double homologous recombination (top). The sizes of the relevant fragments and key restriction sites are indicated. The deletion was first verified by PCR with primers 5'-*ybeY*Mut and 3'-*ybeY*Mut (Table 3.2, Appendix 3) that generate 807-bp and 318-bp amplification products in the parent and mutant strains, respectively. The mutant allele can be restricted with *Hind*III. **(B)** Sequencing of the parent and SmΔ*ybeY* genomes. Mapping of sequencing reads to the genomic *ybeY* region. Genomic DNA was sequenced in an Illumina MiHiSeq System. Mapping of sequencing reads and further screening for SNPs in the mutant strain were performed with the Genomic Workbench tool (https://www.qiagenbioinformatics.com/products/clc-genomics-workbench/). This analysis confirmed the designed in-frame deletion of *ybeY* and did not reveal second site suppressor mutations.

Mutant Sm∆*ybeY* was complemented with plasmid pJBYbeY used for ectopic expression of Sm*ybeY* from its own promoter. For plasmid construction, an 810-bp genomic DNA fragment containing the Sm*ybeY* coding sequence along with 266 nt of its upstream region was PCR-amplified with the primers pair YbeY_PrF/YbeY_PrR (Table 3.2, Appendix 3). The PCR product was first cloned into the pGEM-T® Easy vector (Promega Corporation), then retrieved by *Hind*III-*Eco*RI restriction and finally inserted into the low-copy number plasmid pJB3Tc19 (Blatny *et al.,* 1997) to generate pJBYbeY, which was conjugated into the Sm∆*ybeY* mutant by biparental mating.

3.2.1.2. Construction of *S. meliloti* strain 2B3001*ybeY*^{FLAG}

Tagging of *ybeY* with the FLAG epitope (Sigma-Aldrich) to generate the Sm*ybeY*^{FLAG} strain was done as follows. The full-length Sm*ybeY* coding sequence (devoid of its TAA stop codon) along with 935 nt of its upstream genomic region was amplified by PCR with the pair of primers YbeY_F1/YbeY_XbaI, the latter adding a *Xba*I restriction site to the 3'-end of the fragment (Table 3.2, Appendix 3). The PCR product was first cloned into pGEM-T® Easy, retrieved as a *Sac*II (genomic site internal to the PCR product)-*Xba*I fragment and finally inserted upstream of the DNA sequence coding for three tandem FLAG epitopes in the previously constructed pKS3xFLAG plasmid (Torres-Quesada *et al.,* 2010) to generate pSK5'YbeYFlag. A second 921-bp genomic DNA fragment, starting with the *Sm*YbeY stop codon, was generated by PCR with primers YbeY_iF and YbeY_RK, which add *Hind*III and *Kpn*I restriction sites to its 5'- and 3'-ends, respectively (Table 3.2, Appendix 3). This PCR product was cloned into pGEM-T® Easy, then excised as a *Hind*III-*Kpn*I DNA fragment and inserted immediately downstream of the 3xFLAG DNA sequence in pSK5'YbeYFlag to yield pKSYbeY3xFlag. Finally, a 2,372-bp DNA fragment encoding the C-terminal FLAG-tagged *Sm*YbeY was amplified from pKSYbeY3xFlag with the primers pair YbeY_FS/YbeY_RE (Table 3.2, Appendix 3), then digested with *Eco*RI and inserted into pK18*mobsacB* to generate the suicide plasmid

pK18YbeY3xFlag. This plasmid was mobilized to the Sm2B3001 strain by biparental mating for replacement of wild-type *ybeY* by the modified allele. Presence of *ybeY*^{FLAG} in several double recombinants selected as previously described was confirmed by PCR on genomic DNA with primers YbeY_MutF and YbeY_MutR (Table 3.2, Appendix 3), *Xba*I restriction of the PCR products and Western blot analysis with commercial ANTI-FLAG monoclonal antibodies (Sigma-Aldrich) following a published protocol (Torres-Quesada *et al.,* 2010).

All PCR reactions required for cloning were performed with the proofreading Phusion™ High-Fidelity DNA polymerase (Thermo Scientific). Plasmid inserts were always checked by sequencing to confirm the absence of PCR-introduced mutations.

3.2.2. CoIP-RNA PREPARATION, RNASeq AND DATA ANALYSIS

For co-immunoprecipitation (CoIP) experiments wild-type and Sm*ybeY*^{FLAG} strains were subjected to five different growth/stress conditions in 50 ml broth. Exponential and stationary cultures were obtained in TY medium upon bacterial growth to OD_{600} 0.6 and 2.8, respectively. Salt, cold and heat shocks were applied to exponentially growing bacteria in TY as explained in Material and Methods. 50 ml of each culture were then pooled by sonication before bacterial lysis. CoIP–RNA was obtained from both, control and Sm*ybeY*^{FLAG}lysates, using the ANTI-FLAG M2 affinity gel (Sigma) followed by organic extraction as described in Material and Method. This procedure was performed twice and equivalent quantities of CoIP–RNA from each replicate and strain were finally pooled.

Control and Sm*ybeY*^{FLAG} CoIP-RNA pools were further processed by Vertis Biotechnologie AG to generate two strand-specific cDNA libraries as previously described (Torres-Quesada *et al.,* 2014). Libraries were sequenced on an Illumina MiHiSeq System applying a TruSeq, V3 Chip (2 x 300 bp).

After sequencing, reads were demultiplexed based on their sequence indices and mapped with Bowtie2 version 2.1.0 (Langmead *et al.,* 2012), using standard parameters after quality trimming, to the *S. meliloti* Rm1021 reference sequence (Galibert *et al.,* 2001). Data visualization and analysis based on an updated version of the *S. meliloti* public GenDB project including annotations of identified sRNAs (Schlüter *et al.,* 2010, 2013) were done with the ReadXplorer software (Hilker *et al.,* 2014). Within ReadXplorer, the Express test and DESeq2 (Love *et al.,* 2014) tools were used to identify transcripts differentially represented in control and Sm*ybeY*^{FLAG} CoIP-RNA.

3.2.3. MICROARRAY-BASED TRANSCRIPTOMICS

Total RNA was obtained from four independent exponential (OD_{600} 0.5) and stationary (OD_{600} 2.5) cultures of each strain, i.e. Sm2B3001 and the Sm*ΔybeY* mutant (eight preparations per strain), with the RNeasy Mini Kit (Qiagen). cDNA synthesis, Cy3- and Cy5-labeling, hybridization to Sm14kOLI microarrays, image acquisition and data analysis were performed as previously described in material and methods. The Sm14kOLI microarray (ArrayExpress Accession No. A-MEXP-1760) carries 50mer to 70mer oligonucleotide probes directed against coding regions and both strands of the intergenic regions (*Sinorhizobium meliloti* Rm1021 Sm14kOLI) (Becker *et al.,* 2009). Probes in intergenic regions were separated by approximately 50 to 100 nt. Normalization and t-statistics were carried out using the EMMA 2.8.2 microarray data analysis software (Dondrup *et al.,* 2009). Genes and 5'-/3'-UTRs with P-value ≤0.05 and M ≥1.0 or ≤-1.0 were included in the analysis. The M value represents the $\log_2$ ratio between both channels.

Transcriptome data are available at ArrayExpress under accession number E-MTAB-5233. Functional categories of the differentially expressed genes were established according to the *S. meliloti* Rm1021 genome sequence annotation (Galibert *et al.,* 2001) and the KEGG database (http://www.genome.jp/kegg/).

3.2.4. REVERSE TRANSCRIPTASE PCR

In this work, the use of *S. meliloti hfq* knock-out mutant was required. The *hfq* gene corresponds to ORF SMc01048 (formerly denoted as *nrfA*) of the *S. meliloti* genome project (http://iant.toulouse.inra.fr/bacteria/annotation/cgi/rhime.cgi) of the reference strain Rm1021 (Galibert *et al.*, 2001). The mutant was constructed and further verified in strain Rm1021, as described in Torres-Quesada *et al.*, 2010 (Figure 3.5).

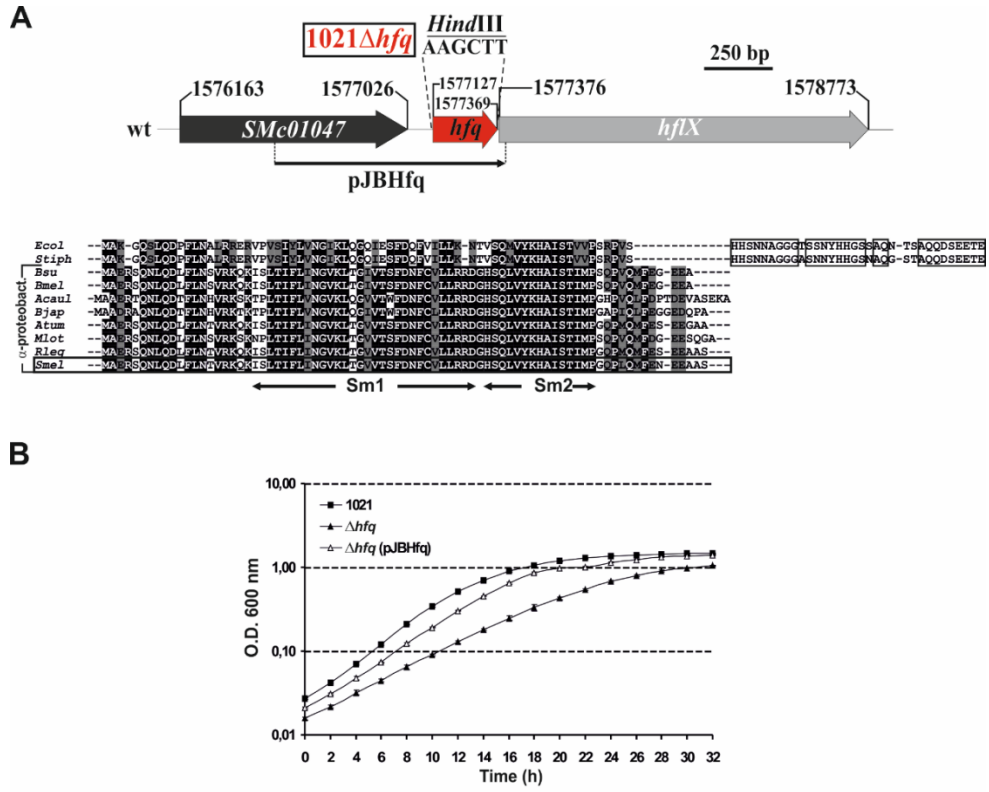

Figure 3.5. Mutational analysis of the *S. meliloti hfq* gene. (A) Arrangement of the genomic *hfq* region, multiple amino acid sequence alignment of Hfq proteins encoded by enterobacterial and α-proteobacterial genomes and details of the *hfq* mutants. The genetic map is drawn to scale. Numbering denotes the gene coordinates in the *S. meliloti* genome database. In the 1021Δ*hfq* mutant the full-length Hfq ORF was replaced by a *Hind*III site. The DNA fragment cloned on complementation plasmid pJBHfq is indicated. In the alignment, Hfq sequences are denoted by the species abbreviation as follows: *Ecol, E. coli; Stiph, Salmonella tiphymurium; Bsu, Brucella suis; Bmel, B. melitensis; Acaul, Azorhizobium caulinodans; Atum, Agrobacterium tumefaciens; Mlot, Mesorhizobium loti; Rleg, Rhizobium leguminosarum; Smel, S. meliloti*. Species belonging to the α-subdivision of the proteobacteria are indicated to the left. Shadowed are the amino acid residues conserved in at least 80% sequences and boxed are the conserved amino acids within the C-terminal extension of Hfq proteins encoded by enterobacteria. The two conserved Sm-like domains are indicated. **(B)** Growth curves in TY broth of the *S. meliloti* wild-type strain Rm1021 and its *hfq* mutant derivative as determined by OD_{600} readings of triplicate cultures in 2 h

intervals. Graphs legends: Rm1021, reference wild-type strain; Δ*hfq*, 1021 *hfq* deletion mutant; Δ*hfq*(pJBHfq), Δ*hfq* complemented with pJBHfq (modified after Torres-Quesada *et al.,* 2010)

Total RNA of the wild-type Rm1021 and 1021Δ*hfq* deletion mutant strains grown under both oxic and microoxic conditions (Material and Methods) was isolated with the RNeasy Mini Kit (Qiagen, Hilden, Germany) following manufacturer instructions. Each RNA sample (5 µg) was reverse transcribed with the AMV reverse transcriptase (Roche Diagnostics, Germany) using random hexamers as primers in 10 µl reaction mixtures. cDNA preparations were diluted to 100 µl and 1 µl of each sample was subjected to 25 cycles of PCR amplification for the detection of NifA and FixK1/K2 transcripts with primer pairs nifAFw/nifARv and fixKFw/fixKRv, respectively (Table 3.2, Appendix 3). As the reference, the abundance of the 16S RNA was assessed by amplification of each cDNA with primers 16SFw/16SRv. Possible contamination of the RNA preparations with DNA was assessed by PCR amplification of the samples with each combination of primers (Torres-Quesada *et al.,* 2010).

3.3. RESULTS

3.3.1. GENOME-WIDE PROFILING OF RNAs BOUND TO *Sm*YbeY

As global approach to explore the proficiency of *Sm*YbeY to bind RNA, the RNA species co-immunoprecipitated (CoIP-RNA) with a C-terminal FLAG-tagged variant of the protein (*Sm*YbeY^{FLAG}) expressed from the chromosome of the Sm2B3001 strain were profiled. Unlike the *ybeY* deletion mutant (SmΔ*ybeY*), the Sm*ybeY*^{FLAG} derivative strain exhibited wild-type growth in complete TY and minimal MM media, indicating that tagging did not compromise *Sm*YbeY function (Figure 3.6).

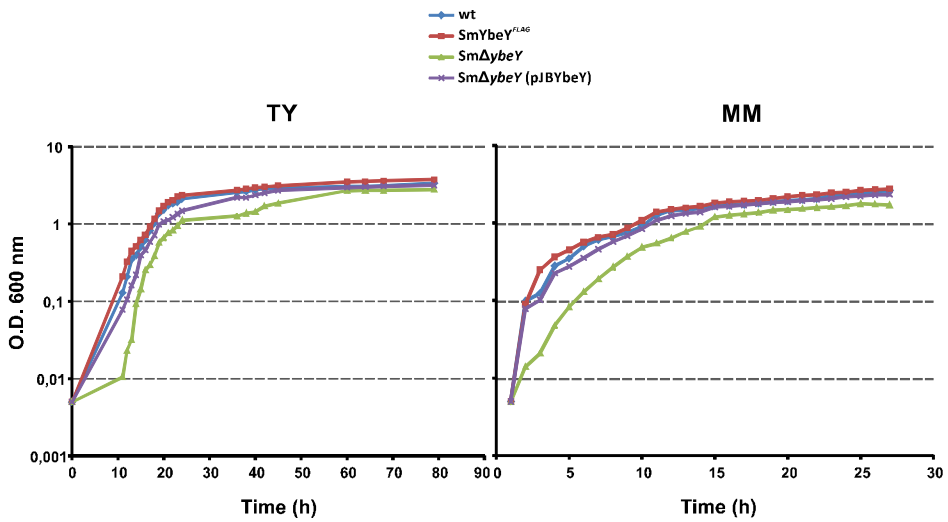

Figure 3.6. Growth curves of the Sm2B3001 strain (wt) and its *Sm*YbeY derivatives in complete TY and minimal MM media.

The untagged wild-type strain was used as control to assess unspecific RNA recovery. Control and *Sm*YbeY^{FLAG} CoIP-RNA, both obtained from bacterial pools representing five different growth conditions (i.e. exponential and stationary cultures and salt, heat and cold shocks), served as the templates to generate strand-specific cDNA libraries that were subjected to paired-end sequencing on an Illumina platform. Before organic extraction of the CoIP-RNA, the presence of *Sm*YbeY^{FLAG} in the RNA-protein complexes was verified by Western-blot (Figure 3.7).

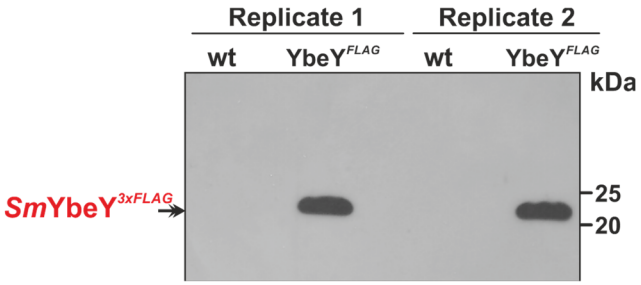

Figure 3.7. Western-blot with anti-FLAG antibody for detection of *Sm*YbeY^{FLAG} in protein-RNA complexes obtained upon CoIP with anti-FLAG.

RNAseq delivered an average of 4,200,000 reads per library of which 1,150,974 (C-wt library) and 1,406,485 (*Sm*YbeY^{FLAG}-derived library) mapped to unique locations within the reference *S. meliloti* Rm1021 genome. The DESeq2 and Express tests, both implemented in the ReadXplorer software (Hilker *et al.*, 2014), were then applied to the sets of uniquely mapped reads in order to identify transcripts differentially represented in both libraries. Transcripts covered by a minimum of 30 reads and enriched at least 2-fold in the *Sm*YbeY^{FLAG} library with respect to the control were scored as *Sm*YbeY-bound (i.e. *Sm*YbeY RNAs).

The combination of both tests rendered a catalog of 271 *Sm*YbeY RNAs (Table S3 in supplementary data of Saramago *et al.*, 2017. https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkw1234). Of those, 255 (94%) derived from mRNAs, 13 (4.8%) were annotated antisense sRNAs (asRNAs) and 3 (1.2%) represent other sRNAs (2 mRNA leaders and 1 *trans*-sRNA) (Figure 3.8). It is worth noting that the mRNA partners of 6 out of the 13 *Sm*YbeY asRNAs were catalogued as *Sm*YbeY-bound. This observation hints at certain affinity of *Sm*YbeY for asRNA-mRNA duplexes.

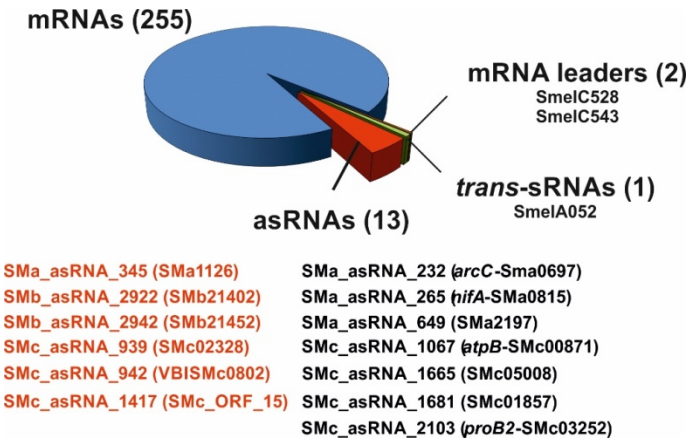

Figure 3.8. Identification of *Sm*YbeY-binding transcripts. The diagram shows the number of different RNA species, i.e. mRNAs, asRNAs, mRNA leaders and *trans*-sRNAs, enriched ≥ 2-fold in *Sm*YbeY^{FLAG} CoIP-RNA with respect to the control. The identified asRNA/mRNA pairs are depicted in red.

A comparison of the *Sm*YbeY and Hfq CoIP-RNAs, the latter obtained using a similar experimental setup in the same culture conditions (Torres-Quesada *et al.,* 2014), suggest that *Sm*YbeY does not have Hfq-like RNA chaperone features.

3.3.2. TRANSCRIPTOMIC ANALYSIS OF THE Δ*ybeY* MUTANT

The *Sm*YbeY-dependent molecular responses of *S. meliloti* were investigated by profiling the transcriptomes of the Sm2B3001 strain and its deletion mutant derivative Sm*ΔybeY* on Sm14kOLI microarrays (Tables S3 and S5 in supplementary data of Saramago *et al*., 2017. https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkw1234).

Total RNA was obtained from bacteria grown to exponential (log RNA) and stationary (stat RNA) growth phase in complete TY medium. These experiments identified 543 and 209 *Sm*YbeY-dependent mRNAs (i.e -1≥M≥1) in log and stat RNA samples, respectively, with 86 of those common to both growth states. In sum, in the experimental conditions lack of *Sm*YbeY altered the expression of 666 protein-coding genes (~11% of the *S. meliloti* Rm1021 ORFs) (Table S4 in supplementary data of Saramago *et al*., 2017. https://academic.oup.com/ nar/article-lookup/doi/10.1093/nar/gkw1234). Functional clustering of these genes revealed that 39.8% encode metabolic functions, 11.1% are related to translation and RNA turnover and 16.5% represent widely diverse cellular processes (e.g. motility, signal transduction and transcription, transposition or nitrogen-fixation) (Figure 3.9). The remaining 32.6% have unpredictable functions.

Sm14kOLI microarrays can also probe most of the recently identified *S. meliloti* non-coding RNAs (ncRNAs). Specifically, an estimation suggested at least 478 known *trans*-sRNAs to have oligonucleotide probes in these microarrays. Analysis of the hybridization signals on this set of probes revealed that lack of *Sm*YbeY altered the expression of 131 and 77 ncRNAs during exponential and stationary bacterial growth, respectively, with 17 of those present in both data sets (Table S5 in supplementary

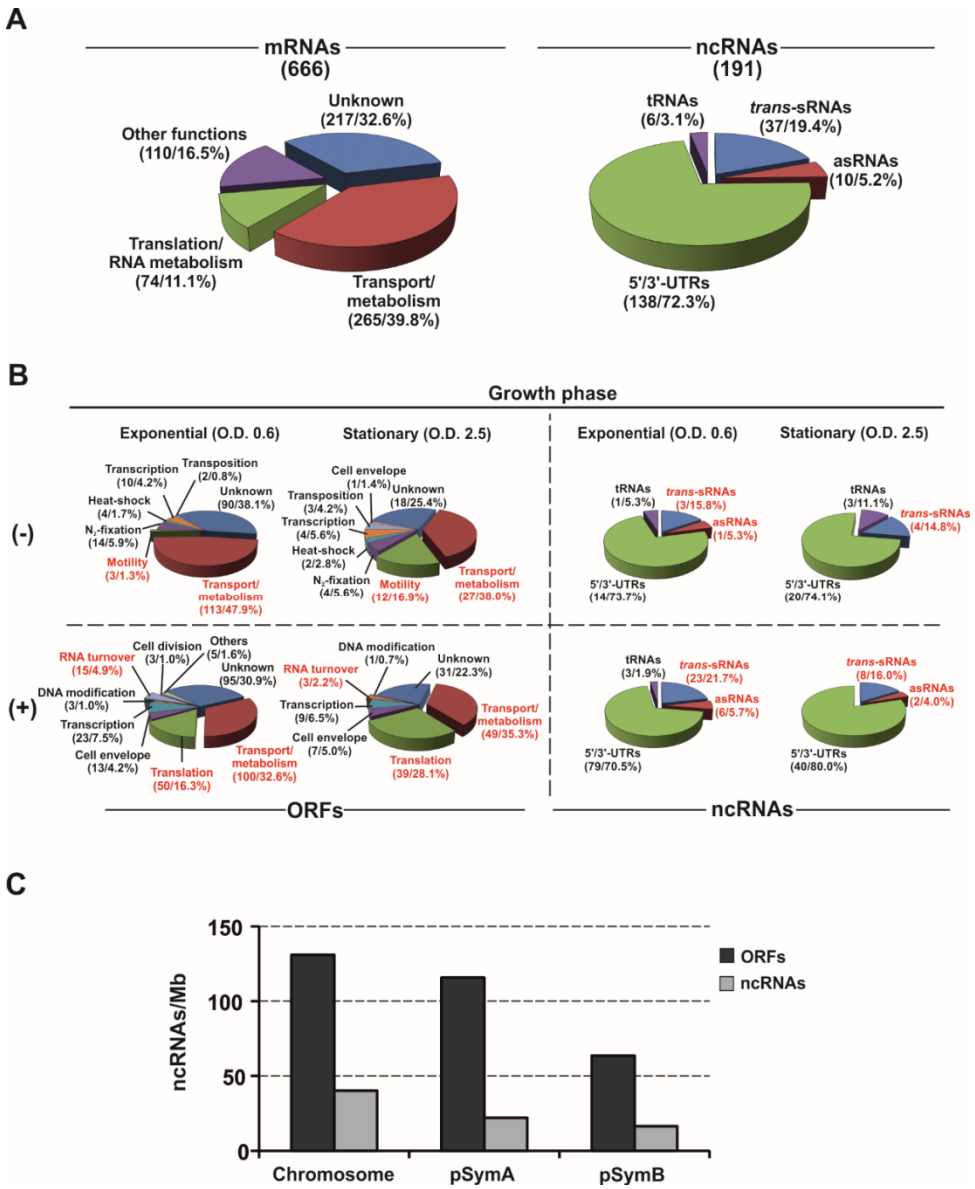

Figure 3.9. *Sm*YbeY-dependent alteration of the *S. meliloti* transcriptome. (A) Number and functional categories of mRNAs and non-coding RNAs (ncRNAs) differentially accumulated in the SmΔ*ybeY* mutant. (**B**) Number and functional categories of mRNAs and ncRNAs negatively (-) and positively (+) influenced by *Sm*YbeY, i.e. up- and down-regulated in SmΔ*ybeY*, respectively, during exponential and stationary growth phases. (**C**) Impact of *Sm*YbeY on the accumulation of chromosomal, pSymA and pSymB mRNAs and ncRNAs. The histogram shows the number of differentially expressed ncRNAs per Mb in each replicon.

data of Saramago *et al*., 2017. https://academic.oup. com/nar/article-lookup/doi/ 10.1093/nar/gkw1234). The majority (72.3%) of these ncRNAs corresponded to

putative *cis*-acting sense transcripts (i.e. 5'/3'-UTRs of mRNAs), whereas *trans*-sRNAs represented 19.4%, asRNAs 5.2% and tRNAs 3.1% (Figure 3.9, A and B).

Based on these numbers, *Sm*YbeY does not seem to have a major role at least in *trans*-sRNA turnover. However, the experimental approach underestimated the expression of asRNAs since only a subset of these transcripts (i.e. those antisense to 5'/3'-UTRs of mRNAs) are represented in the Sm14kOLI microarrays, which precludes a conclusion about the influence of *Sm*YbeY in asRNA accumulation.

The *S. meliloti* genome consists of three large replicons, the chromosome (3.7 Mb) and the two symbiotic megaplasmids pSymA (1.4 Mb) and pSymB (1.7 Mb). Interestingly, distribution of the differentially accumulated mRNAs in this genome revealed a similar relative impact of *Sm*YbeY activity on the expression of chromosomal and pSymA-borne genes (Figure 3.9, C). Chromosomal *Sm*YbeY-dependent genes included those related to ribosome biogenesis, translation, RNA turnover, energy metabolism and flagellum assembly (Figure 3.10, A). Specifically, lack of *Sm*YbeY resulted in pervasive down-regulation during both exponential and stationary growth of a set of genes coding for most of the ribosomal proteins, a set of elongation/translation initiation factors and tRNA synthetases as well as the ribonucleases RNase P, PNPase, RNase III and RNase D. A similar expression profile was exhibited by a number of genes involved in ATP synthesis and cytochrome C oxidase assembly. However, transcripts encoding substrate-dependent dehydrogenases and oxidoreductases showed variable accumulation in the Sm*Δybe*Y mutant. Microarray data also revealed up-regulation in this mutant of gene clusters coding for flagellar structural elements, particularly upon entry of bacteria into stationary phase. *S. meliloti* symbiotic plasmid pSymA mostly encodes functions contributing to ecological specializations of this bacterium, e.g. microaerobic nitrate respiration and symbiotic nitrogen fixation (Figure 3.10, B).

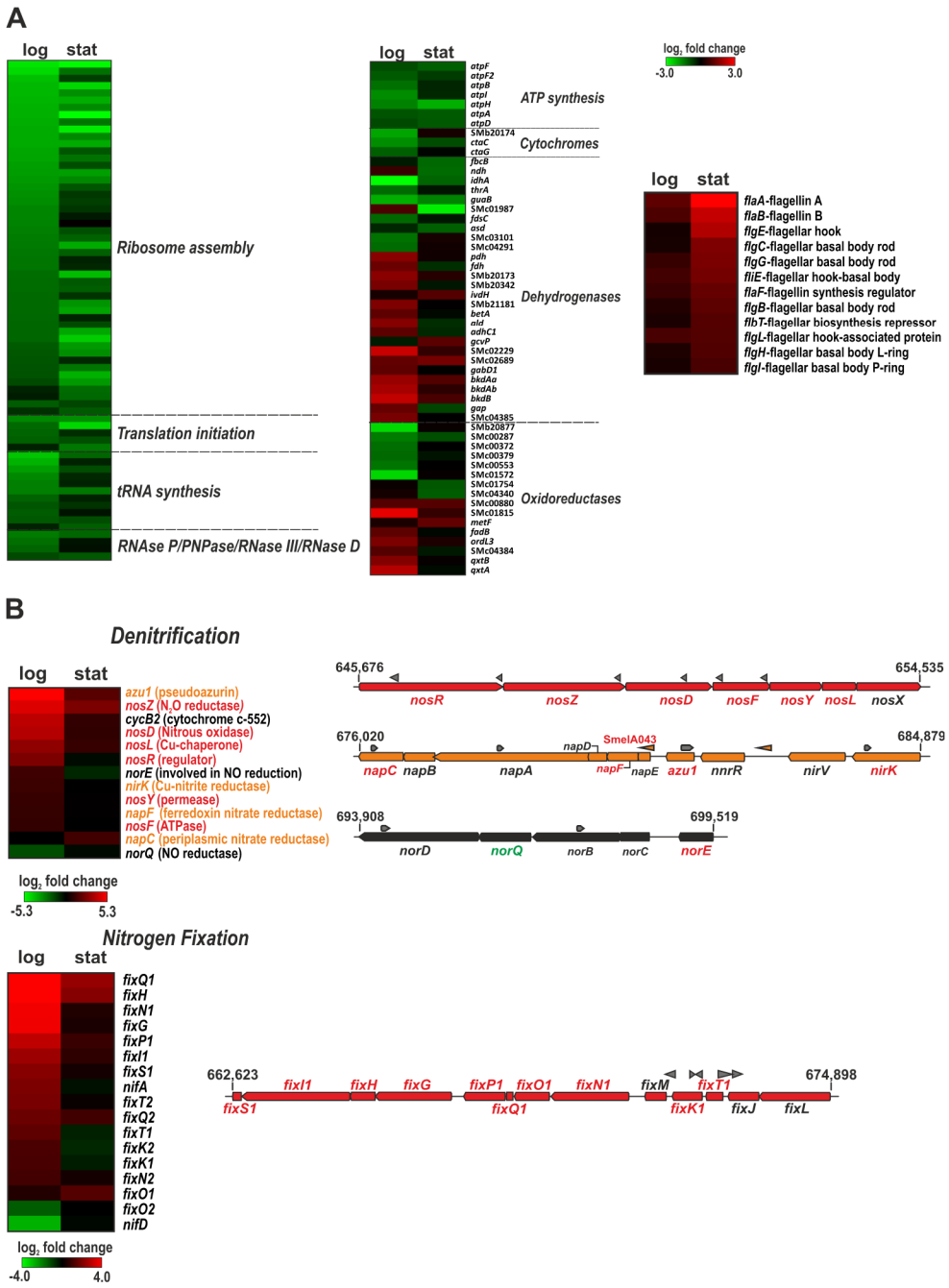

Figure 3.10. Core and plasmid pathways influenced by *Sm*YbeY. Changes in mRNA abundance in exponential (log) and stationary (stat) cultures are visualized in heat maps generated with the MeV tool (http://www.tm4.org/mev.html). In the color scale (log_2 fold changes) green and red stand for down- and up-regulation in the SmΔ*ybeY* mutant, respectively. (**A**) Chromosomal genes related to translation and RNA turnover (left), energy metabolism (middle), and flagella biosynthesis (right). (**B**) Genes of symbiotic plasmid pSymA involved in anaerobic denitrification (upper panel) and nitrogen-fixation (bottom panel). Gene clusters specifying each pathway are depicted to the side of each panel. Gray arrowheads stand for annotated asRNAs. The names of differentially expressed genes within each operon are colored (taken from Saramago *et al.*, 2017).

A number of genes coding for proteins involved in the denitrification pathway including the nitrate, nitrite and nitrous oxide reductases NapC, NirK and NosZ, respectively, were strongly up-regulated in exponentially growing Sm*Δybe*Y mutant bacteria. Similar influence of *Sm*YbeY was observed on the expression of genes coding for the master regulators of nitrogen fixation (i.e. FixK1/2 and NifA) and the elements of the electron transport chain associated with the nitrogenase activity, i.e. most of the genes integrating the *fixNOQPGHIS* operon. The accumulation profiles of these two sets of transcripts indicate an involvement of *Sm*YbeY in the post-transcriptional silencing of denitrification and nitrogen fixation under free-living non-symbiotic conditions.

3.3.3. OVERLAP BETWEEN THE Hfq- AND *Sm*YbeY-DEPENDENT GENES

Published data have evidenced large similarities between the physiological phenotypes associated to *Sm*YbeY and Hfq loss-of-function in *S. meliloti*, which has been interpreted as the consequence of a strong functional relation between the two proteins (Davies *et al*., 2007, 2008; Pandey *et al*., 2011). To further explore these commonalities at the molecular level, the Hfq- and *Sm*YbeY-dependent gene sets were compared in detail (Figure 3.11).

The latter included the 255 *Sm*YbeY-bound mRNAs identified in the CoIP experiments (Table S3 in supplementary data of Saramago *et al*., 2017. https://academic.oup.com/nar /article-lookup/doi/10.1093/nar/gkw1234), of which 23 (9%) were also scored as differentially expressed in the Sm*Δybe*Y mutant. Accumulation of almost half of these 23 transcripts was negatively influenced by *Sm*YbeY and therefore could be preferred substrates for this endoribonuclease (Table S4 in supplementary data of Saramago *et al*., 2017. https://academic .oup.com/nar/ article-lookup/doi/10.1093/nar/gkw1234). Nonetheless, the reduced overlap between these two data sets could be mostly explained by the differences in the experimental setups, i.e. transcriptomics were only performed in two out of the five culture conditions used in the CoIP experiments.

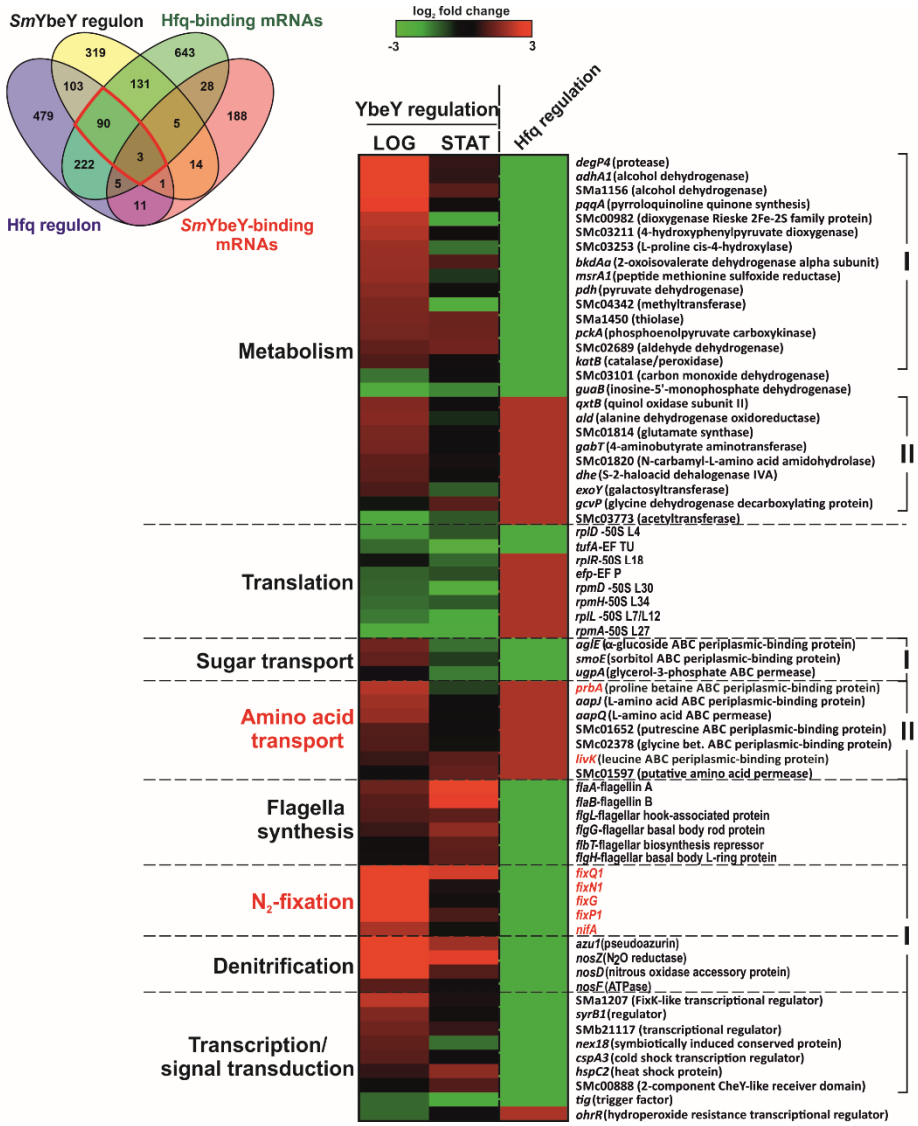

Figure 3.11. **Overlap between the Hfq- and *Sm*YbeY-dependent gene sets in *S. meliloti*.** On the left, Venn-diagram representing Hfq- and *Sm*YbeY-binding mRNAs and the transcripts differentially accumulated in the respective mutant strains. Hfq data sets were compiled from the literature. The red box indicates the 93 Hfq-*Sm*YbeY co-regulated mRNAs that are known to bind the Hfq chaperone. On the right, heat map illustrating accumulation in the *Sm*YbeY and Hfq mutants of a subset of these 93 mRNAs (Mohanty *et al.*, 2003) which encode proteins with predictable function. Functional categories are indicated to the left and the identity of each gene to the right of the panel. In the color scale (log_2 fold changes) green and red stand for down- and up-regulation in the mutants, respectively. Due to the heterogeneity of Hfq data sets, fixed values of -3 and 3 were applied for the Hfq-dependent genes. **I,** genes inversely regulated by Hfq and *Sm*YbeY (i.e. genes involved in nitrogen fixation, in red). **II,** genes negatively influenced by both proteins. In red are indicated *prbA* and *livK* mRNAs (amino acid transport), which are experimentally confirmed targets of the homologous *trans*-sRNAs AbcR1 and AbcR2 (Torres-Quesada *et al.*, 2014).

The compilation of existing transcriptomics and proteomics data uncovered a large Hfq regulon integrated by 917 protein-coding genes (Torres-Quesada *et al.,* 2010; Barra-Bily *et al.,* 2010; Gao *et al.,* 2010; Sobrero *et al.,* 2012) of which 197 (~21%) were also scored in this study as differentially expressed in the Sm∆*ybeY* mutant. Interestingly, 93 of these genes code for mRNAs that have recently been reported to bind Hfq (Torres-Quesada *et al.,* 2014). Conversely, most of the 202 *Sm*YbeY mRNA ligands have been previously catalogued as Hfq-independent. Within the subset of 93 Hfq/*Sm*YbeY co-regulated genes, 69 encode proteins with predictable function, namely 26 involved in metabolism, eight in translation, ten in nutrient uptake, six in flagella biosynthesis, five in symbiotic nitrogen-fixation, four in denitrification and nine in transcription and/or signal transduction (Figure 3.11). These 69 genes can be grouped into two major categories according to their expression patterns in the Hfq and *Sm*YbeY mutants; *i*) genes that exhibited opposite dependence on Hfq and *Sm*YbeY activity, being mostly up-regulated in Sm∆*ybeY* and consistently down-regulated in the different Hfq mutant strains (41 genes), and *ii*) genes whose expression is negatively influenced by both proteins (17 genes) and therefore have been catalogued as up-regulated in the respective mutants.

3.3.3.1. Genes inversely regulated by Hfq and *Sm*YbeY

This category includes large fractions of metabolic and regulatory genes, as well as the full subsets of genes coding for proteins involved in denitrification, nitrogen fixation, biosynthesis of flagella and sugar transport.

In the concrete case of genes involved in nitrogen fixation, previous RT-PCR experiments on RNA from bacteria grown under aerobic and microoxic conditions revealed that Hfq contributes to regulation of *nifA* and *fixK1/K2*, the genes controlling nitrogen fixation. Strikingly, the Hfq-mediated regulation of *fixK* is largely aerobiosis-dependent. Confirming results of microarray experiments where these transcripts were found with a decreased accumulation in the 1021∆*hfq* mutant (Torres-Quesada *et al.,* 2010), *fixK*-derived transcripts were readily detected in RNA from wild-type bacteria

grown under assumed aerobiosis (Figure 3.12; line 1), whereas the 1021Δ*hfq* mutant failed to accumulate these transcripts in these culture conditions (Figure 3.12; line 2). As expected, after 4 h incubation in a microoxic atmosphere (2% O_2) (Material and Methods) wild-type *fixK* expression was clearly induced as compared to aerobiosis (Figure 3.12; compare lines 1 and 3). However, similar amounts of the *fixK* mRNA were detected in the RNA from the *hfq* mutant extracted after the same treatment (Figure 3.12; line 4). In contrast, *nifA* expression was only detected after bacterial incubation in microaerobiosis (Figure 3.12; line 3), further confirming that transcription of this gene demands lower O_2 concentrations than *fixK*. A significant reduced amount of *nifA* amplification product was detected in the 1021Δ*hfq* mutant RNA, although this was still visible in ethidium bromide stained gels (Figure 3.12; lane 4).

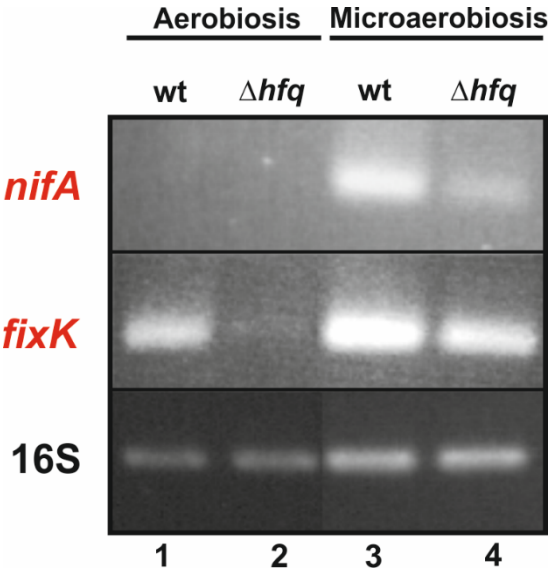

Figure 3.12. Hfq contributes to the regulation of *nifA* and *fixK* expression. RT-PCR analysis on RNA extracted from the wild-type strain Rm1021 (lanes 1 and 3) and the *hfq* mutant (lanes 2 and 4) before (lanes 1 and 2) and after (lanes 3 and 4) culture incubation for 4 h in microaerobiosis (2% O_2). 16S was amplified as constitutive control of expression. Mock-treated (no RT) RNA samples were also PCR amplified with the same primer combinations to check for absence of DNA contamination (not shown).

In sum, inverse gene regulation by Hfq and *Sm*YbeY could be explained by Hfq-mediated protection of mRNAs that are target of *Sm*YbeY degradation. This was exemplified by regulation of nitrogen fixation-related genes *nifA* and *fixK* in free-living *S. meliloti*, since *hfq* favored accumulation of both corresponding mRNAs which was counteracted by presence of Sm*ybeY*.

3.3.3.1.1. Symbiotic phenotype of the Sm∆*ybeY* mutant

Given the great impact of *Sm*YbeY activity on symbiotic and pSymA-encoded pathways we verified the nodulation phenotype associated with the *ybeY* deletion in strain Sm2B3001 (Davies *et al.,* 2007). Nodules induced by the wild-type strain and its Sm∆*ybeY* mutant 30 dpi of alfalfa plants were inspected by optical microscopy (Figure 3.13). The wild-type nodules displayed the typical elongated morphology and were pink-colored indicating active nitrogen fixation (Figure 3.13, A). Microscopy of longitudinal sections of these nodules revealed the successive characteristic areas of the histology of indeterminate nodules: apical meristem or zone I, zone of infection or zone II, interzone II-III where bacteroid differentiation begins, the zone of active nitrogen fixation or zone III occupied by mature bacteroids and, finally, the zone of proximal senescence or zone IV, that contains senescent non-functional bacteroides (Figure 3.13, C) (Vasse e*t al.,* 1990). The Sm∆*ybeY* mutant-induced nodules were white, non-fixing-like, and exhibited disorganized histology. (Figure 3.13, B and D).

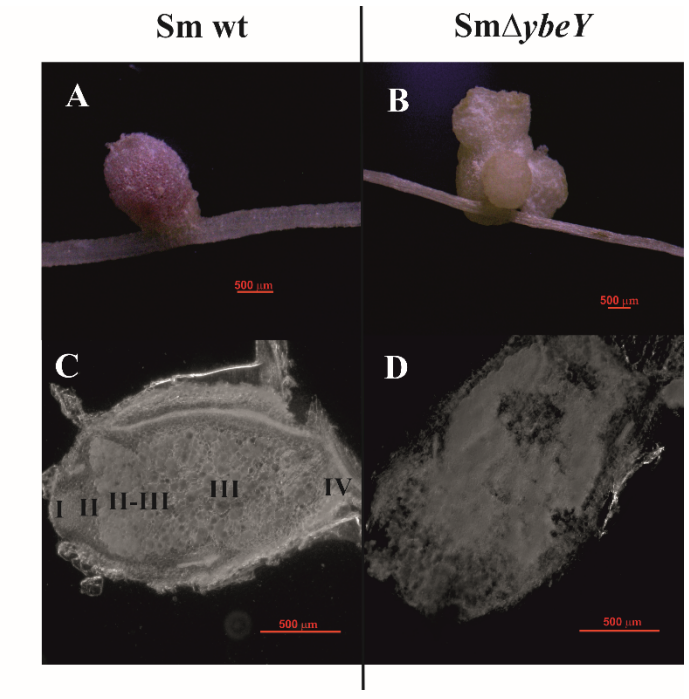

Figure 3.13.Endosymbiotic phenotype of the Sm∆*ybeY* mutant. (**A, B**) Representative enlarged images of nodules induced in alfalfa plants by strain Sm2B3001 strain (wt) and its Sm∆*ybeY* mutant derivative. (**C, D**) Microscopic images of longitudinal nodule sections. Pink nodules were the majority in the case of the wt strain; however, for the deletion mutant there was a majority of white or senescent nodules.

These results confirmed that *Sm*YbeY influences nodule development and nitrogen-fixation efficiency of *S. meliloti* on alfalfa roots. This phenotype is most likely the consequence of the pervasive misregulation of *Sm*YbeY-dependent turnover of symbiotic mRNAs.

3.3.3.2. Genes negatively influenced by Hfq and *Sm*YbeY

This second major group of Hfq/*Sm*YbeY-dependent genes codes for either metabolic proteins or amino acid transporters. *Sm*YbeY would be involved in the decay of these mRNAs upon translational inhibition by Hfq-dependent *trans*-acting sRNA partners.

Accumulation of some known Hfq-binding mRNAs encoding periplasmic components of amino acid ABC transporters is negatively influenced by both Hfq and *Sm*YbeY. The Hfq-dependent *trans*-sRNAs AbcR1 and AbcR2 have among their experimentally confirmed targets two of these mRNAs, namely *prbA* and *livK*, which code for proline betaine and branched-chain amino acid transporters, respectively (chapter 1) (Torres-Quesada *et al.*, 2013, 2014). In particular, this transcriptomics profiling revealed that levels of the *prbA* mRNA increased more than 5-fold in the Sm*Δybe*Y* mutant with respect to the wild-type strain during exponential growth, suggesting that the post-transcriptional silencing of *prbA* requires *Sm*YbeY. However, the *prbA* mRNA was not recovered in the CoIP-RNA. Therefore, the putative role of *Sm*YbeY in the regulation of *prbA* mRNA and if this layer of regulation is AbcR2-dependent was further investigated. Northern hybridization of total RNA from rifampicin-treated cells confirmed that lack of *Sm*YbeY has no effect on either processing or stability of AbcR2 (Figure 3.14). Further, these experiments also revealed that *Sm*YbeY does not influence the rRNA profile, suggesting that it is dispensable for 16S rRNA maturation. Thus, *Sm*YbeY may influence *prbA* mRNA decay upon its predicted antisense interaction with AbcR2. This hypothesis was confirmed by a fluorescence reporter assay that demonstrated decay of *prbA* mRNA to be dependent on both presence of *Sm*YbeY and functional AbcR2. Finally, biochemical evidence for mutual dependence

of *Sm*YbeY and AbcR2 in the regulation of *prbA* was provided by the detection of *prbA* mRNA degradation products as a result of YbeY catalytic activity promoted by AbcR2 *in vitro* (Saramago *et al.*, 2017).

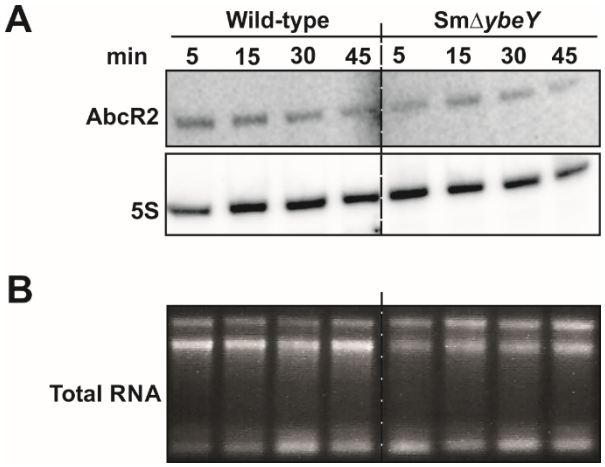

Figure 3.14. SmYbeY does not influence stability of *trans*-sRNA AbcR2. (A) Northern blot analysis of AbcR2 abundance in the wild-type strain and the Sm*ΔybeY* mutant after transcription arrest with rifampicin. Samples of exponentially growing bacteria in TY were collected at the indicated time points (in min) after antibiotic addition. Hybridization to the 5S rRNA is shown as loading control. **(B)** Ethidium bromide-stained agarose gel of total RNA extracts.

Overall, this comparative analysis revealed a discrete overlap between Hfq- and *Sm*YbeY-dependent genes. Nonetheless, it enabled the prediction of putative Hfq-dependent and -independent *Sm*YbeY substrates among the mRNAs exhibiting increased steady-state levels in the Sm*ΔybeY* mutant.

3.4. DISCUSSION

Here, the *S. meliloti* endoribonuclease YbeY was shown to influence core RNA metabolism, energy-producing pathways and plasmid-encoded symbiotic functions. Profiling of the *Sm*YbeY-dependent genes and RNA ligands envisaged a number of Hfq-independent and -dependent substrates for this RNase, e.g. the nitrogen fixation genes. Figure 3.15 summarizes the insights into the cellular pathways influenced by *Sm*YbeY. Although the data presented revealed that Hfq and *Sm*YbeY participate in largely independent RNA networks, evidence was provided that the Hfq-dependent silencing of genes related to nitrogen fixation and amino acid uptake requires *Sm*YbeY.

*Sm*YbeY acts as endoribonuclease rather than as Hfq-like RNA chaperone

Hfq is now viewed as a major RNA chaperone in bacteria that binds to and promotes stability of large sets of mRNA and sRNA transcripts (Vogel and Luisi, 2011). Given the apparent functional overlap between Hfq and *Sm*YbeY in *S. meliloti* and the homology of *Sm*YbeY to the MID RNA-binding domain of AGO proteins (Pandey *et al.,* 2011), the subpopulation of transcripts co-immunoprecipitated with a tagged version of *Sm*YbeY was profiled. This approach has been proved successful for the generation of accurate and reliable genome-wide atlas of Hfq RNA ligands and sRNA-mRNA regulatory pairs in *S. meliloti* and other phylogenetically distant bacterial species (Torres-Quesada *et al.,* 2014; Sittka *et al.,* 2008, 2009; Berghoff *et al.,* 2011; Dambach *et al.,* 2013; Chao *et al.,* 2012; Saadeh *et al.,* 2015). *Sm*YbeY CoIP-RNA was barely enriched in RNA species in comparison to the number of transcripts stably bound by Hfq in the same conditions. This enrichment profile supports a major function of *Sm*YbeY as a catalytic enzyme rather than an Hfq-like role as stabilizer and facilitator of RNA-RNA interactions *in vivo*. Therefore, other methods with enhanced sensitivity are required for the accurate genome-wide mapping of *Sm*YbeY contacts of catalytic nature on RNA substrates.

Supporting this, the catalytic activity of *Sm*YbeY on generic and endogenous RNA molecules has been revealed, which unveiled remarkable differences in its substrate specificity with respect to the well-characterized *Eco*YbeY (Saramago *et al.,* 2017; Vercruysse *et al.,* 2014; Jacob *et al.,* 2013). The latter behaves as a single-strand specific endoribonuclease unable to degrade dsRNA but exhibiting activity on short RNA hairpins and complex structured RNA (e.g. rRNA) (Vercruysse *et al.,* 2014; Jacob *et al.,* 2013). *Sm*YbeY was proficient in cleaving ssRNA, dsRNA and a number of structured RNA substrates. This versatility qualifies *Sm*YbeY an endoribonuclease as unique among bacterial endoribonucleases (Saramago *et al.,* 2017; Arraiano *et al.,* 2010).

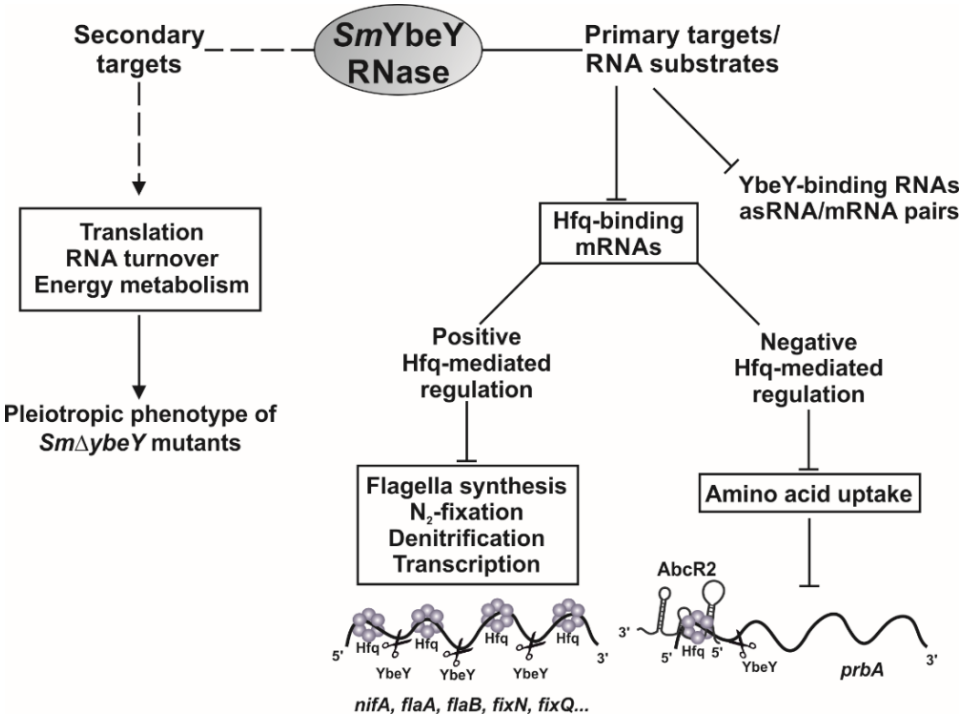

Figure 3.15. The *Sm*YbeY mRNA network. The scheme summarizes the impact of *Sm*YbeY on the *S. meliloti* transcriptome and its functional consequences.

*Sm*YbeY activity influences fundamental and symbiotic functions

Absence or depletion of YbeY in bacteria decreases growth rate, alters rRNA/ribosome profiles, enhances cell sensitivity to stress, and affects virulence and symbiotic traits (Leskinen *et al.,* 2015; Vercruysse *et al.,* 2014; Pandey *et al.,* 2011, 2014; Jacob *et al.,* 2013; Rasouly *et al.,* 2009, 2010; Davies *et al.,* 2007). To explore the molecular basis of this pleiotropic phenotype in *S. meliloti*, the *Sm*YbeY-dependent transcriptome was profiled on oligonucleotide-based microarrays. Contrary to the expected effect of the removal of an endoribonuclease on the RNA steady-state levels, down-regulated transcripts far outnumbered up-regulated transcripts in the *S. meliloti* YbeY mutant. However, this has been a common finding of similar studies addressing the influence of the activity of diverse endo- and exoribonucleases in the bacterial transcriptome, including the *E. coli* and *Thermus thermophilus* YbeY

orthologs, particularly upon stress exposure (Pandey *et al.,* 2014; Pobre and Arraiano, 2015; Stead *et al.,* 2011; Mohanty *et al.,* 2003; Ohyama *et al.,* 2014). Down-regulated genes could be mostly regarded as secondary molecular targets of *Sm*YbeY whose expression is positively influenced by this RNase in an indirect manner, e.g. involving alteration of regulatory intermediates or other yet unknown mechanisms. Nonetheless, a handful of *Sm*YbeY-bound mRNAs identified in the CoIP experiments were also found among this group of down-regulated transcripts. This finding suggests a direct interaction of *Sm*YbeY with these mRNAs in a protective mode that remains to be explored. Such a minor residual protective role in the stabilization of dsRNA and sRNAs has been already proposed for the long considered strictly catalytic ribonucleases RNase III and PNPase, respectively (Arraiano *et al.,* 2010; Bandyra *et al.,* 2016; Gan *et al.,* 2006). Genes positively influenced by *Sm*YbeY included those encoding relevant protein components of energy producing pathways, a number of RNases (e.g. PNPase, RNase III or RNase D) and key elements of the translation machinery. It is well-known that down-regulation of these fundamental physiological functions severely compromises bacterial growth and recovery upon stress exposure.

Protein mistranslation has been reported in *E. coli* YbeY mutants and was mainly attributed to the involvement of YbeY in rRNA maturation and ribosome quality control (Davies *et al.,* 2010; Jacob *et al.,* 2013; Rasouly *et al.,* 2009, 2010; Grinwald *et al.,* 2013). Indeed, misprocessing of the 16S rRNA is a major molecular phenotype linked to YbeY loss-of-function in several bacterial species (Davies *et al.,* 2010; Leskinen *et al.*, 2015; Vercruysse *et al.*, 2014; Jacob *et al.,* 2013). Contrary to what was observed in *E. coli*, *Sm*YbeY does not seem to be involved in 16S rRNA maturation (Figure 3.14, B) (Saramago *et al.,* 2017). This was an unexpected finding since it has been shown that a number of bacterial YbeY orthologs can indistinctly rescue the pleiotropic physiological phenotypes of different Δ*ybeY* mutants (Davies *et al.,* 2010). The differences in substrate specificity and probably in activity mechanisms may explain this apparent discrepancy (Saramago *et al.,* 2017). However, rRNA maturation involves the concerted activity of a suite of RNases some of which

are not essential, can exhibit functional overlap or are interchangeable (Arraiano *et al.,* 2010; Pobre and Arraiano, 2015; Stead *et al.,* 2011). Large genomes such as that of *S. meliloti* are typical sources of genetic redundancy that confers robustness to fundamental physiological processes (Galibert *et al.,* 2001). Therefore, it is not possible to rule out a role of *Sm*YbeY in rRNA maturation that may be efficiently complemented by functionally related RNases.

The transcriptomic analyses evidenced that *Sm*YbeY influenced on chromosomal and pSymA-encoded functions to a similar extent. *S. meliloti* pSymA is a megaplasmid of mosaic origin that mostly host accessory acquired genes that specify relevant strain-specific and symbiotic traits (Galibert *et al.,* 2001). Major late symbiotic functions are coordinated via the two-component regulatory system FixLJ and the master regulators of nitrogen fixation NifA and FixK under microoxic conditions within the root nodule (Bobik *et al.,* 2006). Notably, the *nifA* and *fixK* genes and a number of FixK-dependent gene clusters encoding the microaerobic denitrification pathway and elements of the electron transport chain associated with the nitrogenase complex were found to be up-regulated in the *Sm*YbeY mutant, particularly during exponential growth. Interestingly, YbeY has been also shown to severely disturb regulation of the *Yersinia* virulence plasmid pYV (Leskinen *et al.,* 2015). These independent findings hint at a major universal role of YbeY in the post-transcriptional control of prokaryotic gene networks relevant to the interaction with eukaryotic hosts.

*Sm*YbeY is involved in the regulation of Hfq-dependent mRNAs

Extensive comparison of the Hfq and *Sm*YbeY post-transcriptional regulons revealed a discrete overlap between the arrays of molecular targets of the two proteins. However, a number of known Hfq-binding mRNAs (Torres-Quesada *et al.,* 2014) exhibited accumulation patterns in the Hfq and *Sm*YbeY defective mutants compatible with a role of this RNase in the decay of these transcripts. It has been reported that Hfq actively competes for binding to the sites of the major endoribonuclease RNase E

(typically AU-rich regions) within sRNAs and mRNAs, thereby protecting these transcripts from RNase E cleavage and further exoribonucleolytic degradation (Folichon *et al.,* 2005; Moll *et al.,* 2003). The results obtained in this study revealed that sets of mRNAs with increased steady-state levels in the Sm∆*ybeY* mutant are encoded by genes positively regulated by Hfq, namely metabolic, carbohydrate transport, regulatory, flagellar, nitrogen-fixation and denitrification genes. Further supporting an Hfq-mediated stabilization, these mRNAs have been shown to be mostly recovered in their entire length by pull down with Hfq (Torres-Quesada *et al.,* 2014). Therefore, it can be hypothesized that *Sm*YbeY has a role in the silencing of non-functional transcriptional output of flagellar and oxygen-regulated FixLJ-dependent mRNAs during stationary growth and under free-living non-symbiotic conditions, respectively (Sourjik *et al.,* 2000; Rotter *et al.,* 2006). Therefore, the activity of both, the endoribonuclease YbeY and the Hfq RNA chaperone, has a pleiotropic effect on the biology of *S. meliloti*. Among the many processes in which these proteins are involved, it is worthy to note the post-transcriptional regulation of symbiotic nitrogen-fixation.

Ribonucleases are also key active players at different levels in the post-transcriptional gene silencing mediated by antisense and *trans*-acting sRNAs (Saramago *et al.,* 2014). In this work efficient cleavage of dsRNA and remarkable abundance of asRNA-mRNA duplexes in CoIP-RNA were evidenced. Thus, *Sm*YbeY-mediated silencing of some Hfq-protected mRNAs can be triggered by antisense interaction with asRNAs. In this regard, recent RNAseq-based surveys of the *S. meliloti* transcriptome have uncovered functionally significant pervasive antisense transcription of pSymA-borne symbiotic genes (Schlüter *et al.,* 2010, 2013). Therefore, the involvement of *Sm*YbeY in the asRNA-mediated decay of nitrogen fixation mRNAs is a plausible scenario that merits further investigation.

In contrast to what has been described for *E. coli* (Pandey *et al.,* 2014), transcriptomics data uncovered a scarce influence of *Sm*YbeY on the steady-state levels of *trans*-sRNAs. However, among the transcripts up-regulated in the *Sm*YbeY

mutant, a number of mRNAs coding for amino acid transporters that are putative targets of the Hfq-dependent homologous α-proteobacterial AbcR1 and AbcR2 sRNAs were found (Torres-Quesada *et al.,* 2013; Sobrero *et al.,* 2012; Wilms *et al.,* 2011; Caswell *et al.,* 2012; Overlöper *et al.,* 2014). Co-immunoprecipitation with Hfq typically recovers a specific stretch of these mRNAs, mostly derived from their 5' regions, rather than the full-length transcripts (Torres-Quesada *et al.,* 2014). This likely indicates that these mRNAs undergo ribonucleolytic degradation upon antisense interaction with their sRNA partners at sites within the stretch bound by Hfq. Recently, this hypothesis has been tested with the proline betaine *prbA* mRNA, which is an experimentally confirmed target of both AbcR1 and AbcR2 *trans*-sRNAs (Torres-Quesada *et al.,* 2014). These new data thus add *Sm*YbeY to the repertoire of bacterial ribonucleases involved in RNA-mediated silencing (Saramago *et al.,* 2017).

In summary, the highly conserved *S. meliloti* YbeY protein is a versatile endoribonuclease that influences turnover of bulk and sRNA-regulated mRNAs. The *Sm*YbeY-dependent mRNA network presented here provides a solid resource for the forthcoming investigation of *Sm*YbeY activity mechanisms underlying the post-transcriptional regulation of core RNA metabolism, energy producing pathways and late symbiotic functions in *S. meliloti*.

APPENDIX 3

Table 3.1. Bacterial strains and plasmids used in this work.

***S. meliloti* strains**	Description	Reference/Source
Rm1021	Wild-type SU47 derivative, Sm^r	Meade *et al.*, 1982
1021Δ*hfq*	1021 *hfq* mutant strain; Sm^r	Torres-Quesada *et al.*, 2010
Sm*ybeY*^{FLAG}	Sm2B3001 derivative expressing 3 x FLAG-tagged; YbeY; Sm^r	This work
SmΔ*ybeY*	Sm2B3001 *ybeY* mutant strain; Sm^r	This work
Plasmids	**Description**	**Reference/Source**
pK18mobsacB	Suicide plasmid in S. meliloti, sacB, oriV, Kmr	Schafer *et al.*, 1994
pK18Δ*ybeY*	Suicide plasmid for *ybeY* deletion; Km^r	This work
pK18YbeY3xFlag	Suicide plasmid for YbeY tagging; Km^r	This work
pJBYbeY	pJB3Tc20 derivative expressing YbeY from its native promoter; Ap^r, Tc^r	This work
pBB*syn-eGFP*	pBBR1MCS-2 derivative expressing eGFP constitutively; (Km)	This work
pR*prbA::egfp*	pR-EGFP expressing the prbA::egfp translational fusion; Apr, Tcr	This work

Table 3.2. Oligonucleotides used in this work.

Name	Sequence (5'-3')
YbeY_F1	ATTGCTGGAAGAGCGATTGC
YbeY_R1	ATCCTCGAGTTGCTCGCGTA
YbeY_iR	GTCCAAAAGCTTCATGATAAACGCGGCCGC
YbeY_iF	GTTGGGAAGCTTTAACAGTTTGGAACGATG
YbeY_RE	CGGAATTCATCCTCGAGTTGCTC
YbeY_FS	GCGAATTCTGACGTCGTCGCAACCA
YbeY_F2	CGCGTTTCATATGACGGCATTGG
YbeY_R2	CGTTCCGGATCCTTAATGCGGG
YbeY_MutF	GAGGCGCTGCAGATACTCAA
YbeY_MutR	GATGATGTGGATTTGCTGCC
YbeY_PrF	AAGCTTGATGTTCCTGACCCGTCTCG
YbeY_PrR	GAATTCCCGGCTGTGTCTTGAAGTCG
YbeY_XbaI	TCTAGAATGCGGGGGTTGGTCCCC
YbeY_RK	GGTACCATCCTCGAGTTGCTCGCGTA
5'-*ybeY*Mut	GAGGCGCTGCAGATACTCAA
3'-*ybeY*Mut	GATGATGTGGATTTGCTGCC
nifAFw	TCGTCTTGAGACCACGCTTA
nifARv	CATGACTTGGTCTATTGCGG

fixKFw	TCCATCGAGGTCGAACACCT
fixKRv	CATTTCGCCTGGGAGATGAA
16SFw	GGCTAGCGTTGTTCGGAATT
16SRv	TCCGATCCAGCCGAACTGAA

Restriction sites are underlined

Additional supplementary data can be found on the Nucleic Acids Research homepage by following this link:

https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkw1234#59935072

GENERAL DISCUSSION

In recent years, the increasing number of sequenced bacterial genomes and their functional characterization using high throughput technologies have revealed the existence of transcripts that are not translated into proteins but have different functions in the maintenance of prokaryotic physiology (Storz *et al*., 2009). This group includes bacterial riboregulators whose mechanism of action generally involves mating with complementary sequences located in the 5'-UTR regions of the target mRNAs (Majdalani *et al*., 2005). The sRNA-mRNA duplex may be a substrate or signal recognized by different cellular RNases for mRNA degradation, or can have a positive effect on mRNA stability by protecting it from the action of these ribonucleases (Majdalani *et al*., 2005). These interactions are generally facilitated and/or stabilized by the Hfq protein, mostly studied in γ-Proteobacteria and in some α-Proteobacteria (Valentin-Hansen *et al*., 2004; Chao *et al*. 2010). Bacterial sRNAs are now regarded as hubs of post-transcriptional regulons coupling perception of biotic and abiotic stress factors to generate adequate cellular adaptive responses.

The *S. meliloti* lifecycle includes either a free-living state in the soil and the rhizosphere, where the bacteria respond mainly to abiotic signals, or an endosymbiotic relationship with a leguminous plant that triggers alterations in bacterial gene expression (Patriarca *et al*., 2004). The versatility of *S. meliloti* to respond to environmental stimuli of diverse nature predicts that its genome encodes a large number of regulators. The repertoire of non-coding RNAs expressed by the legume endosymbiont *S. meliloti* belongs to the best characterized ones among those of its α-proteobacterial counterparts (del Val *et al*., 2007, 2012; Schlüter *et al*., 2010, 2013; Sallet *et al*., 2013). However, current knowledge on the function of these transcripts is limited.

This work broadens the knowledge on the importance of riboregulation in the adaptive flexibility of this diazotrophic endosymbiont by approaching the function of two *trans*-sRNAs previously identified by our research group. On top of that, we further approached the function of the highly conserved protein *Sm*YbeY genetically,

regarding its RNA-binding properties, associated phenotypes and influence in the *S. meliloti* transcriptome.

Deepening into the transcriptional regulation and targeting potential of the stress-induced AbcR2 sRNA

The *S. meliloti* AbcR1 and AbcR2 sRNAs are homologous in primary nucleotide sequence and predicted secondary structure. Both belong to the α-proteobacterial sRNA family designated αr15, whose members exist in multiple copies in the genomes of the Rhizobiaceae and Brucellaceae (del Val *et al*. 2012). Homologous sRNAs can act either redundantly to increase robustness of critical pathways, additively each contributing to different extent to a single adaptive response/phenotype, hierarchically upon each other in the same regulatory cascade or independently, influencing different or scarcely overlapping response pathways and arrays of target mRNAs. Several experimental evidences suggest that the *A. tumefaciens* and *S. meliloti* AbcR1 and AbcR2 sRNAs act independently to regulate nutrient uptake, whereas their *Brucella* homologs have a rather redundant function (Wilms *et al*., 2011; Caswell *et al*., 2012; Torres-Quesada *et al*., 2013; Overlöper *et al*., 2014). In *S. meliloti* Rm1021, AbcR1 and AbcR2 exhibit a divergent unrelated expression profile. AbcR2 is induced upon entry of bacteria into stationary phase and under a number of abiotic stresses (e.g., osmotic upshift, EtOH induced membrane stress and pH oscillations), whereas AbcR1 is transcribed in actively dividing bacteria, either in culture, rhizosphere or within the invasion zone of mature alfalfa nodules. (Torres-Quesada *et al*., 2013). According to this, our analysis of the AbcR1/2 promoter regions as well as a previous study by Schlüter *et al*. (2010), suggested a differential regulation of the *abcR1/2* genes, while our Northern experiments also revealed the dependence of *abcR2* transcription on the alternative RNA polymerase sigma factor RpoH1.

In this work, we also validated the AbcR2-mediated translational control of the *SMa0495* and *prbA* mRNAs *in vivo*, both coding for periplasmic amino acid-binding

proteins of ABC transport systems and identified in the Hfq CoIP-RNA (Torres-Quesada *et al*., 2014). Computational predictions suggested that downregulation of *SMa0495* translation involves the interaction of the mRNA with an alternative aSD sequence (M2) that remains single-stranded between the two first hairpins of AbcR1 and AbcR2. M2 has been also shown to be a functional targeting motif in the *A. tumefaciens* AbcR1 homolog (Overlöper *et al.*, 2014). Interestingly, in spite of sharing an aSD motif, the 5′ loops of AbcR1 and AbcR2 differ in some nucleotides and therefore, it is tempting to speculate that this sequence stretch is the functional discriminatory domain for the targeting of specific sets of mRNAs (del Val *et al.*, 2012; Torres-Quesada *et al.,* 2013).

In summary, our data suggest that the *S. meliloti* AbcR2 *trans*-sRNA uses canonical Hfq-dependent antisense mechanisms for the selective post-transcriptional silencing of amino acid ABC transporters to optimize nutrient uptake under diverse environmental conditions, and under the control of an alternative RNA polymerase sigma factor.

The conserved NfeR1 sRNA contributes to osmoadaption and symbiotic efficiency

In addition to the possible interaction with cellular proteins, functional characterization of *trans*-sRNAs should address their structure and phylogeny, their expression pattern and related phenotypes, as well as the identification of their target mRNAs. The initial alignments of the *trans*-encoded NfeR1 sRNA with its group of homologs identified by BLASTN (identity of primary nucleotide sequences) allowed establishing its secondary consensus structures defined by the nucleotide covariance model (CM). This model was used to interrogate bacterial genomes to search new members of this sRNA family called αr14. These comparisons showed that NfeR1 homologs are present only in the subgroup of α-proteobacteria. The phylogeny of this sRNA, established according to its secondary structure, is correlated with the phylogeny of classical chromosomal markers (e.g., 16S RNA), so that NfeR1 was

always grouped with its homologs from the phylogenetically closest species of *S. meliloti,* including members of the *Rhizobiaceae*, *Brucellaceae* and *Phyllobacteriaceae* families. The subgroup of α-proteobacteria includes bacterial species of diverse biology although many of them share the ability to establish long-term symbiotic or pathogenic interactions with eukaryotic hosts (Batut *et al*., 2004). The αr14 CM also identified up to five additional predicted copies of the query NfeR1-encoding gene in the *S. meliloti* Rm1021 genome, although the chromosomally encoded 123-nt long NfeR1 transcript was the only copy detected in our experimental analysis.

NfeR1 was found to be dispensable for free-living growth of *S. meliloti* in high-salinity minimal medium, whereas it is required for the expression of genes involved in stress adaptation, osmolyte catabolism and membrane trafficking in osmotic stress conditions. Likewise, this sRNA is required for the establishment of an efficient symbiosis with its cognate legume host alfalfa under laboratory conditions, contributing to nodulation competitiveness, infectivity, nodule development and global symbiotic efficiency. The determination of the NfeR1 expression profile revealed *nfeR1* transcription to be dependent on sigma factor RpoD, and enhanced by a conserved motif in the promoter regions of α–proteobacterial homologs. The NfeR1-associated phenotypes provide another piece of genetic evidence supporting the possibility that the endosymbiotic compartments can be hyperosmotic (Ohwada *et al*., 1998; Nogales *et al*., 2002; Sleator and Hill, 2002; Wemekamp-Kamphuis *et al*., 2002; Domínguez-Ferreras *et al*., 2009; Soto *et al*., 2009).

The genome-based prediction analysis of target mRNAs of NfeR1 have revealed the existence of typical interactions characterized by imperfect and discontinuous complementarity of the nucleotide sequences near the Shine-Dalgarno site in the mRNA, and involving the three aSDs of the sRNA in an unprecedented redundant way. The predicted target mRNAs of NfeR1 indicated a massive regulation of ABC transporters by this sRNA.

Overlapping regulon of the *S. meliloti* AbcR1/2 and NfeR1 sRNAs regulating ABC transporter proteins

Our results showed that the two sRNAs under study, NfeR1 and Hfq-dependent AbcR2, target multiple mRNAs encoding ABC transporters in *S. meliloti*. In *S. meliloti*, *A. tumefaciens* and *B. abortus*, homologs of AbcR1/2 have been shown to use aSD motifs for the control of large sets of transporter gene mRNAs (Wilms *et al.*, 2011; Caswell *et al.*, 2012; Torres-Quesada *et al.*, 2013, 2014; Overlöper *et al.*, 2014). AbcR1, AbcR2 and NfeR1 exhibit markedly distinct expression profiles in *S. meliloti*, while sharing part of their targeting potential (Figure D-1) (Torres-Quesada *et al.*, 2013, 2014). Thus, these three sRNAs and their sets of target mRNAs likely integrate a dense overlapping regulon for the coordinated post-transcriptional control of largely common sets of transport genes in response to different biological cues (Beisel and Storz, 2010) (Figure D-1). Nutrient uptake has pivotal roles in rhizobial free-living growth, nodule colonization and nodule functioning (Lodwig *et al.*, 2003; Mauchline *et al.*, 2006; Prell and Poole, 2006; Prell *et al.*, 2010), which further explains both, the sRNAs-associated phenotypes and the prevalence of transport mRNAs as sRNA targets in rhizobia. The regulatory activity of these sRNAs would contribute to coordinate nutrient uptake with the metabolic reprogramming concomitant to symbiotic transitions.

The functional characterization of *S. meliloti* AbcR1/2 and NfeR1 sRNAs might contribute to elucidate the common strategies used by this group of bacteria to colonize and/or survive in the free-living state and in eukaryotic host cells. The full regulatory potential of AbcR1/2 and NfeR1 sRNAs should be addressed in future work, particularly by placing special focus on the experimental validation of more predicted mRNA targets as well as the deciphering of molecular interactions between the verified sRNA-mRNA pairs.

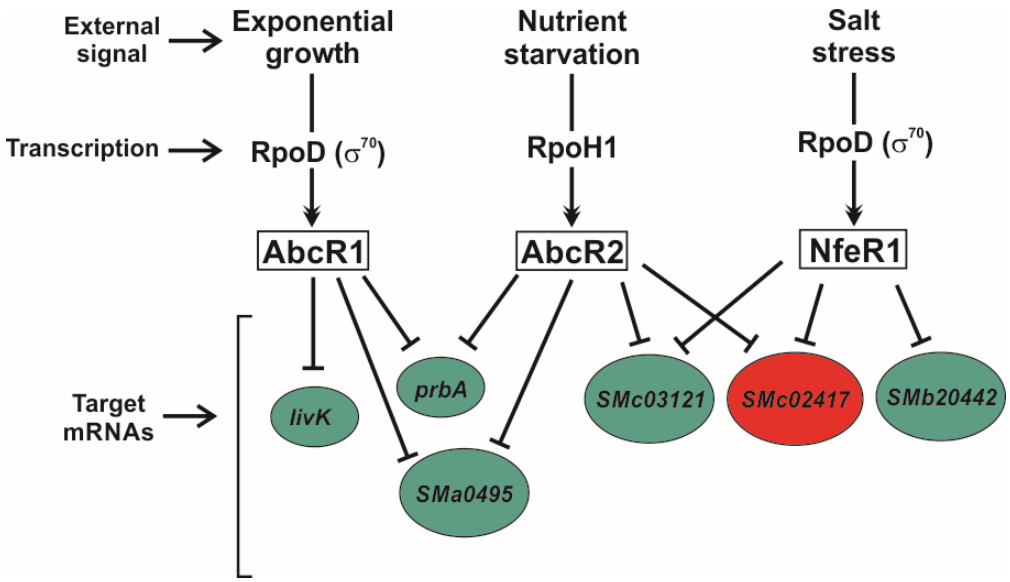

Figure D-1. Dense overlapping regulon of the *S. meliloti* AbcR1/2 and NfeR1 sRNAs. AbcR2, its homolog AbcR1, and NfeR1 govern a large dense overlapping regulon for the control of nutrient uptake in response to different environmental signals. The target mRNA in red was not validated. The mRNAs in green have been experimentally validated: AbcR1 targets (Torres-Quesada *et al*., 2014); AbcR2 targets (in this work); NfeR1 targets (Robledo *et al*., 2017). All indicated mRNAs were also predicted by *in silico* analysis, experimentally identified by proteomic analysis and/or obtained from a genome-wide profiling of Hfq-binding RNAs (Torres-Quesada *et al*., 2014).

YbeY is a novel RNase involved in the riboregulation of amino acid transporters and nitrogen fixation

One of the aims of this work was the identification and characterization of proteins that, in addition or as an alternative to Hfq, could be involved in the activity of AbcR1/2 and NfeR1. The specific study of YbeY (*Sm*YbeY) was proposed for its structural homology with the MID domain of the Argonaute eukaryote protein involved in RNA-mediated silencing and its proposed functional relationship with Hfq (Pandey *et al*., 2011). However, the results presented in this work have contributed to expose that *Sm*YbeY rather is an endoribonuclease than an Hfq-like RNA chaperone (Figure D-2) (Saramago *et al*., 2017). *Sm*YbeY efficiently degrades double- and single-stranded RNA, a versatility that has not been previously described for bacterial

RNases. Among the mRNA substrates of *Sm*YbeY, we have identified *livK* and *prbA*, both targets of AbcR1/2, as well as *nifA*, encoding the transcriptional activator of the structural genes of nitrogenase in microaerobiosis, to which an asRNA (SMa_asRNA_265) is associated (Figure D-2). Subsequent to this work, evidence was provided for *Sm*YbeY contributing to AbcR2-mediated silencing of *prbA* (Saramago *et al*., 2017), possibly by cutting the duplex sRNA-mRNA, a function largely attributed to RNase III as a prototype endoribonuclease of double-stranded substrates (Figure D-2) (Lasa *et al*., 2011).

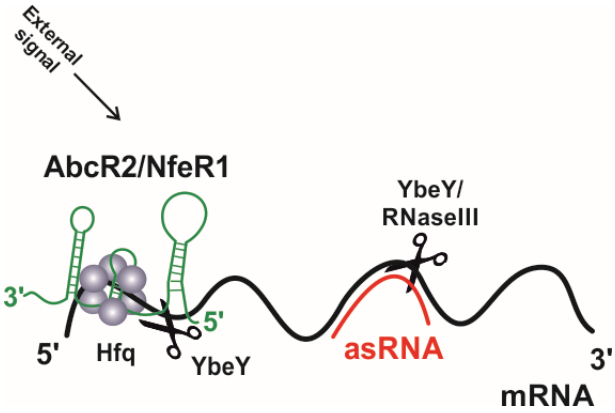

Figure D-2. Activity mechanism of the Hfq chaperone and the YbeY endoribonuclease in riboregulation.

CONCLUSIONS

1. Transcription of the stress-induced Hfq-dependent AbcR2 sRNA is driven by the alternative RNA polymerase sigma factor RpoH1.

2. The catalog of Hfq-binding RNA species is a reliable resource for the identification of sRNA-mRNA regulatory pairs in *S. meliloti*. As a proof of principle, AbcR2 targeting of the Hfq-bound amino acid transporter mRNAs *prbA* and *SMa0495* was confirmed. AbcR2 most likely downregulates these mRNAs by a canonical mechanism involving base-pairing interactions at the translation initiation region with either of the two unpaired anti-Shine-Dalgarno (aSD) motifs within the sRNA, which may serve a discriminatory function for targeting.

3. The *S. meliloti* NfeR1 sRNA is widely conserved in species of the *Rhizobiaceae*, *Brucellaceae* and *Phyllobacteriaceae* families within the large α–subgroup of proteobacteria, integrating the so-called αr14 family of sRNAs. The *S. meliloti* reference strain Rm1021 encodes six αr14 homologs but only the chromosomal 123-nt long NfeR1 transcript was reliably detected by Northern hybridization.

4. Expression of NfeR1 is induced in response to salt stress and throughout the symbiotic interaction of *S. meliloti* with alfalfa. The strength and differential regulation of *nfeR1* transcription are conferred by a motif, which is conserved in the promoter regions of its α–proteobacterial homologs.

5. NfeR1 positively influences osmoadaptation of free-living bacteria and is required for wild-type expression of an array of salt-responsive genes related to stress adaptation, osmolyte catabolism and membrane trafficking. In symbiosis, NfeR1 contributes to nodulation competitiveness, infectivity, nodule development and overall symbiotic efficiency of *S. meliloti* on alfalfa roots. The NfeR1 expression profile and its associated phenotypes provide genetic evidence of the yet undemonstrated hyperosmotic conditions of the endosymbiotic compartments.

6. Comparative computer predictions anticipated a redundant role of three identical NfeR1 unpaired aSD motifs for targeting of multiple mRNAs encoding ABC transporters with diverse substrate preference. This hypothesis has been later confirmed experimentally.

7. AbcR2- and NfeR1-dependent periplasmic proteomes identified multiple mRNAs from ABC transporter genes as additional putative targets of these sRNAs, thus supporting the hypothesis that AbcR2, its homolog AbcR1 and NfeR1 govern a large dense overlapping regulon for the control of nutrient uptake in response to different environmental signals. The regulatory activity of these sRNAs would contribute to coordinate nutrient uptake with the metabolic reprogramming concomitant to symbiotic transitions.

8. *Sm*YbeY does not serve an Hfq-like role as RNA stabilizer and matchmaker in riboregulation. Rather, YbeY acts as a silencing enzyme whose activity profoundly impacts on conserved fundamental chromosomally-encoded functions as well as on certain acquired plasmid-encoded *S. meliloti* pathways.

9. There is a discrete overlap between the Hfq and YbeY regulons. However,The comparison of the Hfq and YbeY RNA networks unveiled a number of putative substrates of this RNase. In particular, YbeY endoribonuclease activity could initiate decay of Hfq-binding transporter and nitrogen-fixation mRNAs upon their base-pairing interaction with the AbcR2 *trans*-sRNA and asRNAs, respectively.

BIBLIOGRAPHY

Afonyushkin, T., Vecerek, B., Moll, I., Blasi, U. and Kaberdin, V. R. (2005) 'Both RNase E and RNase III control the stability of *sodB* mRNA upon translational inhibition by the small regulatory RNA RyhB', *Nucleic Acids Res,* 33(5), pp. 1678-89.

Aiba, H. (2007) 'Mechanism of RNA silencing by Hfq-binding small RNAs', *Curr Opin Microbiol,* 10(2), pp. 134-9.

Akerley, B. J., Rubin, E. J., Novick, V. L., Amaya, K., Judson, N. and Mekalanos, J. J. (2002) 'A genome-scale analysis for identification of genes required for growth or survival of Haemophilus influenzae', *Proc Natl Acad Sci U S A,* 99(2), pp. 966-71.

Albaum, S. P., Neuweger, H., Franzel, B., Lange, S., Mertens, D., Trotschel, C., Wolters, D., Kalinowski, J., Nattkemper, T. W. and Goesmann, A. (2009) 'Qupe-a Rich Internet Application to take a step forward in the analysis of mass spectrometry-based quantitative proteomics experiments', *Bioinformatics,* 25(23), pp. 3128-34.

Altman, R. K., Schwope, I., Sarracino, D. A., Tetzlaff, C. N., Bleczinski, C. F. and Richert, C. (1999) 'Selection of modified oligonucleotides with increased target affinity via MALDI-monitored nuclease survival assays', *J Comb Chem,* 1(6), pp. 493-508.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990) 'Basic local alignment search tool', *J Mol Biol,* 215(3), pp. 403-10.

Ames, P. and Bergman, K. (1981) 'Competitive Advantage Provided by Bacterial Motility in the Formation of Nodules by *Rhizobium meliloti*', *J Bacteriol,* 148(2), pp. 728-9.

Argaman, L., Hershberg, R., Vogel, J., Bejerano, G., Wagner, E. G., Margalit, H. and Altuvia, S. (2001) 'Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*', *Curr Biol,* 11(12), pp. 941-50.

Armitage, J. P. (1992) 'Bacterial motility and chemotaxis', *Sci Prog,* 76(301-302 Pt 3-4), pp. 451-77.

Arraiano, C. M., Andrade, J. M., Domingues, S., Guinote, I. B., Malecki, M., Matos, R. G., Moreira, R. N., Pobre, V., Reis, F. P., Saramago, M., Silva, I. J. and Viegas, S. C. (2010) 'The critical role of RNA processing and degradation in the control of gene expression', *FEMS Microbiol Rev,* 34(5), pp. 883-923.

Arraiano, C. M. and Maquat, L. E. (2003) 'Post-transcriptional control of gene expression: effectors of mRNA decay', *Mol Microbiol,* 49(1), pp. 267-76.

Arraiano, C. M., Yancey, S. D. and Kushner, S. R. (1988) 'Stabilization of discrete mRNA breakdown products in *ams pnp rnb* multiple mutants of *Escherichia coli* K-12', *J Bacteriol,* 170(10), pp. 4625-33.

Babitzke, P., Baker, C. S. and Romeo, T. (2009) 'Regulation of translation initiation by RNA binding proteins', *Annu Rev Microbiol,* 63, pp. 27-44.

Babitzke, P., Granger, L., Olszewski, J. and Kushner, S. R. (1993) 'Analysis of mRNA decay and rRNA processing in *Escherichia coli* multiple mutants carrying a deletion in RNase III', *J Bacteriol,* 175(1), pp. 229-39.

Babitzke, P. and Romeo, T. (2007) 'CsrB sRNA family: sequestration of RNA-binding regulatory proteins', *Curr Opin Microbiol,* 10(2), pp. 156-63.

Bahlawane, C., McIntosh, M., Krol, E. and Becker, A. (2008) *'Sinorhizobium meliloti* regulator MucR couples exopolysaccharide synthesis and motility', *Mol Plant Microbe Interact,* 21(11), pp. 1498-509.

Bailey, T. L. and Elkan, C. (1994) 'Fitting a mixture model by expectation maximization to discover motifs in biopolymers', *Proc Int Conf Intell Syst Mol Biol,* 2, pp. 28-36.

Balbontin, R., Fiorini, F., Figueroa-Bossi, N., Casadesus, J. and Bossi, L. (2010) 'Recognition of heptameric seed sequence underlies multi-target regulation by RybB small RNA in *Salmonella enterica*', *Mol Microbiol,* 78(2), pp. 380-94.

Bandyra, K. J., Sinha, D., Syrjanen, J., Luisi, B. F. and De Lay, N. R. (2016) 'The ribonuclease polynucleotide phosphorylase can interact with small regulatory RNAs in both protective and degradative modes', *Rna,* 22(3), pp. 360-72.

Bar-Nahum, G. and Nudler, E. (2001) 'Isolation and characterization of sigma(70)-retaining transcription elongation complexes from *Escherichia coli*', *Cell,* 106(4), pp. 443-51.

Bardill, J. P. and Hammer, B. K. (2012) 'Non-coding sRNAs regulate virulence in the bacterial pathogen *Vibrio cholerae*', *RNA Biol,* 9(4), pp. 392-401.

Barnett, M. J., Bittner, A. N., Toman, C. J., Oke, V. and Long, S. R. (2012) 'Dual RpoH Sigma Factors and Transcriptional Plasticity in a Symbiotic Bacterium', *J*

Bacteriol, 194(18), pp. 4983-94.

Barnett, M. J., Fisher, R. F., Jones, T., Komp, C., Abola, A. P., Barloy-Hubler, F., Bowser, L., Capela, D., Galibert, F., Gouzy, J., Gurjal, M., Hong, A., Huizar, L., Hyman, R. W., Kahn, D., Kahn, M. L., Kalman, S., Keating, D. H., Palm, C., Peck, M. C., Surzycki, R., Wells, D. H., Yeh, K. C., Davis, R. W., Federspiel, N. A. and Long, S. R. (2001) 'Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti* pSymA megaplasmid', *Proc Natl Acad Sci U S A,* 98(17), pp. 9883-8.

Barra-Bily, L., Fontenelle, C., Jan, G., Flechard, M., Trautwetter, A., Pandey, S. P., Walker, G. C. and Blanco, C. (2010a) 'Proteomic alterations explain phenotypic changes in *Sinorhizobium meliloti* lacking the RNA chaperone Hfq', *J Bacteriol,* 192(6), pp. 1719-29.

Barra-Bily, L., Pandey, S. P., Trautwetter, A., Blanco, C. and Walker, G. C. (2010b) 'The *Sinorhizobium meliloti* RNA chaperone Hfq mediates symbiosis of *S. meliloti* and alfalfa', *J Bacteriol,* 192(6), pp. 1710-8.

Barrick, J. E., Sudarsan, N., Weinberg, Z., Ruzzo, W. L. and Breaker, R. R. (2005) '6S RNA is a widespread regulator of eubacterial RNA polymerase that resembles an open promoter', *Rna,* 11(5), pp. 774-84.

Batut, J., Andersson, S. G. and O'Callaghan, D. (2004) 'The evolution of chronic infection strategies in the alpha-proteobacteria', *Nat Rev Microbiol,* 2(12), pp. 933-45.

Baumgardt, K., Smidova, K., Rahn, H., Lochnit, G., Robledo, M. and Evguenieva-Hackenberg, E. (2016) 'The stress-related, rhizobial small RNA RcsR1 destabilizes the autoinducer synthase encoding mRNA sinI in *Sinorhizobium meliloti'*, *RNA Biol,* 13(5), pp. 486-99.

Becker, A., Barnett, M. J., Capela, D., Dondrup, M., Kamp, P. B., Krol, E., Linke, B., Ruberg, S., Runte, K., Schroeder, B. K., Weidner, S., Yurgel, S. N., Batut, J., Long, S. R., Puhler, A. and Goesmann, A. (2009) 'A portal for rhizobial genomes: RhizoGATE integrates a *Sinorhizobium meliloti* genome annotation update with postgenome data', *J Biotechnol,* 140(1-2), pp. 45-50.

Becker, A., Overloper, A., Schluter, J. P., Reinkensmeier, J., Robledo, M., Giegerich, R., Narberhaus, F. and Evguenieva-Hackenberg, E. (2014) 'Riboregulation in plant-associated alpha-proteobacteria', *RNA Biol,* 11(5), pp. 550-62.

Beckstette, M., Homann, R., Giegerich, R. and Kurtz, S. (2009) 'Significant speedup of database searches with HMMs by search space reduction with PSSM family models', *Bioinformatics,* 25(24), pp. 3251-8.

Beisel, C. L. and Storz, G. (2010) 'Base pairing small RNAs and their roles in global regulatory networks', *FEMS Microbiol Rev,* 34(5), pp. 866-82.

Berghoff, B. A., Glaeser, J., Sharma, C. M., Zobawa, M., Lottspeich, F., Vogel, J. and Klug, G. (2011) 'Contribution of Hfq to photooxidative stress resistance and global regulation in *Rhodobacter sphaeroides*', *Mol Microbiol,* 80(6), pp. 1479-95.

Beringer, J. E. (1974) 'R factor transfer in *Rhizobium leguminosarum*', *J Gen Microbiol,* 84(1), pp. 188-98.

Bernhart, S. H., Hofacker, I. L., Will, S., Gruber, A. R. and Stadler, P. F. (2008) 'RNAalifold: improved consensus structure prediction for RNA alignments', *BMC Bioinformatics,* 9, pp. 474.

Biase, F. H., Franco, M. M., Goulart, L. R. and Antunes, R. C. (2002) 'Protocol for extraction of genomic DNA from swine solid tissues', *Genet. Mol. Biol.,* 25(3), pp. 313-315.

Birnboim, H. C. and Doly, J. (1979) 'A rapid alkaline extraction procedure for screening recombinant plasmid DNA', *Nucleic Acids Res,* 7(6), pp. 1513-23.

Blanca-Ordóñez, H., Oliva-García, J. J., Pérez-Mendoza, D., Soto, M. J., Olivares, J., Sanjuán, J. and Nogales, J. (2010) 'pSymA-Dependent Mobilization of the *Sinorhizobium meliloti* pSymB Megaplasmid'.

Blatny, J. M., Brautaset, T., Winther-Larsen, H. C., Haugan, K. and Valla, S. (1997) 'Construction and use of a versatile set of broad-host-range cloning and expression vectors based on the RK2 replicon', *Appl Environ Microbiol,* 63(2), pp. 370-9.

Bobik, C., Meilhoc, E. and Batut, J. (2006) 'FixJ: a Major Regulator of the Oxygen Limitation Response and Late Symbiotic Functions of *Sinorhizobium meliloti*', *J. Bacteriol.*

Boisset, S., Geissmann, T., Huntzinger, E., Fechter, P., Bendridi, N., Possedko, M., Chevalier, C., Helfer, A. C., Benito, Y., Jacquier, A., Gaspin, C., Vandenesch, F. and Romby, P. (2007) *'Staphylococcus aureus* RNAIII

coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism', *Genes Dev,* 21(11), pp. 1353-66.

Boland, A., Tritschler, F., Heimstädt, S., Izaurralde, E. and Weichenrieder, O. (2010) 'Crystal structure and ligand binding of the MID domain of a eukaryotic Argonaute protein', *EMBO Rep: Vol. 7*, pp. 522-7.

Bradford, M. M. (1976) 'A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding', *Anal Biochem,* 72, pp. 248-54.

Brennan, R. G. and Link, T. M. (2007) 'Hfq structure, function and ligand binding', *Curr Opin Microbiol,* 10(2), pp. 125-33.

Brockwell, J. and Hely, F. (1966) 'Symbiotic characteristics of *Rhizobium meliloti*: and appraisal of the systematic treatment of nodulation and nitrogen fixation interactions between hosts and Rhizobia of diverse origins', *Australian Journal of Agricultural Research,* 17(6), pp. 885-899.

Brownlee, G. G. (1971) 'Sequence of 6S RNA of *E. coli*', *Nat New Biol,* 229(5), pp. 147-9.

Burge, S. W., Daub, J., Eberhardt, R., Tate, J., Barquist, L., Nawrocki, E. P., Eddy, S. R., Gardner, P. P. and Bateman, A. (2013) 'Rfam 11.0: 10 years of RNA families', *Nucleic Acids Res,* 41(Database issue), pp. D226-32.

Busch, A., Richter, A. S. and Backofen, R. (2008) 'IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions', *Bioinformatics,* 24(24), pp. 2849-56.

Capela, D., Barloy-Hubler, F., Gouzy, J., Bothe, G., Ampe, F., Batut, J., Boistard, P., Becker, A., Boutry, M., Cadieu, E., Dreano, S., Gloux, S., Godrie, T., Goffeau, A., Kahn, D., Kiss, E., Lelaure, V., Masuy, D., Pohl, T., Portetelle, D., Puhler, A., Purnelle, B., Ramsperger, U., Renard, C., Thebault, P., Vandenbol, M., Weidner, S. and Galibert, F. (2001) 'Analysis of the chromosome sequence of the legume symbiont *Sinorhizobium meliloti* strain 1021', *Proc Natl Acad Sci U S A,* 98(17), pp. 9877-82.

Carpousis, A. J. (2002) 'The *Escherichia coli* RNA degradosome: structure, function and relationship in other ribonucleolytic multienzyme complexes', *Biochem Soc Trans,* 30(2), pp. 150-5.

Carpousis, A. J., Luisi, B. F. and McDowall, K. J. (2009) 'Endonucleolytic initiation of mRNA decay in *Escherichia coli*', *Prog Mol Biol Transl Sci,* 85, pp. 91-135.

Carpousis, A. J., Van Houwe, G., Ehretsmann, C. and Krisch, H. M. (1994) 'Copurification of *E. coli* RNAase E and PNPase: evidence for a specific association between two enzymes important in RNA processing and degradation', *Cell,* 76(5), pp. 889-900.

Carter, R. J., Dubchak, I. and Holbrook, S. R. (2001) 'A computational approach to identify genes for functional RNAs in genomic sequences', *Nucleic Acids Res: Vol. 19*, pp. 3928-38.

Casse, F., Boucher, C., Julliot, J. S., Michel, M. and Dénarié, J. (1979) 'Identification and characterization of large plasmids in *Rhizobium meliloti* using agarose gel electrophoresis'.

Caswell, C. C., Gaines, J. M., Ciborowski, P., Smith, D., Borchers, C. H., Roux, C. M., Sayood, K., Dunman, P. M. and Roop Ii, R. M. (2012) 'Identification of two small regulatory RNAs linked to virulence in *Brucella abortus* 2308', *Mol Microbiol,* 85(2), pp. 345-60.

Caswell, C. C., Oglesby-Sherrouse, A. G. and Murphy, E. R. (2014) 'Sibling rivalry: related bacterial small RNAs and their redundant and non-redundant roles', *Front Cell Infect Microbiol,* 4.

Cavanagh, A. T., Klocko, A. D., Liu, X. and Wassarman, K. M. (2008) 'Promoter specificity for 6S RNA regulation of transcription is determined by core promoter sequences and competition for region 4.2 of σ70', *Mol Microbiol,* 67(6), pp. 1242-56.

Ceizel Borella, G., Lagares, A., Jr. and Valverde, C. (2016) 'Expression of the *Sinorhizobium meliloti* small RNA gene *mmgR* is controlled by the nitrogen source', *FEMS Microbiol Lett,* 363(9).

Chang, C., Damiani, I., Puppo, A. and Frendo, P. (2009) 'Redox changes during the legume-rhizobium symbiosis', *Mol Plant,* 2(3), pp. 370-7.

Chao, Y., Papenfort, K., Reinhardt, R., Sharma, C. M. and Vogel, J. (2012) 'An atlas of Hfq-bound transcripts reveals 3′ UTRs as a genomic reservoir of regulatory small RNAs', *EMBO J,* 31(20), pp. 4005-19.

Charoenpanich, P., Meyer, S., Becker, A. and McIntosh, M. (2013) 'Temporal Expression Program of Quorum Sensing-Based Transcription Regulation in *Sinorhizobium meliloti*', J *Bacteriol: Vol. 14*, pp. 3224-36.

Church, G. M. and Gilbert, W. (1984) 'Genomic sequencing', *Proc Natl Acad Sci U S A*, 81(7), pp. 1991-5.

Condon, C., Squires, C. and Squires, C. L. (1995) 'Control of rRNA transcription in *Escherichia coli*', *Microbiol Rev,* 59(4), pp. 623-45.

Conesa, A., Gotz, S., Garcia-Gomez, J. M., Terol, J., Talon, M. and Robles, M. (2005) 'Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research', *Bioinformatics*, 21(18), pp. 3674-6.

Crick, F. (1970) 'Central dogma of molecular biology', *Nature,* 227(5258), pp. 561-3.

Dambach, M., Irnov, I. and Winkler, W. C. (2013) 'Association of RNAs with *Bacillus subtilis* Hfq', *PLoS One,* 8(2), pp. e55156.

Davies, B. W., Kohrer, C., Jacob, A. I., Simmons, L. A., Zhu, J., Aleman, L. M., Rajbhandary, U. L. and Walker, G. C. (2010) 'Role of *Escherichia coli* YbeY, a highly conserved protein, in rRNA processing', *Mol Microbiol,* 78(2), pp. 506-18.

Davies, B. W. and Walker, G. C. (2007) 'Identification of Novel *Sinorhizobium meliloti* Mutants Compromised for Oxidative Stress Protection and Symbiosis▽', *J Bacteriol: Vol. 5*, pp. 2110-3.

Davies, B. W. and Walker, G. C. (2008) 'A highly conserved protein of unknown function is required by *Sinorhizobium meliloti* for symbiosis and environmental stress protection', *J Bacteriol,* 190(3), pp. 1118-23.

De Lay, N., Schu, D. J. and Gottesman, S. (2013) 'Bacterial Small RNA-based Negative Regulation: Hfq and Its Accomplices', J *Biol Chem: Vol. 12*, pp. 7996-8003.

del Val, C., Ernst, P., Falkenhahn, M., Fladerer, C., Glatting, K. H., Suhai, S. and Hotz-Wagenblatt, A. (2007) 'ProtSweep, 2Dsweep and DomainSweep: protein analysis suite at DKFZ', *Nucleic Acids Res,* 35(Web Server issue), pp. W444-50.

del Val, C., Romero-Zaliz, R., Torres-Quesada, O., Peregrina, A., Toro, N. and

Jimenez-Zurdo, J. I. (2012) 'A survey of sRNA families in alpha-proteobacteria', *RNA Biol,* 9(2), pp. 119-29.

Ding, Y., Kalo, P., Yendrek, C., Sun, J., Liang, Y., Marsh, J. F., Harris, J. M. and Oldroyd, G. E. (2008) 'Abscisic Acid Coordinates Nod Factor and Cytokinin Signaling during the Regulation of Nodulation in *Medicago truncatula*', *Plant Cell,* 20(10), pp. 2681-95.

Ditta, G., Stanfield, S., Corbin, D. and Helinski, D. R. (1980) 'Broad host range DNA cloning system for gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*', *Proc Natl Acad Sci U S A,* 77(12), pp. 7347-51.

Dominguez-Ferreras, A., Munoz, S., Olivares, J., Soto, M. J. and Sanjuan, J. (2009) 'Role of potassium uptake systems in *Sinorhizobium meliloti* osmoadaptation and symbiotic performance', *J Bacteriol,* 191(7), pp. 2133-43.

Domonkos, A., Horvath, B., Marsh, J. F., Halasz, G., Ayaydin, F., Oldroyd, G. E. and Kalo, P. (2013) 'The identification of novel loci required for appropriate nodule development in *Medicago truncatula'*, *BMC Plant Biol,* 13, pp. 157.

Domínguez-Ferreras, A., Pérez-Arnedo, R., Becker, A., Olivares, J., Soto, M. J. and Sanjuán, J. (2006) 'Transcriptome Profiling Reveals the Importance of Plasmid pSymB for Osmoadaptation of *Sinorhizobium meliloti*'.

Dondrup, M., Albaum, S. P., Griebel, T., Henckel, K., Junemann, S., Kahlke, T., Kleindt, C. K., Kuster, H., Linke, B., Mertens, D., Mittard-Runte, V., Neuweger, H., Runte, K. J., Tauch, A., Tille, F., Puhler, A. and Goesmann, A. (2009) 'EMMA 2--a MAGE-compliant system for the collaborative analysis and integration of microarray data', *BMC Bioinformatics,* 10, pp. 50.

Duan, X., Young, R., Straubinger, R. M., Page, B., Cao, J., Wang, H., Yu, H., Canty, J. M. and Qu, J. (2009) 'A straightforward and highly efficient precipitation/on-pellet digestion procedure coupled with a long gradient nano-LC separation and Orbitrap mass spectrometry for label-free expression profiling of the swine heart mitochondrial proteome', *J Proteome Res,* 8(6), pp. 2838-50.

Casida, L.E. (1982) *'Ensifer adhaerens* gen. nov., sp. nov.: A Bacterial Predator of Bacteria in Soil', *Microbiology Society*.

Ebeling, S., Kundig, C. and Hennecke, H. (1991) 'Discovery of a rhizobial RNA that is essential for symbiotic root nodule development', *J Bacteriol,* 173(20), pp. 6373-82.

Eddy, S. R. (2001) 'Non-coding RNA genes and the modern RNA world', *Nat Rev Genet,* 2(12), pp. 919-29.

Evguenieva-Hackenberg, E. and Klug, G. (2000) 'RNase III Processing of Intervening Sequences Found in Helix 9 of 23S rRNA in the Alpha Subclass of Proteobacteria', *J Bacteriol: Vol. 17*, pp. 4719-29.

Fan, Y., Thompson, J. W., Dubois, L. G., Moseley, M. A. and Wernegreen, J. J. (2013) 'Proteomic analysis of an unculturable bacterial endosymbiont (Blochmannia) reveals high abundance of chaperonins and biosynthetic enzymes', *J Proteome Res,* 12(2), pp. 704-18.

Finan, T. M., Weidner, S., Wong, K., Buhrmester, J., Chain, P., Vorholter, F. J., Hernandez-Lucas, I., Becker, A., Cowie, A., Gouzy, J., Golding, B. and Puhler, A. (2001) 'The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endosymbiont *Sinorhizobium meliloti'*, *Proc Natl Acad Sci U S A,* 98(17), pp. 9889-94.

Finan, T. M., Wood, J. M. and Jordan, D. C. (1983) 'Symbiotic properties of C4-dicarboxylic acid transport mutants of *Rhizobium leguminosarum'*, *J Bacteriol,* 154(3), pp. 1403-13.

Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Heger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E. L., Tate, J. and Punta, M. (2014) 'Pfam: the protein families database', *Nucleic Acids Res,* 42(Database issue), pp. D222-30.

Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E., Gavin, O. L., Gunasekaran, P., Ceric, G., Forslund, K., Holm, L., Sonnhammer, E. L., Eddy, S. R. and Bateman, A. (2010) 'The Pfam protein families database', *Nucleic Acids Res,* 38(Database issue), pp. D211-22.

Fischer, H. M. (1994) 'Genetic regulation of nitrogen fixation in rhizobia', *Microbiol Rev,* 58(3), pp. 352-86.

Flechard, M., Fontenelle, C., Blanco, C., Goude, R., Ermel, G. and Trautwetter, A. (2010) 'RpoE2 of *Sinorhizobium meliloti* is necessary for trehalose synthesis and growth in hyperosmotic media', *Microbiology,* 156(Pt 6), pp. 1708-18.

Folichon, M., Allemand, F., Regnier, P. and Hajnsdorf, E. (2005) 'Stimulation of poly(A) synthesis by *Escherichia coli* poly(A)polymerase I is correlated with Hfq binding to poly(A) tails', *Febs j,* 272(2), pp. 454-63.

Frank, D. N. and Pace, N. R. (1998) 'Ribonuclease P: unity and diversity in a tRNA processing ribozyme', *Annu Rev Biochem,* 67, pp. 153-80.

Franze de Fernandez, M. T., Eoyang, L. and August, J. T. (1968) 'Factor fraction required for the synthesis of bacteriophage Qbeta-RNA', *Nature,* 219(5154), pp. 588-90.

Franze de Fernandez, M. T., Hayward, W. S. and August, J. T. (1972) 'Bacterial proteins required for replication of phage Q ribonucleic acid. Purification and properties of host factor I, a ribonucleic acid-binding protein', *J Biol Chem,* 247(3), pp. 824-31.

Galibert, F., Finan, T. M., Long, S. R., Puhler, A., Abola, P., Ampe, F., Barloy-Hubler, F., Barnett, M. J., Becker, A., Boistard, P., Bothe, G., Boutry, M., Bowser, L., Buhrmester, J., Cadieu, E., Capela, D., Chain, P., Cowie, A., Davis, R. W., Dreano, S., Federspiel, N. A., Fisher, R. F., Gloux, S., Godrie, T., Goffeau, A., Golding, B., Gouzy, J., Gurjal, M., Hernandez-Lucas, I., Hong, A., Huizar, L., Hyman, R. W., Jones, T., Kahn, D., Kahn, M. L., Kalman, S., Keating, D. H., Kiss, E., Komp, C., Lelaure, V., Masuy, D., Palm, C., Peck, M. C., Pohl, T. M., Portetelle, D., Purnelle, B., Ramsperger, U., Surzycki, R., Thebault, P., Vandenbol, M., Vorholter, F. J., Weidner, S., Wells, D. H., Wong, K., Yeh, K. C. and Batut, J. (2001) 'The composite genome of the legume symbiont *Sinorhizobium meliloti'*, *Science,* 293(5530), pp. 668-72.

Gama-Castro, S., Salgado, H., Peralta-Gil, M., Santos-Zavaleta, A., Muñiz-Rascado, L., Solano-Lira, H., Jimenez-Jacinto, V., Weiss, V., García-Sotelo, J. S., López-Fuentes, A., Porrón-Sotelo, L., Alquicira-Hernández, S., Medina-Rivera, A., Martínez-Flores, I., Alquicira-Hernández, K., Martínez-Adame, R., Bonavides-Martínez, C., Miranda-Ríos, J., Huerta, A. M., Mendoza-Vargas, A., Collado-Torres, L., Taboada, B., Vega-Alvarado, L., Olvera, M., Olvera, L., Grande, R., Morett, E. and Collado-Vides, J. (2011) 'RegulonDB version 7.0: transcriptional regulation of *Escherichia coli* K-12 integrated within genetic sensory response units (Gensor Units)', *Nucleic Acids Res: Vol. Database issue*, pp. D98-D105.

Gan, J., Tropea, J. E., Austin, B. P., Court, D. L., Waugh, D. S. and Ji, X. (2006) 'Structural insight into the mechanism of double-stranded RNA processing by ribonuclease III', *Cell,* 124(2), pp. 355-66.

Gao, M., Barnett, M. J., Long, S. R. and Teplitski, M. (2010) 'Role of the

Sinorhizobium meliloti global regulator Hfq in gene regulation and symbiosis', *Mol Plant Microbe Interact,* 23(4), pp. 355-65.

Geissmann, T. A. and Touati, D. (2004) 'Hfq, a new chaperoning role: binding to messenger RNA determines access for small RNA regulator', *Embo j,* 23(2), pp. 396-405.

Georg, J. and Hess, W. R. (2011) 'cis-antisense RNA, another level of gene regulation in bacteria', *Microbiol Mol Biol Rev,* 75(2), pp. 286-300.

Giacomini, A., Ollero, F. J., Squartini, A. and Nuti, M. P. (1994) 'Construction of multipurpose gene cartridges based on a novel synthetic promoter for high-level gene expression in gram-negative bacteria', *Gene,* 144(1), pp. 17-24.

Gil, R., Silva, F. J., Pereto, J. and Moya, A. (2004) 'Determination of the core of a minimal bacterial gene set', *Microbiol Mol Biol Rev,* 68(3), pp. 518-37, table of contents.

Giller, G., Tasara, T., Angerer, B., Muhlegger, K., Amacker, M. and Winter, H. (2003) 'Incorporation of reporter molecule-labeled nucleotides by DNA polymerases. I. Chemical synthesis of various reporter group-labeled 2'-deoxyribonucleoside-5'-triphosphates', *Nucleic Acids Res,* 31(10), pp. 2630-5.

Gottesman, S. (2004) 'The small RNA regulators of *Escherichia coli*: roles and mechanisms*', *Annu Rev Microbiol,* 58, pp. 303-28.

Gottesman, S. and Storz, G. (2011) 'Bacterial small RNA regulators: versatile roles and rapidly evolving variations', *Cold Spring Harb Perspect Biol,* 3(12).

Grinwald, M. and Ron, E. Z. (2013) 'The *Escherichia coli* translation-associated heat shock protein YbeY is involved in rRNA transcription antitermination', *PLoS One,* 8(4), pp. e62297.

Gross, G. and Dunn, J. J. (1987) 'Structure of secondary cleavage sites of *E. coli* RNAaseIII in A3t RNA from bacteriophage T7', *Nucleic Acids Res,* 15(2), pp. 431-42.

Gruber, A. R., Lorenz, R., Bernhart, S. H., Neubock, R. and Hofacker, I. L. (2008) 'The Vienna RNA websuite', *Nucleic Acids Res,* 36(Web Server issue), pp. W70-4.

Gruber, T. M. and Gross, C. A. (2003) 'Multiple sigma subunits and the partitioning

of bacterial transcription space', *Annu Rev Microbiol,* 57, pp. 441-66.

Gulash, M., Ames, P., Larosiliere, R. C. and Bergman, K. (1984) 'Rhizobia are attracted to localized sites on legume roots', *Appl Environ Microbiol,* 48(1), pp. 149-52.

Haakonsen, D. L., Yuan, A. H. and Laub, M. T. (2015) 'The bacterial cell cycle regulator GcrA is a σ70 cofactor that drives gene expression from a subset of methylated promoters', *Genes Dev,* 29(21), pp. 2272-86.

Harfouche, L., Haichar Fel, Z. and Achouak, W. (2015) 'Small regulatory RNAs and the fine-tuning of plant-bacteria interactions', *New Phytol,* 206(1), pp. 98-106.

Haseth, P. L. D. and Uhlenbeck, O. C. (2002) 'Interaction of *Escherichia coli* host factor protein with oligoriboadenylates'.

Herridge, D. F., Peoples, M. B. and Boddey, R. M. (2008) 'Global inputs of biological nitrogen fixation in agricultural systems', *Plant and soil,* 311(1), pp. 1-18.

Heuer, H., Krsek, M., Baker, P., Smalla, K. and Wellington, E. M. (1997) 'Analysis of actinomycete communities by specific amplification of genes encoding 16S rRNA and gel-electrophoretic separation in denaturing gradients', *Appl Environ Microbiol,* 63(8), pp. 3233-41.

Hilker, R., Stadermann, K. B., Doppmeier, D., Kalinowski, J., Stoye, J., Straube, J., Winnebald, J. and Goesmann, A. (2014) 'ReadXplorer--visualization and analysis of mapped sequences', *Bioinformatics,* 30(16), pp. 2247-54.

Hindley, J. (1967) 'Fractionation of 32P-labelled ribonucleic acids on polyacrylamide gels and their characterization by fingerprinting', *J Mol Biol,* 30(1), pp. 125-36.

Hoffmann, T., Wensing, A., Brosius, M., Steil, L., Volker, U. and Bremer, E. (2013) 'Osmotic control of *opuA* expression in *Bacillus subtilis* and its modulation in response to intracellular glycine betaine and proline pools', *J Bacteriol,* 195(3), pp. 510-22.

Hofmann, N., Wurm, R. and Wagner, R. (2011) 'The *E. coli* Anti-Sigma Factor Rsd: Studies on the Specificity and Regulation of Its Expression', *PLoS One,* 6(5).

Hori, K. and Yanazaki, Y. (1974) 'Nucleotide sequence specific interaction of host factor I with bacteriophage Q beta RNA', *FEBS Lett,* 43(1), pp. 20-2.

Hunter, S., Apweiler, R., Attwood, T. K., Bairoch, A., Bateman, A., Binns, D., Bork, P., Das, U., Daugherty, L., Duquenne, L., Finn, R. D., Gough, J., Haft, D., Hulo, N., Kahn, D., Kelly, E., Laugraud, A., Letunic, I., Lonsdale, D., Lopez, R., Madera, M., Maslen, J., McAnulla, C., McDowall, J., Mistry, J., Mitchell, A., Mulder, N., Natale, D., Orengo, C., Quinn, A. F., Selengut, J. D., Sigrist, C. J., Thimma, M., Thomas, P. D., Valentin, F., Wilson, D., Wu, C. H. and Yeats, C. (2009) 'InterPro: the integrative protein signature database', *Nucleic Acids Res,* 37(Database issue), pp. D211-5.

Huntzinger, E., Boisset, S., Saveanu, C., Benito, Y., Geissmann, T., Namane, A., Lina, G., Etienne, J., Ehresmann, B., Ehresmann, C., Jacquier, A., Vandenesch, F. and Romby, P. (2005) *'Staphylococcus aureus* RNAIII and the endoribonuclease III coordinately regulate *spa* gene expression', *Embo j,* 24(4), pp. 824-35.

Huttenhofer, A. (2006) 'RNomics: identification and function of small non-protein-coding RNAs in model organisms', *Cold Spring Harb Symp Quant Biol,* 71, pp. 135-40.

I. L. Hofacker, W. Fontana, P. F. Stadler, L. S. Bonhoeffer, Tacker, M. and Schuster, P. (1994) 'Fast folding and comparison of RNA secondary structures', *Monatsh Chem,* 125(2), pp. 167–188.

Ibragimova, M. V., Rumyantseva, M. L., Onishchuk, O. P., Belova, V. S., Kurchak, O. N., Andronov, E. E., Dzyubenko, N. I. and Simarov, B. V. (2006) 'Symbiosis Between the Root-nodule Bacterium *Sinorhizobium meliloti* and Alfalfa (medicago sativa) Under Salinization Conditions', *Microbiology,* 75(1), pp. 77-81.

Ish-Horowicz, D. and Burke, J. F. (1981) 'Rapid and efficient cosmid cloning', *Nucleic Acids Res,* 9(13), pp. 2989-98.

Jacob, A. I., Kohrer, C., Davies, B. W., RajBhandary, U. L. and Walker, G. C. (2013) 'Conserved bacterial RNase YbeY plays key roles in 70S ribosome quality control and 16S rRNA maturation', *Mol Cell,* 49(3), pp. 427-38.

Jebbar, M., Sohn-Bösser, L., Bremer, E., Bernard, T. and Blanco, C. (2005) 'Ectoine-Induced Proteins in *Sinorhizobium meliloti* Include an Ectoine ABC-Type Transporter Involved in Osmoprotection and Ectoine Catabolism', *Journal of Bacteriology,* 187(4), pp. 1293-1304.

Jimenez-Zurdo, J. I., Valverde, C. and Becker, A. (2013) 'Insights into the

noncoding RNome of nitrogen-fixing endosymbiotic alpha-proteobacteria', *Mol Plant Microbe Interact,* 26(2), pp. 160-7.

Jones, K. M., Kobayashi, H., Davies, B. W., Taga, M. E. and Walker, G. C. (2007) 'How rhizobial symbionts invade plants: the *Sinorhizobium-Medicago* model', *Nat Rev Microbiol,* 5(8), pp. 619-33.

Kawano, M., Reynolds, A. A., Miranda-Rios, J. and Storz, G. (2005) 'Detection of 5′- and 3′-UTR-derived small RNAs and cis-encoded antisense RNAs in Escherichia coli', *Nucleic Acids Res,* 33(3), pp. 1040-50.

Keiler, K. C., Waller, P. R. and Sauer, R. T. (1996) 'Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA', *Science,* 271(5251), pp. 990-3.

Khan, S. R., Gaines, J., Roop, R. M., 2nd and Farrand, S. K. (2008) 'Broad-host-range expression vectors with tightly regulated promoters and their use to examine the influence of TraR and TraM expression on Ti plasmid quorum sensing', *Appl Environ Microbiol,* 74(16), pp. 5053-62.

Klocko, A. D. and Wassarman, K. M. (2009) '6S RNA binding to Esigma(70) requires a positively charged surface of sigma(70) region 4.2', *Mol Microbiol,* 73(2), pp. 152-64.

Kobayashi, K., Ehrlich, S. D., Albertini, A., Amati, G., Andersen, K. K., Arnaud, M., Asai, K., Ashikaga, S., Aymerich, S., Bessieres, P., Boland, F., Brignell, S. C., Bron, S., Bunai, K., Chapuis, J., Christiansen, L. C., Danchin, A., Débarbouillé, M., Dervyn, E., Deuerling, E., Devine, K., Devine, S. K., Dreesen, O., Errington, J., Fillinger, S., Foster, S. J., Fujita, Y., Galizzi, A., Gardan, R., Eschevins, C., Fukushima, T., Haga, K., Harwood, C. R., Hecker, M., Hosoya, D., Hullo, M. F., Kakeshita, H., Karamata, D., Kasahara, Y., Kawamura, F., Koga, K., Koski, P., Kuwana, R., Imamura, D., Ishimaru, M., Ishikawa, S., Ishio, I., Coq, D. L., Masson, A., Mauël, C., Meima, R., Mellado, R. P., Moir, A., Moriya, S., Nagakawa, E., Nanamiya, H., Nakai, S., Nygaard, P., Ogura, M., Ohanan, T., O'Reilly, M., O'Rourke, M., Pragai, Z., Pooley, H. M., Rapoport, G., Rawlins, J. P., Rivas, L. A., Rivolta, C., Sadaie, A., Sadaie, Y., Sarvas, M., Sato, T., Saxild, H. H., Scanlan, E., Schumann, W., Seegers, J. F. M. L., Sekiguchi, J., Sekowska, A., Séror, S. J., Simon, M., Stragier, P., Studer, R., Takamatsu, H., Tanaka, T., Takeuchi, M., Thomaides, H. B., Vagner, V., Dijl, J. M. v., Watabe, K., Wipat, A., Yamamoto, H., Yamamoto, M., Yamamoto, Y., Yamane, K., Yata,

K., Yoshida, K., Yoshikawa, H., Zuber, U. and Ogasawara, N. (2003) 'Essential *Bacillus subtilis* genes', *Proc.Natl.Acad.Aci.USA*.

Kovach, M. E., Phillips, R. W., Elzer, P. H., Roop, R. M., 2nd and Peterson, K. M. (1994) 'pBBR1MCS: a broad-host-range cloning vector', *Biotechniques,* 16(5), pp. 800-2.

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

Kuster, H., Becker, A., Firnhaber, C., Hohnjec, N., Manthey, K., Perlick, A. M., Bekel, T., Dondrup, M., Henckel, K., Goesmann, A., Meyer, F., Wipf, D., Requena, N., Hildebrandt, U., Hampp, R., Nehls, U., Krajinski, F., Franken, P. and Puhler, A. (2007) 'Development of bioinformatic tools to support EST-sequencing, in silico- and microarray-based transcriptome profiling in mycorrhizal symbioses', *Phytochemistry,* 68(1), pp. 19-32.

Laane, C., Haaker, H. and Veeger, C. (1978) 'Involvement of the cytoplasmic membrane in nitrogen fixation by *Rhizobium leguminosarum* bacteroids', *Eur J Biochem,* 87(1), pp. 147-53.

Laemmli, U. K. (1970) 'Cleavage of structural proteins during the assembly of the head of bacteriophage T4', *Nature,* 227(5259), pp. 680-5.

Lagares, A., Jr., Ceizel Borella, G., Linne, U., Becker, A. and Valverde, C. (2017) 'Regulation of polyhydroxybutyrate accumulation in *Sinorhizobium meliloti* by the trans-encoded small RNA MmgR', *J Bacteriol*.

Lamontagne, B. and Elela, S. A. (2004) 'Evaluation of the RNA determinants for bacterial and yeast RNase III binding and cleavage', *J Biol Chem,* 279(3), pp. 2231-41.

Lang, C. and Long, S. R. (2015) 'Transcriptomic Analysis of *Sinorhizobium meliloti* and *Medicago truncatula* Symbiosis Using Nitrogen Fixation-Deficient Nodules', *Mol Plant Microbe Interact,* 28(8), pp. 856-68.

Langmead, B. and Salzberg, S. L. (2012) 'Fast gapped-read alignment with Bowtie 2', *Nature Methods,* 9, pp. 357-359.

Lasa, I., Toledo-Arana, A., Dobin, A., Villanueva, M., de los Mozos, I. R., Vergara-Irigaray, M., Segura, V., Fagegaltier, D., Penades, J. R., Valle, J.,

Solano, C. and Gingeras, T. R. (2011) 'Genome-wide antisense transcription drives mRNA processing in bacteria', *Proc Natl Acad Sci U S A*, 108(50), pp. 20172-7.

Lelandais-Briere, C., Moreau, J., Hartmann, C. and Crespi, M. (2016) 'Noncoding RNAs, Emerging Regulators in Root Endosymbioses', *Mol Plant Microbe Interact*, 29(3), pp. 170-80.

Lenz, D. H., Mok, K. C., Lilley, B. N., Kulkarni, R. V., Wingreen, N. S. and Bassler, B. L. (2004) 'The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in *Vibrio harveyi* and *Vibrio cholerae*', *Cell*, 118(1), pp. 69-82.

Leskinen, K., Varjosalo, M. and Skurnik, M. (2015) 'Absence of YbeY RNase compromises the growth and enhances the virulence plasmid gene expression of *Yersinia enterocolitica* O: 3', *Microbiology*, 161(Pt 2), pp. 285-99.

Letunic, I., Doerks, T. and Bork, P. (2009) 'SMART 6: recent updates and new developments', *Nucleic Acids Res*, 37(Database issue), pp. D229-32.

Li, H. L., Chelladurai, B. S., Zhang, K. and Nicholson, A. W. (1993) 'Ribonuclease III cleavage of a bacteriophage T7 processing signal. Divalent cation specificity, and specific anion effects', *Nucleic Acids Res*, 21(8), pp. 1919-25.

Limpens, E., Mirabella, R., Fedorova, E., Franken, C., Franssen, H., Bisseling, T. and Geurts, R. (2005) 'Formation of organelle-like N2-fixing symbiosomes in legume root nodules is controlled by DMI2', *Proc Natl Acad Sci U S A*, 102(29), pp. 10375-80.

Lindell, M., Romby, P. and Wagner, E. G. H. (2002) 'Lead(II) as a probe for investigating RNA structure in vivo', *RNA*, 8(4), pp. 534-41.

Lindstrom, K. and Martinez-Romero, M. E. (2002) 'International Committee on Systematics of Prokaryotes: Subcommittee on the taxonomy of *Agrobacterium* and *Rhizobium*'.

Lodwig, E. M., Hosie, A. H., Bourdes, A., Findlay, K., Allaway, D., Karunakaran, R., Downie, J. A. and Poole, P. S. (2003) 'Amino-acid cycling drives nitrogen fixation in the legume-*Rhizobium* symbiosis', *Nature*, 422(6933), pp. 722-6.

Loh, E., Dussurget, O., Gripenland, J., Vaitkevicius, K., Tiensuu, T., Mandin, P., Repoila, F., Buchrieser, C., Cossart, P. and Johansson, J. (2009) 'A trans-

acting riboswitch controls expression of the virulence regulator PrfA in *Listeria monocytogenes*', *Cell,* 139(4), pp. 770-9.

Long, S. R. (1996) '*Rhizobium* symbiosis: nod factors in perspective', *Plant Cell,* 8(10), pp. 1885-98.

Love, M. I., Huber, W. and Anders, S. (2014) 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2', *Genome Biol,* 15(12), pp. 550.

MacCoss, M. J., Wu, C. C., Liu, H., Sadygov, R. and Yates, J. R., (2003) 'A correlation algorithm for the automated quantitative analysis of shotgun proteomics data', *Anal Chem,* 75(24), pp. 6912-21.

MacLellan, S. R., MacLean, A. M. and Finan, T. M. (2006) 'Promoter prediction in the rhizobia', *Microbiology,* 152(Pt 6), pp. 1751-63.

Majdalani, N., Vanderpool, C. K. and Gottesman, S. (2005) 'Bacterial small RNA regulators', *Crit Rev Biochem Mol Biol,* 40(2), pp. 93-113.

Maki, K., Uno, K., Morita, T. and Aiba, H. (2008) 'RNA, but not protein partners, is directly responsible for translational silencing by a bacterial Hfq-binding small RNA', *Proc Natl Acad Sci U S A,* 105(30), pp. 10332-7.

Masse, E., Escorcia, F. E. and Gottesman, S. (2003) 'Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli*', *Genes Dev,* 17(19), pp. 2374-83.

Masse, E. and Gottesman, S. (2002) 'A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*', *Proc Natl Acad Sci U S A,* 99(7), pp. 4620-5.

Mauchline, T. H., Fowler, J. E., East, A. K., Sartor, A. L., Zaheer, R., Hosie, A. H. F., Poole, P. S. and Finan, T. M. (2006a) 'Mapping the *Sinorhizobium meliloti* 1021 solute-binding protein-dependent transportome', 103(47), pp. 17933-17938.

Meade, H. M., Long, S. R., Ruvkun, G. B., Brown, S. E. and Ausubel, F. M. (1982) 'Physical and genetic characterization of symbiotic and auxotrophic mutants of *Rhizobium meliloti* induced by transposon Tn5 mutagenesis', *J Bacteriol,* 149(1), pp. 114-22.

Meade, H. M. and Signer, E. R. (1977) 'Genetic mapping of *Rhizobium meliloti*',

Proc Natl Acad Sci U S A, 74(5), pp. 2076-8.

Mitsui, H., Sato, T., Sato, Y., Ito, N. and Minamisawa, K. (2004) *'Sinorhizobium meliloti* RpoH1 is required for effective nitrogen-fixing symbiosis with alfalfa', *Mol Genet Genomics,* 271(4), pp. 416-25.

Moawad, H. and Beck, D. P. (1991) 'Some characteristics of *Rhizobium leguminosarum* isolates from uninoculated field-grown lentil', *Soil biology and biochemistry,* 23(10), pp. 933-937.

Mohanty, B. K. and Kushner, S. R. (2003) 'Genomic analysis in *Escherichia coli* demonstrates differential roles for polynucleotide phosphorylase and RNase II in mRNA abundance and decay', *Mol Microbiol,* 50(2), pp. 645-58.

Moll, I., Afonyushkin, T., Vytvytska, O., Kaberdin, V. R. and Blasi, U. (2003) 'Coincident Hfq binding and RNase E cleavage sites on mRNA and small regulatory RNAs', *Rna,* 9(11), pp. 1308-14.

Moore, S. D. and Sauer, R. T. (2007) 'The tmRNA system for translational surveillance and ribosome rescue', *Annu Rev Biochem,* 76, pp. 101-24.

Morita, T., Maki, K. and Aiba, H. (2005) 'RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs', Genes *Dev: Vol. 18*, pp. 2176-86.

Morita, T., Mochizuki, Y. and Aiba, H. (2006) 'Translational repression is sufficient for gene silencing by bacterial small noncoding RNAs in the absence of mRNA destruction', *Proc Natl Acad Sci U S A,* 103(13), pp. 4858-63.

Møller, T., Franch, T., Udesen, C., Gerdes, K. and Valentin-Hansen, P. (2002) 'Spot 42 RNA mediates discoordinate expression of the *E. coli* galactose operon', *Genes Dev,* 16(13), pp. 1696-706.

Nawrocki, E. P. and Eddy, S. R. (2007) 'Query-dependent banding (QDB) for faster RNA similarity searches', *PLoS Comput Biol,* 3(3), pp. e56.

Nawrocki, E. P., Kolbe, D. L. and Eddy, S. R. (2009) 'Infernal 1.0: inference of RNA alignments', *Bioinformatics,* 25(10), pp. 1335-7.

Nicholson, A. W. (1999) 'Function, mechanism and regulation of bacterial ribonucleases', *FEMS Microbiol Rev,* 23(3), pp. 371-90.

Nicholson, A. W. (2014) 'Ribonuclease III mechanisms of double-stranded RNA

cleavage', *Wiley Interdiscip Rev RNA,* 5(1), pp. 31-48.

Nogales, J., Bernabéu-Roda, L., Cuéllar, V. and Soto, M. J. (2012) 'ExpR Is Not Required for Swarming but Promotes Sliding in *Sinorhizobium meliloti'*, J *Bacteriol: Vol. 8*, pp. 2027-35.

Novichkov, P. S., Rodionov, D. A., Stavrovskaya, E. D., Novichkova, E. S., Kazakov, A. E., Gelfand, M. S., Arkin, A. P., Mironov, A. A. and Dubchak, I. (2010) 'RegPredict: an integrated system for regulon inference in prokaryotes by comparative genomics approach', *Nucleic Acids Res,* 38(Web Server issue), pp. W299-307.

Oganesyan, V., Busso, D., Brandsen, J., Chen, S., Jancarik, J., Kim, R. and Kim, S. H. (2003) 'Structure of the hypothetical protein AQ_1354 from *Aquifex aeolicus*', *Acta Crystallogr D Biol Crystallogr,* 59(Pt 7), pp. 1219-23.

Ohwada, T., Sasaki, Y., Koike, H., Igawa, K. and Sato, T. (2014) 'Correlation between NaCl Sensitivity of *Rhizobium* Bacteria and Ineffective Nodulation of Leguminous Plants', 62(11), pp. 2086-2090.

Ohyama, H., Sakai, T., Agari, Y., Fukui, K., Nakagawa, N., Shinkai, A., Masui, R. and Kuramitsu, S. (2014) 'The role of ribonucleases in regulating global mRNA levels in the model organism *Thermus thermophilus* HB8', *BMC Genomics*, pp. 386.

Oke, V., Rushing, B. G., Fisher, E. J., Moghadam-Tabrizi, M. and Long, S. R. (2001) 'Identification of the heat-shock sigma factor RpoH and a second RpoH-like protein in *Sinorhizobium meliloti'*, *Microbiology,* 147(Pt 9), pp. 2399-408.

Oldroyd, G. E. and Downie, J. A. (2008) 'Coordinating nodule morphogenesis with rhizobial infection in legumes', *Annu Rev Plant Biol*, 59, pp. 519-46.

Oldroyd, G. E., Murray, J. D., Poole, P. S. and Downie, J. A. (2011) 'The rules of engagement in the legume-rhizobial symbiosis', *Annu Rev Genet,* 45, pp. 119-44.

Oliva, G., Sahr, T. and Buchrieser, C. (2015) 'Small RNAs, 5' UTR elements and RNA-binding proteins in intracellular bacteria: impact on metabolism and virulence', *FEMS Microbiol Rev,* 39(3), pp. 331-49.

Ono, Y., Mitsui, H., Sato, T. and Minamisawa, K. (2001) 'Two RpoH homologs responsible for the expression of heat shock protein genes in *Sinorhizobium meliloti*', *Mol Gen Genet,* 264(6), pp. 902-12.

Oono, R., Anderson, C. G. and Denison, R. F. (2011) 'Failure to fix nitrogen by non-reproductive symbiotic rhizobia triggers host sanctions that reduce fitness of their reproductive clonemates', *Proceedings: Biological Sciences,* 278(1718), pp. 2698-2703.

Overlöper, A., Kraus, A., Gurski, R., Wright, P. R., Georg, J., Hess, W. R. and Narberhaus, F. (2014) 'Two separate modules of the conserved regulatory RNA AbcR1 address multiple target mRNAs in and outside of the translation initiation region', *RNA Biol,* 11(5), pp. 624-40.

Ow, M. C., Perwez, T. and Kushner, S. R. (2003) 'RNase G of *Escherichia coli* exhibits only limited functional overlap with its essential homologue, RNase E', *Mol Microbiol,* 49(3), pp. 607-22.

Pandey, S. P., Minesinger, B. K., Kumar, J. and Walker, G. C. (2011) 'A highly conserved protein of unknown function in *Sinorhizobium meliloti* affects sRNA regulation similar to Hfq', *Nucleic Acids Res: Vol. 11*, pp. 4691-708.

Pandey, S. P., Winkler, J. A., Li, H., Camacho, D. M., Collins, J. J. and Walker, G. C. (2014) 'Central role for RNase YbeY in Hfq-dependent and Hfq-independent small-RNA regulation in bacteria', *BMC Genomics,* 15, pp. 121.

Papenfort, K., Bouvier, M., Mika, F., Sharma, C. M. and Vogel, J. (2010) 'Evidence for an autonomous 5′ target recognition domain in an Hfq-associated small RNA', *Proc Natl Acad Sci U S A,* 107(47), pp. 20435-40.

Papenfort, K., Pfeiffer, V., Mika, F., Lucchini, S., Hinton, J. C. D. and Vogel, J. (2006) 'σE-dependent small RNAs of *Salmonella* respond to membrane stress by accelerating global omp mRNA decay', *Mol Microbiol,* 62(6), pp. 1674-88.

Parker, J. S., Parizotto, E. A., Wang, M., Roe, S. M. and Barford, D. (2009) 'Enhancement of the Seed-Target Recognition Step in RNA Silencing by a PIWI/MID Domain Protein', *Mol Cell: Vol. 2*, pp. 204-14.

Patriarca, E. J., Tatè, R. and Iaccarino, M. (2002) 'Key Role of Bacterial $NH4^+$ Metabolism in *Rhizobium*-Plant Symbiosis', *Microbiol Mol Biol Rev,* 66(2), pp. 203-22.

Peck, M. C., Fisher, R. F. and Long, S. R. (2006) 'Diverse Flavonoids Stimulate NodD1 Binding to nod Gene Promoters in *Sinorhizobium meliloti*'.

Phadtare, S. and Severinov, K. (2010) 'RNA remodeling and gene regulation by

cold shock proteins', *RNA Biol,* 7(6), pp. 788-95.

Pobre, V. and Arraiano, C. M. (2015) 'Next generation sequencing analysis reveals that the ribonucleases RNase II, RNase R and PNPase affect bacterial motility and biofilm formation in *E. coli'*, *BMC Genomics,* 16, pp. 72.

Preiss, J., Yung, S. G. and Baecker, P. A. (1983) 'Regulation of bacterial glycogen synthesis', *Mol Cell Biochem,* 57(1), pp. 61-80.

Prell, J., Bourdes, A., Kumar, S., Lodwig, E., Hosie, A., Kinghorn, S., White, J. and Poole, P. (2010) 'Role of symbiotic auxotrophy in the *Rhizobium*-legume symbioses', *PLoS One,* 5(11), pp. e13933.

Prell, J. and Poole, P. (2006) 'Metabolic changes of rhizobia in legume nodules', *Trends Microbiol,* 14(4), pp. 161-8.

Judicial Commission of the International Committee on Systematics of Prokaryotes. (2008) 'The genus name *Sinorhizobium* Chen *et al*. 1988 is a later synonym of *Ensifer Casida* 1982 and is not conserved over the latter genus name, and the species name *'Sinorhizobium adhaerens'* is not validly published. Opinion 84', *Int J Syst Evol Microbiol,* 58(Pt 8), pp. 1973.

Pulvermacher, S. C., Stauffer, L. T. and Stauffer, G. V. (2009) 'Role of the *Escherichia coli* Hfq protein in GcvB regulation of *oppA* and *dppA* mRNAs', *Microbiology,* 155(Pt 1), pp. 115-23.

Qu, J., Lesse, A. J., Brauer, A. L., Cao, J., Gill, S. R. and Murphy, T. F. (2010) 'Proteomic expression profiling of *Haemophilus influenzae* grown in pooled human sputum from adults with chronic obstructive pulmonary disease reveal antioxidant and stress responses', *BMC Microbiol,* 10, pp. 162.

Rasouly, A., Davidovich, C. and Ron, E. Z. (2010) 'The Heat Shock Protein YbeY Is Required for Optimal Activity of the 30S Ribosomal Subunit▽', *J Bacteriol: Vol. 18*, pp. 4592-6.

Rasouly, A., Schonbrun, M., Shenhar, Y. and Ron, E. Z. (2009) 'YbeY, a Heat Shock Protein Involved in Translation in *Escherichia coli*', *J Bacteriol: Vol. 8*, pp. 2649-55.

Regalia, M., Rosenblad, M. A. and Samuelsson, T. (2002) 'Prediction of signal recognition particle RNA genes', *Nucleic Acids Res,* 30(15), pp. 3368-77.

Regnier, P. and Arraiano, C. M. (2000) 'Degradation of mRNA in bacteria: emergence of ubiquitous features', *Bioessays,* 22(3), pp. 235-44.

Regnier, P. and Grunberg-Manago, M. (1990) 'RNase III cleavages in non-coding leaders of *Escherichia coli* transcripts control mRNA stability and genetic expression', *Biochimie,* 72(11), pp. 825-34.

Reinkensmeier, J. and Giegerich, R. (2015) 'Thermodynamic matchers for the construction of the cuckoo RNA family', *RNA Biol,* 12(2), pp. 197-207.

Reinkensmeier, J., Schluter, J. P., Giegerich, R. and Becker, A. (2011) 'Conservation and Occurrence of *Trans*-Encoded sRNAs in the Rhizobiales', *Genes (Basel)*, 2(4), pp. 925-56.

Richter, A. S. and Backofen, R. (2012) 'Accessibility and conservation: general features of bacterial small RNA-mRNA interactions?', *RNA Biol,* 9(7), pp. 954-65.

Rivas, E., Clements, J. and Eddy, S. R. (2017) 'A statistical test for conserved RNA structure shows lack of evidence for structure in lncRNAs', *Nat Methods,* 14(1), pp. 45-48.

Rivas, E., Klein, R. J., Jones, T. A. and Eddy, S. R. (2001) 'Computational identification of noncoding RNAs in *E. coli* by comparative genomics', *Curr Biol,* 11(17), pp. 1369-73.

Robertson, G. T. and Roop, R. M., Jr. (1999) 'The Brucella abortus host factor I (HF-I) protein contributes to stress resistance during stationary phase and is a major determinant of virulence in mice', *Mol Microbiol,* 34(4), pp. 690-700.

Robledo, M., Frage, B., Wright, P. R. and Becker, A. (2015) 'A stress-induced small RNA modulates alpha-rhizobial cell cycle progression', *PLoS Genet,* 11(4), pp. e1005153.

Robledo, M., Jimenez-Zurdo, J. I., Soto, M. J., Velazquez, E., Dazzo, F., Martinez-Molina, E. and Mateos, P. F. (2011) 'Development of functional symbiotic white clover root hairs and nodules requires tightly regulated production of rhizobial cellulase CelC2', *Mol Plant Microbe Interact,* 24(7), pp. 798-807.

Robledo, M., Peregrina, A., Millan, V., Garcia-Tomsig, N. I., Torres-Quesada, O., Mateos, P. F., Becker, A. and Jimenez-Zurdo, J. I. (2017) 'A conserved

alpha-proteobacterial small RNA contributes to osmoadaptation and symbiotic efficiency of rhizobia on legume roots', *Environ Microbiol*.

Romby, P. and Charpentier, E. (2010) 'An overview of RNAs with regulatory functions in gram-positive bacteria', *Cell Mol Life Sci,* 67(2), pp. 217-37.

Rotter, C., Muhlbacher, S., Salamon, D., Schmitt, R. and Scharf, B. (2006) 'Rem, a new transcriptional activator of motility and chemotaxis in *Sinorhizobium meliloti*', *J Bacteriol,* 188(19), pp. 6932-42.

Roux, B., Rodde, N., Jardinaud, M. F., Timmers, T., Sauviac, L., Cottret, L., Carrere, S., Sallet, E., Courcelle, E., Moreau, S., Debelle, F., Capela, D., de Carvalho-Niebel, F., Gouzy, J., Bruand, C. and Gamas, P. (2014) 'An integrated analysis of plant and bacterial gene expression in symbiotic root nodules using laser-capture microdissection coupled to RNA sequencing', *Plant J,* 77(6), pp. 817-37.

Ruberg, S., Tian, Z. X., Krol, E., Linke, B., Meyer, F., Wang, Y., Puhler, A., Weidner, S. and Becker, A. (2003) 'Construction and validation of a *Sinorhizobium meliloti* whole genome DNA microarray: genome-wide profiling of osmoadaptive gene expression', *J Biotechnol,* 106(2-3), pp. 255-68.

Saadeh, B., Caswell, C. C., Chao, Y., Berta, P., Wattam, A. R., Roop, R. M., 2nd and O'Callaghan, D. (2015) 'Transcriptome-Wide Identification of Hfq-Associated RNAs in *Brucella* suis by Deep Sequencing', *J Bacteriol,* 198(3), pp. 427-35.

Sahr, T., Bruggemann, H., Jules, M., Lomma, M., Albert-Weissenberger, C., Cazalet, C. and Buchrieser, C. (2009) 'Two small ncRNAs jointly govern virulence and transmission in *Legionella pneumophila*', *Mol Microbiol,* 72(3), pp. 741-62.

Sallet, E., Roux, B., Sauviac, L., Jardinaud, M. F., Carrere, S., Faraut, T., de Carvalho-Niebel, F., Gouzy, J., Gamas, P., Capela, D., Bruand, C. and Schiex, T. (2013a) 'Next-generation annotation of prokaryotic genomes with EuGene-P: application to *Sinorhizobium meliloti* 2011', *DNA Res,* 20(4), pp. 339-54.

Santos, J. M., Drider, D., Marujo, P. E., Lopez, P. and Arraiano, C. M. (1997) 'Determinant role of *E. coli* RNase III in the decay of both specific and heterologous mRNAs', *FEMS Microbiol Lett,* 157(1), pp. 31-8.

Saramago, M., Barria, C., Dos Santos, R. F., Silva, I. J., Pobre, V., Domingues, S., Andrade, J. M., Viegas, S. C. and Arraiano, C. M. (2014) 'The role of RNases in the regulation of small RNAs', *Curr Opin Microbiol,* 18, pp. 105-15.

Saramago, M., Peregrina, A., Robledo, M., Matos, R. G., Hilker, R., Serrania, J., Becker, A., Arraiano, C. M. and Jimenez-Zurdo, J. I. (2017) *'Sinorhizobium meliloti* YbeY is an endoribonuclease with unprecedented catalytic features, acting as silencing enzyme in riboregulation', *Nucleic Acids Res,* 45(3), pp. 1371-1391.

Sauer, E. and Weichenrieder, O. (2011) 'Structural basis for RNA 3'-end recognition by Hfq', *Proc Natl Acad Sci U S A,* 108(32), pp. 13065-70.

Sauviac, L., Philippe, H., Phok, K. and Bruand, C. (2007) 'An Extracytoplasmic Function Sigma Factor Acts as a General Stress Response Regulator in *Sinorhizobium meliloti*', *J Bacteriol,* 189(11), pp. 4204-16.

Schlüter, J. P., Reinkensmeier, J., Barnett, M. J., Lang, C., Krol, E., Giegerich, R., Long, S. R. and Becker, A. (2013) 'Global mapping of transcription start sites and promoter motifs in the symbiotic α-proteobacterium *Sinorhizobium meliloti* 1021', *BMC Genomics,* 14, pp. 156.

Schlüter, J. P., Reinkensmeier, J., Daschkey, S., Evguenieva-Hackenberg, E., Janssen, S., Jänicke, S., Becker, J. D., Giegerich, R. and Becker, A. (2010) 'A genome-wide survey of sRNAs in the symbiotic nitrogen-fixing alpha-proteobacterium *Sinorhizobium meliloti'*, *BMC Genomics*, pp. 245.

Schulz, A., Stoveken, N., Binzen, I. M., Hoffmann, T., Heider, J. and Bremer, E. (2017) 'Feeding on compatible solutes: A substrate-induced pathway for uptake and catabolism of ectoines and its genetic control by EnuR', *Environ Microbiol,* 19(3), pp. 926-946.

Schumacher, M. A., Pearson, R. F., Moller, T., Valentin-Hansen, P. and Brennan, R. G. (2002) 'Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: a bacterial Sm-like protein', *Embo j,* 21(13), pp. 3546-56.

Schwarz, G., Mendel, R. R. and Ribbe, M. W. (2009) 'Molybdenum cofactors, enzymes and pathways', *Nature,* 460(7257), pp. 839-847.

Schäfer, A., Tauch, A., Jager, W., Kalinowski, J., Thierbach, G. and Puhler, A. (1994) '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(1), pp. 69-73.

Seibold, G. M. and Eikmanns, B. J. (2007) 'The glgX gene product of *Corynebacterium glutamicum* is required for glycogen degradation and for fast adaptation to hyperosmotic stress', *Microbiology,* 153(Pt 7), pp. 2212-20.

Serganov, A. (2009) 'The long and the short of riboswitches', *Curr Opin Struct Biol,* 19(3), pp. 251-9.

Serganov, A., Huang, L. and Patel, D. J. (2009) 'Coenzyme recognition and gene regulation by a flavin mononucleotide riboswitch', *Nature,* 458(7235), pp. 233-7.

Serrania, J., Vorholter, F. J., Niehaus, K., Puhler, A. and Becker, A. (2008) 'Identification of *Xanthomonas campestris* pv. campestris galactose utilization genes from transcriptome data', *J Biotechnol,* 135(3), pp. 309-17.

Shao, S., von der Malsburg, K. and Hegde, R. S. (2013) 'Listerin-dependent nascent protein ubiquitination relies on ribosome subunit dissociation', *Mol Cell,* 50(5), pp. 637-48.

Sharma, C. M., Darfeuille, F., Plantinga, T. H. and Vogel, J. (2007) 'A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites', *Genes Dev,* 21(21), pp. 2804-17.

Sharma, C. M., Hoffmann, S., Darfeuille, F., Reignier, J., Findeiss, S., Sittka, A., Chabas, S., Reiche, K., Hackermuller, J., Reinhardt, R., Stadler, P. F. and Vogel, J. (2010) 'The primary transcriptome of the major human pathogen *Helicobacter pylori'*, *Nature,* 464(7286), pp. 250-5.

Sharma, C. M., Papenfort, K., Pernitzsch, S. R., Mollenkopf, H. J., Hinton, J. C. and Vogel, J. (2011) 'Pervasive post-transcriptional control of genes involved in amino acid metabolism by the Hfq-dependent GcvB small RNA', *Mol Microbiol,* 81(5), pp. 1144-65.

Sharma, C. M. and Vogel, J. (2009) 'Experimental approaches for the discovery and characterization of regulatory small RNA', *Curr Opin Microbiol,* 12(5), pp. 536-46.

Sharma, R. P., Grayson, D. R. and Gavin, D. P. (2008) 'Histone deactylase 1 expression is increased in the prefrontal cortex of schizophrenia subjects: analysis of the National Brain Databank microarray collection', *Schizophr Res,* 98(1-3),

pp. 111-7.

Shrivastava, P. and Kumar, R. (2015) 'Soil salinity: A serious environmental issue and plant growth promoting bacteria as one of the tools for its alleviation', *Saudi J Biol Sci,* 22(2), pp. 123-31.

Simon, R., Priefer, U. and Pühler, A. (1983) 'A Broad Host Range Mobilization System for *in vivo* Genetic Engineering: Transposon Mutagenesis in Gram Negative Bacteria', *Nature Biotechnology,* 1(9), pp. 784-791.

Sittka, A., Lucchini, S., Papenfort, K., Sharma, C. M., Rolle, K., Binnewies, T. T., Hinton, J. C. and Vogel, J. (2008) 'Deep Sequencing Analysis of Small Noncoding RNA and mRNA Targets of the Global Post-Transcriptional Regulator, Hfq', *PLoS Genet,* 4(8).

Sittka, A., Pfeiffer, V., Tedin, K. and Vogel, J. (2007) 'The RNA chaperone Hfq is essential for the virulence of *Salmonella typhimurium'*, *Mol Microbiol,* 63(1), pp. 193-217.

Sittka, A., Sharma, C. M., Rolle, K. and Vogel, J. (2009) 'Deep sequencing of *Salmonella* RNA associated with heterologous Hfq proteins in vivo reveals small RNAs as a major target class and identifies RNA processing phenotypes', *RNA Biol,* 6(3), pp. 266-75.

Siu, F. Y., Spanggord, R. J. and Doudna, J. A. (2007) 'SRP RNA provides the physiologically essential GTPase activation function in cotranslational protein targeting', *RNA,* 13(2), pp. 240-50.

Sleator, R. D. and Hill, C. (2002) 'Bacterial osmoadaptation: the role of osmolytes in bacterial stress and virulence', *FEMS Microbiol Rev,* 26(1), pp. 49-71.

Smith, C., Heyne, S., Richter, A. S., Will, S. and Backofen, R. (2010) 'Freiburg RNA Tools: a web server integrating IntaRNA, ExpaRNA and LocaRNA', *Nucleic Acids Res,* 38(Web Server issue), pp. W373-7.

Sobrero, P., Schluter, J. P., Lanner, U., Schlosser, A., Becker, A. and Valverde, C. (2012) 'Quantitative proteomic analysis of the Hfq-regulon in *Sinorhizobium meliloti* 2011', *PLoS One,* 7(10), pp. e48494.

Sobrero, P. and Valverde, C. (2011) 'Evidences of autoregulation of hfq expression in *Sinorhizobium meliloti* strain 2011', *Arch Microbiol,* 193(9), pp. 629-39.

Sobrero, P. and Valverde, C. (2012a) 'The bacterial protein Hfq: much more than a mere RNA-binding factor', *Crit Rev Microbiol,* 38(4), pp. 276-99.

Sourjik, V., Muschler, P., Scharf, B. and Schmitt, R. (2000) 'VisN and VisR Are Global Regulators of Chemotaxis, Flagellar, and Motility Genes in *Sinorhizobium (Rhizobium) meliloti*', *J Bacteriol: Vol. 3*, pp. 782-8.

Stanek, K. A., West, J. P., Randolph, P. S. and Mura, C. (2016) 'Crystal structure and RNA-binding properties of an Hfq homolog from the deep-branching *Aquificae*: Conservation of the lateral RNA-binding mode'.

Starczynowski, D. T., Morin, R., McPherson, A., Lam, J., Chari, R., Wegrzyn, J., Kuchenbauer, F., Hirst, M., Tohyama, K., Humphries, R. K., Lam, W. L., Marra, M. and Karsan, A. (2011) 'Genome-wide identification of human microRNAs located in leukemia-associated genomic alterations', *Blood,* 117(2), pp. 595-607.

Starker, C. G., Parra-Colmenares, A. L., Smith, L., Mitra, R. M. and Long, S. R. (2006) 'Nitrogen Fixation Mutants of *Medicago truncatula* Fail to Support Plant and Bacterial Symbiotic Gene Expression1[W][OA]', *Plant Physiol,* 140(2), pp. 671-80.

Stead, M. B., Marshburn, S., Mohanty, B. K., Mitra, J., Pena Castillo, L., Ray, D., van Bakel, H., Hughes, T. R. and Kushner, S. R. (2011) 'Analysis of *Escherichia coli* RNase E and RNase III activity in vivo using tiling microarrays', *Nucleic Acids Res,* 39(8), pp. 3188-203.

Storz, G. (2002) 'An expanding universe of noncoding RNAs', *Science,* 296(5571), pp. 1260-3.

Storz, G., Altuvia, S. and Wassarman, K. M. (2005) 'An abundance of RNA regulators', *Annu Rev Biochem*, 74, pp. 199-217.

Storz, G., Vogel, J. and Wassarman, K. M. (2011) 'Regulation by small RNAs in bacteria: expanding frontiers', *Mol Cell,* 43(6), pp. 880-91.

Studier, F. W. (1991) 'Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system', *J Mol Biol,* 219(1), pp. 37-44.

Sulthana, S., Basturea, G. N. and Deutscher, M. P. (2016) 'Elucidation of pathways of ribosomal RNA degradation: an essential role for RNase E', *Rna,* 22(8), pp. 1163-71.

Sun, X., Zhulin, I. and Wartell, R. M. (2002) 'Predicted structure and phyletic distribution of the RNA-binding protein Hfq', *Nucleic Acids Res: Vol. 17*, pp. 3662-71.

Suzuki, K., Babitzke, P., Kushner, S. R. and Romeo, T. (2006) 'Identification of a novel regulatory protein (CsrD) that targets the global regulatory RNAs CsrB and CsrC for degradation by RNase E', *Genes Dev,* 20(18), pp. 2605-17.

Tatusov, R. L., Fedorova, N. D., Jackson, J. D., Jacobs, A. R., Kiryutin, B., Koonin, E. V., Krylov, D. M., Mazumder, R., Mekhedov, S. L., Nikolskaya, A. N., Rao, B. S., Smirnov, S., Sverdlov, A. V., Vasudevan, S., Wolf, Y. I., Yin, J. J. and Natale, D. A. (2003) 'The COG database: an updated version includes eukaryotes', *BMC Bioinformatics*, pp. 41.

Temin, H. M. (1985) 'Reverse transcription in the eukaryotic genome: retroviruses, pararetroviruses, retrotransposons, and retrotranscripts.', *Molecular Biology and Evolution,* 2(6), pp. 455-468.

Tjaden, B., Goodwin, S. S., Opdyke, J. A., Guillier, M., Fu, D. X., Gottesman, S. and Storz, G. (2006) 'Target prediction for small, noncoding RNAs in bacteria', *Nucleic Acids Res,* 34(9), pp. 2791-802.

Tomizawa, J. (1984) 'Control of ColE1 plasmid replication: the process of binding of RNA I to the primer transcript', *Cell,* 38(3), pp. 861-70.

Tomizawa, J. (1985) 'Control of ColE1 plasmid replication: initial interaction of RNA I and the primer transcript is reversible', *Cell,* 40(3), pp. 527-35.

Torres-Quesada, O., Millan, V., Nisa-Martinez, R., Bardou, F., Crespi, M., Toro, N. and Jimenez-Zurdo, J. I. (2013) 'Independent activity of the homologous small regulatory RNAs AbcR1 and AbcR2 in the legume symbiont *Sinorhizobium meliloti'*, *PLoS One,* 8(7), pp. e68147.

Torres-Quesada, O., Oruezabal, R. I., Peregrina, A., Jofré, E., Lloret, J., Rivilla, R., Toro, N. and Jiménez-Zurdo, J. I. (2010) 'The *Sinorhizobium meliloti* RNA chaperone Hfq influences central carbon metabolism and the symbiotic interaction with alfalfa', *BMC Microbiology,* 10(1), pp. 71.

Torres-Quesada, O., Reinkensmeier, J., Schluter, J. P., Robledo, M., Peregrina, A., Giegerich, R., Toro, N., Becker, A. and Jimenez-Zurdo, J. I. (2014) 'Genome-wide profiling of Hfq-binding RNAs uncovers extensive post-transcriptional rewiring of major stress response and symbiotic regulons in

Sinorhizobium meliloti', *RNA Biol,* 11(5), pp. 563-79.

Triplett, E. W. and Sadowsky, M. J. (1992) 'Genetics of competition for nodulation of legumes', *Annu Rev Microbiol,* 46, pp. 399-428.

Trotochaud, A. E. and Wassarman, K. M. (2005) 'A highly conserved 6S RNA structure is required for regulation of transcription', *Nat Struct Mol Biol,* 12(4), pp. 313-9.

Tsui, H. C., Leung, H. C. and Winkler, M. E. (1994) 'Characterization of broadly pleiotropic phenotypes caused by an *hfq* insertion mutation in *Escherichia coli* K-12', *Mol Microbiol,* 13(1), pp. 35-49.

Uchiyama, I. (2007) 'MBGD: a platform for microbial comparative genomics based on the automated construction of orthologous groups', *Nucleic Acids Res,* 35(Database issue), pp. D343-6.

Ulve, V. M., Cheron, A., Trautwetter, A., Fontenelle, C. and Barloy-Hubler, F. (2007a) 'Characterization and expression patterns of *Sinorhizobium meliloti* tmRNA (*ssrA*)', *FEMS Microbiol Lett,* 269(1), pp. 117-23.

Ulve, V. M., Sevin, E. W., Cheron, A. and Barloy-Hubler, F. (2007b) 'Identification of chromosomal alpha-proteobacterial small RNAs by comparative genome analysis and detection in Sinorhizobium meliloti strain 1021', *BMC Genomics,* 8, pp. 467.

Urban, J. H. and Vogel, J. (2007) 'Translational control and target recognition by *Escherichia coli* small RNAs *in vivo*', *Nucleic Acids Res,* 35(3), pp. 1018-37.

Urban, J. H. and Vogel, J. (2008) 'Two seemingly homologous noncoding RNAs act hierarchically to activate *glmS* mRNA translation', *PLoS Biol,* 6(3), pp. e64.

Valentin-Hansen, P., Eriksen, M. and Udesen, C. (2004) 'The bacterial Sm-like protein Hfq: a key player in RNA transactions', *Mol Microbiol,* 51(6), pp. 1525-33.

Valverde, C., Livny, J., Schlüter, J. P., Reinkensmeier, J., Becker, A. and Parisi, G. (2008) 'Prediction of *Sinorhizobium meliloti* sRNA genes and experimental detection in strain 2011', *BMC Genomics*, pp. 416.

van Dillewijn, P., Soto, M. J., Villadas, P. J. and Toro, N. (2001) 'Construction and Environmental Release of a *Sinorhizobium meliloti* Strain Genetically Modified

To Be More Competitive for Alfalfa Nodulation', *Appl Environ Microbiol,* 67(9), pp. 3860-5.

Vasse, J., de Billy, F., Camut, S. and Truchet, G. (1990) 'Correlation between ultrastructural differentiation of bacteroids and nitrogen fixation in alfalfa nodules', *J Bacteriol,* 172(8), pp. 4295-306.

Vercruysse, M., Kohrer, C., Davies, B. W., Arnold, M. F., Mekalanos, J. J., RajBhandary, U. L. and Walker, G. C. (2014) 'The highly conserved bacterial RNase YbeY is essential in *Vibrio cholerae*, playing a critical role in virulence, stress regulation, and RNA processing', *PLoS Pathog,* 10(6), pp. e1004175.

Viegas, S. C. and Arraiano, C. M. (2008) 'Regulating the regulators: How ribonucleases dictate the rules in the control of small non-coding RNAs', *RNA Biol,* 5(4), pp. 230-43.

Viegas, S. C., Fernandez De Palencia, P., Amblar, M., Arraiano, C. M. and Lopez, P. (2004) 'Development of an inducible system to control and easily monitor gene expression in *Lactococcus lactis*', *Plasmid,* 51(3), pp. 256-64.

Viegas, S. C., Pfeiffer, V., Sittka, A., Silva, I. J., Vogel, J. and Arraiano, C. M. (2007) 'Characterization of the role of ribonucleases in *Salmonella* small RNA decay', *Nucleic Acids Res,* 35(22), pp. 7651-64.

Viegas, S. C., Schmidt, D., Kasche, V., Arraiano, C. M. and Ignatova, Z. (2005) 'Effect of the increased stability of the penicillin amidase mRNA on the protein expression levels', *FEBS Lett,* 579(22), pp. 5069-73.

Vinayagam, A., del Val, C., Schubert, F., Eils, R., Glatting, K. H., Suhai, S. and Konig, R. (2006) 'GOPET: a tool for automated predictions of Gene Ontology terms', *BMC Bioinformatics,* 7, pp. 161.

Vitreschak, A. G., Rodionov, D. A., Mironov, A. A. and Gelfand, M. S. (2002) 'Regulation of riboflavin biosynthesis and transport genes in bacteria by transcriptional and translational attenuation', *Nucleic Acids Res,* 30(14), pp. 3141-51.

Vogel, J. (2009) 'An RNA trap helps bacteria get the most out of chitosugars', *Mol Microbiol,* 73(5), pp. 737-41.

Vogel, J., Argaman, L., Wagner, E. G. and Altuvia, S. (2004) 'The small RNA IstR

inhibits synthesis of an SOS-induced toxic peptide', *Curr Biol,* 14(24), pp. 2271-6.

Vogel, J., Bartels, V., Tang, T. H., Churakov, G., Slagter-Jäger, J. G., Hüttenhofer, A. and Wagner, E. G. H. (2003) 'RNomics in *Escherichia coli* detects new sRNA species and indicates parallel transcriptional output in bacteria', *Nucleic Acids Res,* 31(22), pp. 6435-43.

Vogel, J. and Luisi, B. F. (2011) 'Hfq and its constellation of RNA', *Nat Rev Microbiol,* 9(8), pp. 578-89.

Vogel, J. and Wagner, E. G. (2007) 'Target identification of small noncoding RNAs in bacteria', *Curr Opin Microbiol,* 10(3), pp. 262-70.

Voss, B., Holscher, M., Baumgarth, B., Kalbfleisch, A., Kaya, C., Hess, W. R., Becker, A. and Evguenieva-Hackenberg, E. (2009) 'Expression of small RNAs in Rhizobiales and protection of a small RNA and its degradation products by Hfq in *Sinorhizobium meliloti'*, *Biochem Biophys Res Commun,* 390(2), pp. 331-6.

Wagner, E. G. H. (2013) 'Cycling of RNAs on Hfq', *RNA Biol,* 10(4), pp. 619-26.

Wais, R. J., Wells, D. H. and Long, S. R. (2002) 'Analysis of differences between *Sinorhizobium meliloti* 1021 and 2011 strains using the host calcium spiking response', *Mol Plant Microbe Interact,* 15(12), pp. 1245-52.

Wang, W., Wang, L., Wu, J., Gong, Q. and Shi, Y. (2013) 'Hfq-bridged ternary complex is important for translation activation of *rpoS* by DsrA', *Nucleic Acids Res,* 41(11), pp. 5938-48.

Washietl, S., Hofacker, I. L. and Stadler, P. F. (2005) 'Fast and reliable prediction of noncoding RNAs'.

Wassarman, K. M. (2007) '6S RNA: a regulator of transcription', *Molecular Microbiology,* 65(6), pp. 1425-1431.

Wassarman, K. M., Repoila, F., Rosenow, C., Storz, G. and Gottesman, S. (2001) 'Identification of novel small RNAs using comparative genomics and microarrays', *Genes Dev,* 15(13), pp. 1637-51.

Wassarman, K. M. and Saecker, R. M. (2006) 'Synthesis-mediated release of a small RNA inhibitor of RNA polymerase', *Science,* 314(5805), pp. 1601-3.

Wassarman, K. M. and Storz, G. (2000) '6S RNA regulates *E. coli* RNA polymerase

activity', *Cell,* 101(6), pp. 613-23.

Wassarman, K. M., Zhang, A. and Storz, G. (1999) 'Small RNAs in *Escherichia coli*', *Trends Microbiol,* 7(1), pp. 37-45.

Waters, L. S. and Storz, G. (2009) 'Regulatory RNAs in bacteria', *Cell,* 136(4), pp. 615-28.

Weber, A. and Jung, K. (2002) 'Profiling early osmostress-dependent gene expression in *Escherichia coli* using DNA macroarrays', *J Bacteriol,* 184(19), pp. 5502-7.

Weber, A., Kogl, S. A. and Jung, K. (2006) 'Time-dependent proteome alterations under osmotic stress during aerobic and anaerobic growth in *Escherichia coli*', *J Bacteriol,* 188(20), pp. 7165-75.

Wei, B. L., Brun-Zinkernagel, A. M., Simecka, J. W., Pruss, B. M., Babitzke, P. and Romeo, T. (2001) 'Positive regulation of motility and *flhDC* expression by the RNA-binding protein CsrA of *Escherichia coli*', *Mol Microbiol,* 40(1), pp. 245-56.

Wemekamp-Kamphuis, H. H., Wouters, J. A., Sleator, R. D., Gahan, C. G. M., Hill, C. and Abee, T. (2002) 'Multiple Deletions of the Osmolyte Transporters BetL, Gbu, and OpuC of *Listeria monocytogenes* Affect Virulence and Growth at High Osmolarity', *Appl Environ Microbiol,* 68(10), pp. 4710-6.

White, D., Hart, M. E. and Romeo, T. (1996) 'Phylogenetic distribution of the global regulatory gene *csrA* among eubacteria', *Gene,* 182(1-2), pp. 221-3.

Will, S., Reiche, K., Hofacker, I. L., Stadler, P. F. and Backofen, R. (2007) 'Inferring Noncoding RNA Families and Classes by Means of Genome-Scale Structure-Based Clustering', *PLoS Comput Biol,* 3(4).

Wilms, I., Voss, B., Hess, W. R., Leichert, L. I. and Narberhaus, F. (2011) 'Small RNA-mediated control of the *Agrobacterium tumefaciens* GABA binding protein', *Mol Microbiol,* 80(2), pp. 492-506.

Wright, P. R., Georg, J., Mann, M., Sorescu, D. A., Richter, A. S., Lott, S., Kleinkauf, R., Hess, W. R. and Backofen, R. (2014) 'CopraRNA and IntaRNA: predicting small RNA targets, networks and interaction domains', *Nucleic Acids Res,* 42(Web Server issue), pp. W119-23.

Wright, P. R., Richter, A. S., Papenfort, K., Mann, M., Vogel, J., Hess, W. R., Backofen, R. and Georg, J. (2013) 'Comparative genomics boosts target prediction for bacterial small RNAs', *Proc Natl Acad Sci U S A,* 110(37), pp. E3487-96.

Wurm, R., Neusser, T. and Wagner, R. (2010) '6S RNA-dependent inhibition of RNA polymerase is released by RNA-dependent synthesis of small de novo products', *Biol Chem,* 391(2-3), pp. 187-96.

Yang, W. C., de Blank, C., Meskiene, I., Hirt, H., Bakker, J., van Kammen, A., Franssen, H. and Bisseling, T. (1994) '*Rhizobium* nod factors reactivate the cell cycle during infection and nodule primordium formation, but the cycle is only completed in primordium formation', *Plant Cell: Vol. 10*, pp. 1415-26.

Young, J. M. (2003) 'The genus name *Ensifer Casida* 1982 takes priority over *Sinorhizobium* Chen *et al*. 1988, and *Sinorhizobium morelense* Wang *et al*. 2002 is a later synonym of *Ensifer adhaerens* Casida 1982. Is the combination "*Sinorhizobium adhaerens*" (Casida 1982) Willems *et al*. 2003 legitimate? Request for an Opinion', *Int J Syst Evol Microbiol,* 53(Pt 6), pp. 2107-10.

Yurgel, S. N., Rice, J. and Kahn, M. L. (2013) 'Transcriptome analysis of the role of GlnD/GlnBK in nitrogen stress adaptation by *Sinorhizobium meliloti* Rm1021', *PLoS One,* 8(3), pp. e58028.

Zahran, H. H. (1991) 'Conditions for successful *Rhizobium*-legume symbiosis in saline environments', *Biology and fertility of soils,* 12(1), pp. 73-80.

Zahran, H. H. (1999) '*Rhizobium*-legume symbiosis and nitrogen fixation under severe conditions and in an arid climate', *Microbiol Mol Biol Rev,* 63(4), pp. 968-89, table of contents.

Zhan, C., Fedorov, E. V., Shi, W., Ramagopal, U. A., Thirumuruhan, R., Manjasetty, B. A., Almo, S. C., Fiser, A., Chance, M. R. and Fedorov, A. A. (2005) 'The ybeY protein from *Escherichia coli* is a metalloprotein', *Acta Crystallogr Sect F Struct Biol Cryst Commun: Vol. Pt 11*, pp. 959-63.

Zhang, A., Altuvia, S., Tiwari, A., Argaman, L., Hengge-Aronis, R. and Storz, G. (1998) 'The OxyS regulatory RNA represses *rpoS* translation and binds the Hfq (HF-I) protein', *Embo j,* 17(20), pp. 6061-8.

Zhang, A., Wassarman, K. M., Rosenow, C., Tjaden, B. C., Storz, G. and Gottesman, S. (2003) 'Global analysis of small RNA and mRNA targets of Hfq',

Mol Microbiol, 50(4), pp. 1111-24.