Universidad de Granada

Estación Experimental del Zaidín

(Consejo Superior de Investigaciones Científicas)

RNA Regulation of Metabolism in the Legume Symbiont *Sinorhizobium* (*Ensifer*) *meliloti*

Natalia Isabel García Tomsig

Tesis Doctoral / *Doctoral Thesis*

2022

Programa de Doctorado en Biología Fundamental y de Sistemas

Doctoral Program in Fundamental and Systems Biology

Directores de la Tesis/*Thesis supervisors:*

Dr. José I. Jiménez Zurdo

Dra. Marta Robledo Garrido

Editor: Universidad de Granada. Tesis Doctorales
Autor: Natalia Isabel García Tomsig
ISBN: 978-84-1117-362-9
URI: http://hdl.handle.net/10481/75449

Esta Tesis Doctoral ha sido realizada en el Departamento del Suelo y de la Planta (Grupo de Estructura, Dinámica y Función de Genomas de Rizobacterias), de la Estación Experimental del Zaidín (CSIC). Para su realización, se contó con la siguiente financiación:

- Beca predoctoral FPU, disfrutada entre el año 2017 y 2022 y adscrita a los proyectos BFU2017-82645-P y PID2020-114782GB-I00 financiados por MCIN/AEI/10.13039/501100011033/ y por FEDER Una manera de hacer Europa, siendo Investigador Principal de los mismos el Dr. José Ignacio Jiménez Zurdo.

- Beca de movilidad FPU para estancias breves del Ministerio de Educación y Formación profesional, disfrutada en *Zentrum für Synthetische Mikrobiologie* (Marburg, Alemania) bajo la dirección de la Dra. Anke Beker en el año 2018.

Parte del contenido de esta Tesis Doctoral ha sido publicado en en los siguientes artículos científicos, revisiones o protocolos de laboratorio:

Marta Robledo, Alexandra Peregrina, Vicenta Millán, **Natalia I. García-Tomsig**, Omar Torres-Quesada, Pedro F. Mateos, Anke Becker, José I. Jiménez-Zurdo. 2017. A conserved α-proteobacterial small RNA contributes to osmoadaptation and symbiotic efficiency of rhizobia on legume roots. *Environmental Microbiology* **19 (7),** 2661-2680. https://doi.org/10.1111/1462-2920.13757

Marta Robledo, **Natalia I. García-Tomsig**, José I. Jiménez-Zurdo. 2018. Primary Characterization of Small RNAs in Symbiotic Nitrogen-Fixing Bacteria. In: Medina C., López-Baena F. (eds) Host-Pathogen Interactions. *Methods in Molecular Biology*, vol. 1734. Humana Press, New York, NY. https://doi.org/10.1007/978-1-4939-7604-1_22

Marta Robledo, Ana M. Matia-González, **Natalia I. García-Tomsig**, José I. Jiménez-Zurdo. 2018. Identification of Small RNA–Protein Partners in Plant Symbiotic

Bacteria. In: Arluison V., Valverde C. (eds) Bacterial Regulatory RNA. *Methods in Molecular Biology*, vol. 1737. Humana Press, New York, NY. https://doi.org/10.1007/978-1-4939-7634-8_20

Marta Robledo, **Natalia I. García-Tomsig**, José I. Jiménez-Zurdo. 2020. Riboregulation in Nitrogen-Fixing Endosymbiotic Bacteria. *Microorganisms*. **8(3)**, 384. https://doi.org/10.3390/microorganisms8030384

Marta Robledo*, **Natalia I. García-Tomsig*,** Fernando M. García-Rodríguez, José I. Jiménez-Zurdo. 2021. Synthetase of the methyl donor S-adenosylmethionine from nitrogen-fixing α-rhizobia can bind functionally diverse RNA species. *RNA Biology* **18:8**, 1111-1123. https://doi.org/10.1080/15476286.2020.1829365. * Equal contribution

Natalia I. García-Tomsig, Marta Robledo, George C. diCenzo, Alessio Mengoni, Vicenta Millán, Alexandra Peregrina, Alejandro Uceta, José I. Jiménez-Zurdo. 2022. Pervasive RNA regulation of metabolism enhances the root colonization ability of nitrogen-fixing symbiotic α-rhizobia. *mBio* **13 (1)**, e03576-21. https://doi.org/10.1128/mbio.03576-21

Acknowledgements / Agradecimientos

En primer lugar, quiero agradecer a mis directores por todas las contribuciones que han hecho a mi formación investigadora.

Esta Tesis no podría haberse llevado a cabo sin la excepcional dirección del Dr. José I. Jiménez Zurdo en todos los aspectos. Espero haber incorporado a mi aprendizaje algo de su extenso conocimiento sobre el mundo del ARN, de su ingenio en la búsqueda de nuevas hipótesis a abordar y de su capacidad de liderazgo en el grupo. Debo también agradecer su infinita paciencia.

Por otra parte, esta Tesis ha contado con la co-diracción de la Dra. Marta Robledo Garrido, una investigadora sobresaliente, que contribuyó enormemente a mi aprendizaje experimental en el laboratorio. De ella aprendí la importancia de tratar los experimentos con rigor y de cuidar los detalles. Los resultados obtenidos aquí son también fruto de las buenas prácticas que me ha inculcado.

Agradezco también al Grupo de investigación "Estructura, Dinámica y Función de Genomas de Rizobacterias" de la EEZ, dónde se ha desarrollado esta Tesis, las facilidades que me han ofrecido. He de decir que todos ellos me han hecho sentir acogida dentro del grupo.

Quiero hacer especial mención a los investigadores con los que he compartido el día a día durante estos años: Fernando García, la Dra. María Dolores Molina, Vicenta Millán, la Dra. Ana Vicente, Ascensión Martos y José M. del Arco. Concretamente, Fernando García ha contribuido con mi aprendizaje en la manipulación de proteínas y Vicenta Millán contribuyó enormemente a mi formación en las diferentes técnicas de biología molecular. Por su parte, María Dolores me ha ayudado miles de veces a solventar problemas durante el

desarrollo de esta Tesis además de enseñarme minuciosamente algunas de las metodologías con las se han obtenido los resultados de esta Tesis. Todos ellos son grandes profesionales y han contribuido con este trabajo. También quiero agradecer a los técnicos del laboratorio Ascensión Martos y José M. del Arco por su gran ayuda en las tareas dairias de laboratorio.

Quiero agradecer a la Dra. María Trinidad Gallegos por su ayuda y asesoría en ensayos de interacción ARN/ADN-proteína y al grupo de la Dra. María J. Soto Misffut (EEZ) por su colaboración en los ensayos de colonización y de motilidad de *S. meliloti.*

I would like to thank Dra. Anke Becker for let me to work in her group during my international stay. I would also like to thank Dra. Elizaveta Krol for her assistance in Chapter 2.

I would like to thank Dr. George diCenzo and Dr. Alessio Mengoni for their very important contributions to data analysis of Chapter 1.

Personalmente, me tomo la licencia de agradecer a mi familia por su paciencia y cuidados durante tantos años. Esta Tesis es la culminación de muchos años de estudio durante lo que he recibido su incondicional apoyo a pesar de que las condiciones no fueran favorables. La gran capidad de trabajo duro y constante de mi abuela y mi madre, dos mujeres excepcionales, ha dejado una impronta en este trabajo. Quiero agredecer a mi tío, y padrino, toda la ayuda que nos dío, y el gran ejemplo de cordura que me ha dado. También quiero hacer mención a mi tío que aún no siendo de mi familia de nacimiento, me trató como su sobrina desde pequeña.

Más recientemente, quiero agradecer a mi pareja Luis y a sus padres por su apoyo en esta recta final. Agradezco a mi pareja la paciencia y el cariño que me brinda, permitiéndome compaginar este trabajo con la vida personal. Es una gran persona a la que admiro, y su enorme ayuda también ha contribuido al rendimiento de mi trabajo en esta Tesis.

Por úlitmo, quiero agradecer a cualquiera que se tome el tiempo de leer esta Tesis, espero que nuestro trabajo te resulte intereante y aporte algo a tus investigaciones.

Finally, I would like thank anyone who takes the time to read this Thesis, I hope our work is interesting to you and contributes to your research

Resumen

En las últimas dos décadas, se ha consolidado la idea de que los RNAs que no codifican proteínas asumen diversas funciones estructurales y reguladoras en todos los organismos vivos. Las tecnologías de secuenciación masiva de los transcriptomas (RNA-Seq) han identificado extensas y heterogéneas poblaciones de RNAs pequeños (sRNAs) con un potencial papel regulador en procariotas. Estos sRNAs actúan mayoritariamente mediante el apareamiento de bases asistido por proteínas con sus mRNAs diana para la regulación post-transcripcional de la reprogramación de la expresión génica como mecanismo ubicuo de adaptación de las bacterias en entornos fluctuantes. Sin embargo, este nivel de regulación genética es aún escasamente estudiado en la amplia mayoría de bacterias no-modelo. El conocimiento sobre el papel de los sRNAs en la simbiosis fijadora de nitrógeno entre rhizobios y leguminosas, deriva mayoritariamente de los estudios en la alpha-proteobacteria S. meliloti, pero incluso para este simbionte modelo, sólo se han caracterizado unos pocos sRNAs y proteínas que asisten la riboregulación. Una extensa clase de sRNAs son los llamados trans-sRNAs que se expresan diferencialmente desde regiones intergénicas y que suelen regular la traducción y/o estabilidad de sus mRNAs diana mediante el apareamiento de series cortas y discontinuas de nucleótidos complementarios, en un mecanismo que generalmente requiere la participación de proteínas (e.g., chaperonas de RNA y RNasas). Los *trans*-sRNAs homólogos AbcR1 y AbcR2 (ABC Regulators) y NfeR1 (Nodule Formation Efficiency Regulator) presentan motivos anti-Shine-Dalgarno (aSD) conservados en sus regiones desapareadas, que se predice que interaccionan con sus mRNAs diana para el bloqueo de la traducción. La caracterización primaria de éstos anticipa un gran impacto

en la riboregulación del metabolismo de *S. meliloti* a lo largo de la interacción simbiótica. Sin embargo, sus funciones no se han abordado aún en profundidad.

En esta Tesis, hemos descifrado la arquitectura de la red reguladora gobernada por los trans-sRNAs AbcR1/2 y NfeR1 y su función en el metabolismo adaptativo de S. meliloti. Esta caracterización incluye la identificación de las proteínas reguladoras que modulan su expresión y la disección, a escala genómica, de sus respectivos interactomas de mRNA. Hemos explorado también el papel de las proteínas de unión a RNA que podrían participar en la regulación mediada por estos sRNAs.

Para ello, hemos empleado técnicas generales de biología molecular y microbiología, siendo de especial interés la implementación de la cromatografía de afinidad de los trans-sRNAs marcados, lo que permite la búsqueda de mRNAs y proteínas que interactúan con ellos, a lo largo de todo el genoma, en condiciones que estimulan la expresión endógena de cada sRNA.

Nuestros datos muestran que AbcR1/2 y NfeR1, sus reguladores transcripcionales y sus mRNAs diana forman parte de un regulon post-transcripcional particularmente extenso, que tiene un gran impacto en la reprogramación metabólica de *S. meliloti* durante su transición simbiótica. Concluimos que el regulador simbiótico LsrB y el factor sigma alternativo RpoH1 son responsables de la transcripción diferencial de AbcR1 y AbcR2, respectivamente. Además, mostramos que la modulación del metabolismo mediado por la regulación post-transcripcional de AbcR1 potencia la capacidad de *S. meliloti* para colonizar el rizoplano de la raíz de la planta, lo que es un rasgo simbiótico biotecnológicamente es muy relevante.

También demostramos que NfeR1 es un sRNA que se induce por estrés de nitrógeno y que su transcripción ocurre desde un promotor complejo, siendo activada por LsrB y reprimida por el regulador maestro de la respuesta a estrés de nitrógeno, NtrC. Este sRNA potencia la fuerza de la respuesta a estrés de nitrógeno de S. meliloti mediante el alivio de la (auto)represión del mRNA bicistrónico ntrBC.

Por otro lado, demostramos que la sintetasa de S-adenosil metionina, el principal donador de grupos metilo, MetK, es una nueva proteína de unión a RNA con funciones no canónicas en la riboregulación mediada por los sRNAs AbcR2 y NfeR1. También buscamos mRNAs diana de AbcR1/2 y NfeR1 que estuvieran sometidos a la regulación por las endoribonucleasas RNasaIII e YbeY, como una aproximación global a la función de éstas en el silenciamiento del RNA.

En conjunto, los datos presentados en esta Tesis describen una extensa red reguladora asistida por RNAs que optimiza el metabolismo simbiótico adaptativo de *S. meliloti* y que se predice que opera en otros rizobios. Debido a que la riboregulación se basa en la plasticidad funcional de los RNAs y la posibilidad de modificación de las interacciones entre nucleótidos complementarios, estas redes reguladoras podrían reprogramarse a diferentes niveles, abriendo así vías aún inexploradas para la ingeniería de biofertilizantes altamente competitivos y la fijación simbiótica de nitrógeno en la práctica agrícolas sostenible.

Abstract

In the last two decades, it has become evident that non-coding protein RNAs assume diverse structural and regulatory functions in all kingdoms of life. High-throughput transcriptome profiling (RNA-Seq)

has uncovered large and heterogeneous populations of small RNAs (sRNAs) with potential regulatory roles in bacteria. These sRNAs act mostly by protein-assisted base-pairing with target mRNAs to fine-tune post-transcriptional reprogramming of gene expression as a major and ubiquitous adaptive trait, contributing greatly to bacterial fitness in fluctuating environments. However, this level of genetic regulation is still poorly explored in most of non-model bacteria. The knowledge about the role of sRNAs in the N-fixing rhizobia-legume symbiosis mostly derives from work on the α-proteobacterium *S. meliloti*, but even in this model plant symbiont only a handful of sRNAs and proteins that assist riboregulation have been characterized. A large class of sRNAs are the so-called *trans*-sRNAs that are differentially expressed from intergenic regions and most commonly modulate translation and/or stability of their target mRNAs by short and discontinuous antisense interactions and ribonucleases recruitment. The homologous ABC Regulators AbcR1 and AbcR2 and the Nodule Formation Efficiency Regulator NfeR1 *trans*-sRNAs exhibit conserved unpaired anti-Shine-Dalgarno (aSD) motifs that are predicted to interact with their target mRNAs blocking translation. Primary characterization anticipates a major impact of these *trans*-sRNAs in the regulation of *S. meliloti* metabolism throughout symbiosis. However, their functions are not yet delineated with detail.

In this Thesis, we have deciphered the architecture of the AbcR1/2 and NfeR1 regulatory network and their function in the *S. meliloti* adaptive metabolism. This characterization includes identification of regulatory proteins that modulate their expression and dissection of their respective mRNA interactomes at a genome-wide scale. We have also

explored the role of RNA-binding proteins that could assist the these riboregulation by these sRNAs.

For that, we have employed general techniques of molecular biology and of microbiology. Remarkably, we have implemented the affinity chromatography of aptamer-tagged *trans*-sRNAs that allow tackling the comprehensive genome-wide profiling of their interactomes (mRNAs and proteins) in growth conditions that stimulate endogenous upregulation of each sRNA.

Our data show that AbcR1/2 and NfeR1 sRNAs, their transcriptional regulators and their target mRNAs are arranged into an exceptionally large and overlapping post-transcriptional regulon that pervasively contributes to *S. melitoti* metabolic reprogramming throughout the symbiotic transition. We concluded that the LysR-type symbiotic regulator LsrB and the alternative σ factor RpoH1 are responsible for the differential transcription of AbcR1 and AbcR2, respectively. Further, we show that AbcR1-mediated post-transcriptional fine-tuning of metabolism enhances the ability of *S. meliloti* to colonize the root rhizoplane, a biotechnologically relevant symbiotic trait.

We also demonstrated that NfeR1 is a nitrogen stress induced sRNA, which is transcribed from a dual-mode promoter activated by LsrB and repressed by the master regulator of the nitrogen stress response, NtrC. This *trans*-sRNA likely strengthens the *S. meliloti* nitrogen stress response by alleviating the (auto)repression of the *ntrBC* bicistronic mRNA.

On the other hand, we demonstrated that MetK, the synthetase of the major methyl donor S-adenosyl methionine, is a novel RNA-binding protein with a non-canonical function in regulation by the AbcR2 and NfeR1 *trans*-sRNAs. We also mined the RNase III and YbeY-dependent

transcriptomes to search for misregulated AbcR1/2 and NfeR1 target mRNAs, as a first global approach to the function of both endoribonucleases in RNA silencing.

Taken together, data presented in this Thesis depict a singularly large RNA network that fine-tunes the *S. meliloti* adaptive symbiotic metabolism and is predicted to operate in diverse rhizobial species. Because riboregulation relies on the functional plasticity of the RNA molecules and on modifiable base pairing interactions, this network could be rewired at different levels, thereby opening yet unexplored avenues for the engineering of highly competitive biofertilizers and symbiotic N-fixation in the sustainable agricultural practices.

Índex

Figures Índex

Table Índex

Abbreviations

C – carbon
N – nitrogen
QS – *quorum sensing*
PHB – polyhydroxibutyrate
AHL – N-acyl-homoserine-lactone
ROS – reactive oxygen species
NCR – Nodule-specific Cysteine Rich peptides
SAM - S-adenosylmethionine
NSR – Nitrogen Stress Response
ABC – ATP-Binding Cassette
SBP – Solute Binding Proteins
RBP – RNA-Binding Proteins
DNA – Deoxyribonucleic acids
RNA – Ribonucleic acids
ssRNA – single-stranded RNA
dsRNA – double- stranded RNA
mRNA – messenger RNA
rRNA – ribosomal RNA
tRNA – transfer RNA
tmRNA – transfer-messenger RNA
snRNA – small nuclear RNA
snoRNA – small nucleolar RNA
sRNA – small RNA
asRNA – antisense RNA
sense-RNA – mRNA-derived sense-sRNA
aSD – anti-Shine Dalgarno
αr – α-proteobacterial sRNA families
TSS – Transcription Start Sites
UTR – Untranslated Region
CDS – Coding Sequence
RBS – Ribosome Binding Site
ORF – Open Reading Frame
IGR – genomic regions between ORFs
RNAP – RNA Polymerase
eGFP – Enhanced Green Fluorescence Protein
PCR – Polymerase chain reaction
RT-PCR – Reverse-transcription PCR
RT-qPCR – Quantitative reverse transcription PCR
RNA-Seq – RNA-Sequencing
dRNA-Seq – differential RNA-Seq
MAPS – MS2 affinity purification coupled with RNA-Seq
MBP – Maltose Binding Protein
RACE – Rapid Amplification of cDNA Ends
SDS – Sodium Dodecyl Sulfate
SDS-PAGE – SDS–PolyAcrylamide Gel Electrophoresis
LC-MS/MS – liquid chromatography mass spectrometry
CoIP – Co-inmunoprecipitation
LB – Lysogeny Broth
TY – tryptone-yeast
MM – minimal medium
Sm – streptomycin
Tc – tetracycline
Km – kanamycin
Gm – gentamycin
IPTG – Isopropyl β-D-1-thiogalactopyranoside
MOPS – morpholino propanesulfonic acid
EDTA – Ethylenediaminetetraacetic acid
TEMED – Tetrametiletilendiamina
DTT – Ditiotreitol
NTP – nucleotides
PVDF – polyvinylidene difluoride membrane
BSA – bovine serum albumin
HEPES – 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
CFU – colony-forming units
SD – standard deviation
CM – Covariance Model
FBA – flux balance analysis

Introduction

The perception of RNA as a mere labile intermediate (messenger RNA; mRNA) between the genetic information from DNA to the function executed by proteins is out of date. Many structural and regulatory functions have been early attributed to these nucleic acids upon the findings of their key involvement in mRNA translation (ribosomal and transfer RNAs; rRNAs/tRNAs), splicing (small nuclear RNAs; snRNAs), post-transcriptional modification (small nucleolar RNAs; snoRNAs), or as a scaffold allowing the assembly of complex molecular machines (e.g., ribosome and spliceosome) [1–3]. Later, the discovery of the catalytic potential of RNA, in the form of ribonucleic enzymes, marked a turn in RNA biology by providing evidence that nanomachines such as the ubiquitous RNase P, the spliceosome, and even the ribosome are themselves ribozymes. Larger ribozymes have also been found to accomplish sophisticated RNA splicing reactions, as described for group II introns [1, 3, 4].

In the last two decades, the development of high-throughput sequencing technologies has also revolutionized our classic view of the prokaryotic transcriptome. Different well-established strategies for the generation and deep-sequencing of cDNA libraries (RNA-Seq) allow for genome-wide mapping of Transcription Start Sites (TSS), processing sites and 3'-ends of the cellular transcripts, thereby uncovering unexpected operon structures and novel transcribed regions systematically overlooked in the primary annotation of bacterial genomes (Fig. 1) [5]. Many of these newly discovered transcripts are not translated into proteins but are mostly involved in the regulation of gene expression, as early reported for similar transcripts mediating RNA silencing in eukaryotes (e.g., miRNAs, siRNAs).

The bacterial non-coding transcriptome typically consists of a heterogeneous population of 50 to 250 nt-long RNA species (small RNAs; sRNAs) with diverse biogenesis pathways and regulatory activity mechanisms. The major families of regulatory sRNAs include *cis*- and *trans*-acting transcripts, which regulate mRNA

levels from genes located in the same locus or transcribed from different loci in the genome, respectively [6].

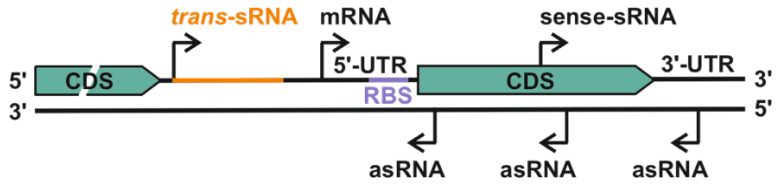


Figure 1. The prokaryotic non-coding transcriptome as revealed by differential RNA-Seq (dRNA-Seq). Identified TSSs (shown by arrows) can be assigned to mRNAs, and the different sRNA types: *trans*-sRNAs, antisense RNAs (asRNAs) and mRNA-derived sense-sRNAs. UTR, Untranslated Region; CDS, Coding Sequence; RBS, Ribosome Binding Site. See text for further details.

1. Activity mechanisms and function of bacterial sRNAs

Cis-acting RNAs include Untranslated Regions (UTRs) of mRNAs and sRNAs transcribed from the strand complementary to the mRNA that they regulate (antisense RNAs; asRNAs). Many 5'-UTRs are riboswitches that directly sense shifts in temperature, pH or intracellular levels of certain metabolites and in turn adopt alternative conformations to modulate transcription termination, translation initiation, RNA processing and stability of the downstream coding sequence (CDS) [7, 8].

The activity of asRNAs, 3'-UTRs and most of the *trans*-sRNAs rely on base-pairing interactions to fine-tune translation and/or turnover rates of their mRNA targets [9–15]. Complementarity between the asRNA and its target is perfect and may extend longer than 200-nt at any mRNA region. Conversely, regulation by *trans*-sRNAs mostly involves short and discontinuous series of complementary nucleotides in both molecules, requiring the assistance of proteins (e.g., Hfq) as RNA matchmakers [16]. These interactions typically occur at the vicinity of the Ribosome Binding Site (RBS) and block mRNA translation (Fig. 2). Regardless of the type of sRNA involved in regulation and the possible effect on translation, formation of RNA duplexes through base-pairing usually promotes degradation

of the mRNA targets by cellular ribonucleases (e.g., RNase E or RNase III) [10, 17].

Nonetheless, it has been also reported positive regulation by both asRNAs and *trans*-sRNAs [15, 18, 19]. Several *trans*-sRNAs act as direct translational activators by an 'anti-antisense mechanism' in the 5' mRNA region to liberate a sequestered RBS (Figure 2). Alternative mechanisms of positive regulation by *trans*-sRNAs involve pairing at the 3'-end of the mRNA or suppression of premature Rho-dependent transcription termination to promote mRNA stability or transcription, respectively [19, 20].

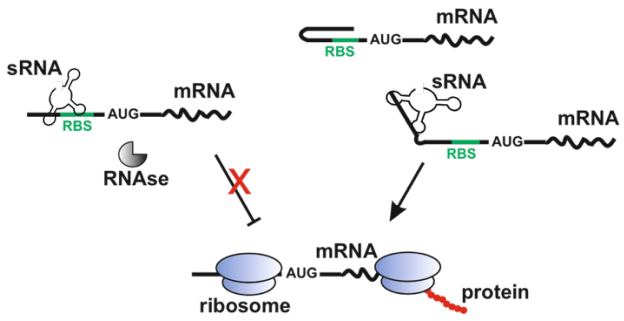


Figure 2. Canonical base-pairing mechanisms of negative (left) and positive (right) post-transcriptional regulation by *trans*-sRNAs.

Instead of acting by base-pairing interactions with mRNAs, several *trans*-sRNAs bind to and antagonize the activity of certain proteins. This is the case of the CsrB family sRNAs that act by target mimicry to outcompete the carbon (C) storage regulator CsrA protein, which is a regulator of translation (Fig. 3) [6, 19].

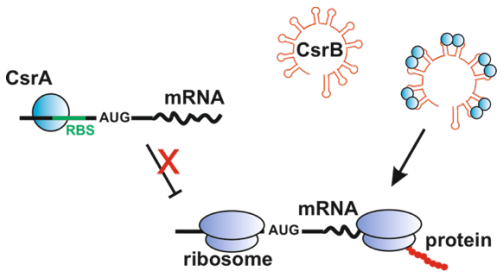


Figure 3. Antagonistic binding of CsrB to CsrA. CsrB forms a large globular ribonucleoprotein complex with CsrA and antagonizes the effects of CsrA.

RNA-mediated post-transcriptional regulation of gene expression is ubiquitous in bacteria, influencing virtually any cellular process, e.g., stress responses, biofilm formation, *quorum sensing* (QS) and different traits of host-microbe interactions[21]. Research on prokaryotic riboregulation has been pioneered in the early post-genomic era by work on the model bacterium *Escherichia coli*. Soon afterwards, studies in related enterobacteria (i.e., *Salmonella* spp.) and other clinically relevant microbes revealed the unprecedented prominent roles of sRNAs in the establishment of host-pathogen interactions[21]. In the last decade, functional RNomics is continuously providing new paradigms about the contribution of sRNAs to the ecological specializations of phylogenetically distant bacteria with complex lifestyles, bringing RNA to the forefront of microbial research.

2. *Rhizobium*-legume symbiosis

Some soil-dwelling species of the large classes of α- and β-proteobacteria, collectively referred to as rhizobia, establish mutualistic symbioses with legumes (Fig. 4). The outcome of these interactions is the organogenesis of specialized nodule structures on the roots or, less frequently, the stems of their specific host plant. Invading rhizobia colonize nodules intracellularly and differentiate to bacteroids that achieve the nitrogenase-mediated reduction of the atmospheric dinitrogen (N_2) to ammonia for the benefit of the plant. N-fixing root nodule

symbioses provide more than half of the combined N incorporated annually into terrestrial ecosystems, rendering plant growth independent of exogenously applied combined N, which is commonly provided to crops in the form of pollutant and costly chemical fertilizers. Besides this doubtless agronomical and ecological significance, the rhizobia-legume symbiosis provides a complex biological experimental model to explore the molecular mechanisms underlying bacterial adaptations during chronic intracellular infection of eukaryotic hosts [22–24].

The saprophytic and symbiotic competence of free-living rhizobia in soil is largely determined by their capacity to cope efficiently with the abiotic variables shaping this environment, e.g., oligotrophy, drought, salinity, or acidity, which are known to negatively influence the physiology of both symbiotic partners and the N fixation process itself [25]. Competitive rhizobial strains actively colonize the rhizosphere of their compatible legume and initiate infection upon synthesis of lipo-chitooligosaccharide signal molecules (i.e., Nod factors) in response to species-specific root exuded flavonoids[23, 26]. Nod factor signaling triggers major nodule developmental pathways (i.e., root hair curling and generation of the nodule primordia within the cortex), and root hair invasion through tubular structures made of plant cell-wall material known as infection threads. Besides Nod factors, other signaling molecules of bacterial origin, such as surface polysaccharides or effector proteins, selectively contribute to the infection of certain legume species by their rhizobial partners[27].

During early symbiotic infection, invading rhizobia elicit a defense response in the host, which is initially featured by the release of reactive oxygen species (ROS) [28]. Compatible rhizobia counteract and survive this oxidative stress by specific mechanisms, being subsequently delivered from the branched infection threads into the cells of the nodule primordia. Inside the plant cells, rhizobia differentiate into N-fixing polyploid bacteroids that end up surrounded by a

membrane of plant origin, thereby generating an organelle-like structure called the symbiosome. In some legume lineages (e.g., Inverted Repeat Lacking Clade from the Papilionoideae subfamily), irreversible terminal bacteroid differentiation is directed by plant-secreted Nodule-specific Cysteine Rich (NCR) peptides [29]. Finally, mature bacteroids sense the microoxic environment inside the plant nodule, which drive expression of gene clusters encoding the nitrogenase. This complex rhizobial lifestyle demands an adaptive flexibility that is supported by both large multipartite genomes and continuous gene expression shifts during host infection [30, 31]. To date, symbiotic regulatory networks in rhizobia have been studied almost exclusively from the perspective of the transcriptional control orchestrated by proteins, i.e., transcription factors and alternative RNA polymerase holoenzymes (σ factors), or specific post-translational modifications, but the underlying post-transcriptional mechanisms are largely unknown [32–35]. sRNA-mediated regulation of gene expression is expected to play major roles in the establishment of these mutualistic symbioses. However, the knowledge about riboregulation in rhizobia is still scarce and rather limited to *Sinorhizobium* (*Ensifer*) *meliloti,* the symbiotic partner of the forage legume alfalfa (*Medicago sativa* L.) and other related *Medicago* species [36–38].

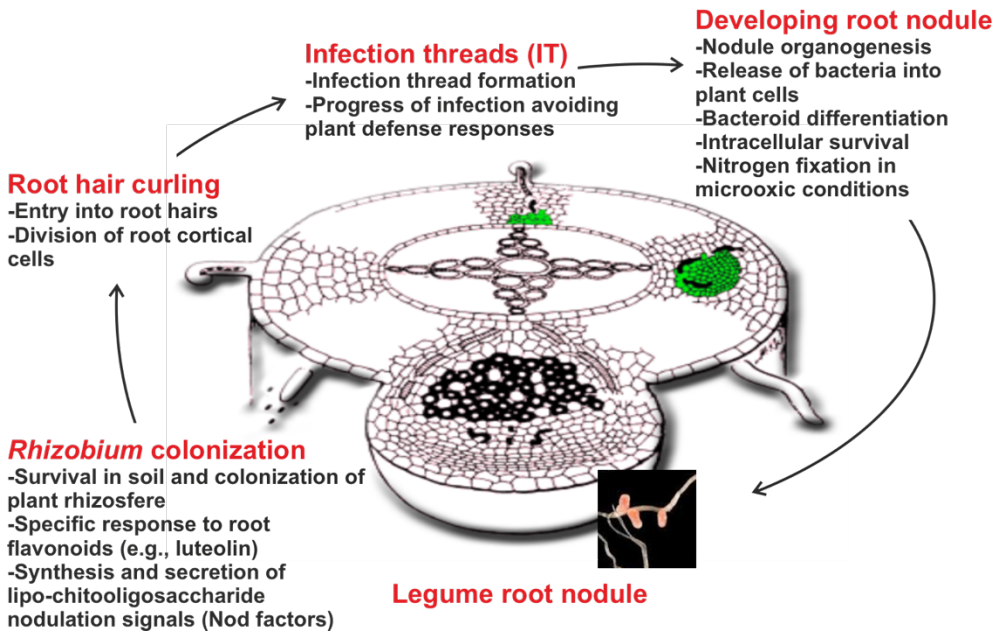


Figure 4. Steps of the *Rhizobium*-legume symbiosis.

3. Deciphering the *S. meliloti* non-coding transcriptome: from comparative genomics to RNA-Seq

Non-protein coding genes largely escape classical genetics screens and the primary annotation of a single bacterial genome sequence, which is essentially limited to the prediction of Open Reading Frames (ORFs), tRNAs and rRNAs. The *S. meliloti* genome (reference strain Sm1021) consists of three replicons: chromosome (3.65 Mb), and the symbiotic megaplasmids pSymA (1.35 Mb) and pSymB (1.68 Mb). The primary genome analysis predicted 6,206 ORFs, 54 tRNAs, 3 rRNA operons, and transfer-messenger RNA (tmRNA) as the only sRNA [39].

Before the advent of high-throughput sequencing, computational comparative genomics was the tool of choice to identify conserved regions with putative functions in the unannotated portions of the genome. Accordingly, pioneering seminal genome-wide searches for sRNAs in *S. meliloti* relied on the

comparison of intergenic sequences (i.e., genomic regions between ORFs; IGRs) from phylogenetically close species to unveil sRNAs of the *trans*-acting class. Genomic comparisons were combined with prediction of other known features of *trans*-sRNAs namely, association with orphan transcription signatures (promoter motifs and/or Rho-independent transcriptional terminators) and conservation of thermodynamically stable secondary structures using different computational tools e.g., QRNA, sRNAPredict2, or RNAz algorithms (Fig. 5) [40–42]. Northern blot hybridization of total RNA, RACE (Rapid Amplification of cDNA Ends) mapping of transcripts boundaries and/or microarray probing experimentally confirmed that a few dozens of these candidate IGRs did express sRNA species from independent transcription units. In the absence of further functional insights, these *trans*-sRNAs were initially referred to as either Smr [40], Sra [41] or Sm [42].

Straightforward experimental identification of TSSs associated to CDS, UTR, and non-coding RNA genes in the prokaryotic genomes is now feasible with the implementation of strand-specific differential RNA-Seq (dRNA-Seq) [43] or Cappable-Seq [44] strategies. In particular, dRNA-Seq surveys rediscovered the early-identified sRNAs in *S. meliloti*, further uncovering the complexity of the transcriptome by adding hundreds of unknown *trans*-sRNAs, as well as thousands of mRNA-derived sRNAs and asRNAs (referred to as Smel, SM_ncRNAs, or SM_asRNAs) (Fig. 5) [45–47]. Other RNA-Seq studies are conceived to profile specific subpopulations of cellular transcripts supposedly enriched in sRNAs, e.g., RNA species co-immunoprecipitated with the bacterial RNA chaperone Hfq [48, 49]. However, this approach resulted in a minor addition to the sRNA landscape revealed by deep-sequencing of *S. meliloti* total RNA [50].

Prokaryotic gene prediction pipelines such as EuGen-P have incorporated the novel gene structural features uncovered by dRNA-Seq for the accurate reannotation of the *S. meliloti* genome (strains Sm1021 and Sm2011) [46, 47].

Similar RNA-Seq and *in silico* workflows have been used for the primary annotation of sRNA genes in other rhizobia (e.g., *Rhizobium etli* CFN42, *Bradyrhizobium japonicum* USDA110) [51–53]. Even though dRNA-Seq mostly serves annotation purposes, comparison of transcripts levels in some datasets identified differentially expressed sRNAs in free-living and nodule endosymbiotic bacteria [45, 47, 50, 53]. In this regard, it is noteworthy the identification of nodule-expressed sRNAs by RNA-Seq profiling of RNA derived from each developing zone of indeterminate nodules induced on the model legume *M. truncatula* by Sm2011 [54]. On the other hand, comprehensive mapping of TSS in *S. meliloti* has facilitated the prediction of motifs putatively recognized by alternative σ factors such as RpoE2 (σ^{E2}) or RpoN (σ^{54}) in the promoter regions of some of the identified sRNAs, thus placing these RNA regulators in major stress response and/or symbiotic regulons [46, 47]. The integration of the updated genome annotation files, primary expression profiles and promoter predictions provides a solid reference resource for the investigation of sRNA function in plant symbiotic bacteria.

4. Conservation of *S. meliloti* sRNAs: α-proteobacterial (αr) sRNA families

Reverse comparative RNomics contributes to unravel sRNA function by identifying either functionally characterized homologs to the query transcripts or conservation patterns potentially related to the lifestyle of phylogenetically close bacterial species. Homology predictions typically rely on stochastic Covariance Models (CMs) capturing both sequence and secondary structure (folding) conservation from a multiple alignment to the query sRNA. CMs can be automatically generated by INFERNAL, which builds RNA families collected by the Rfam database (https://rfam.xfam.org/) [55]. Regardless of their assignment to specific family models, a handful of chromosomally encoded rhizobial sRNAs

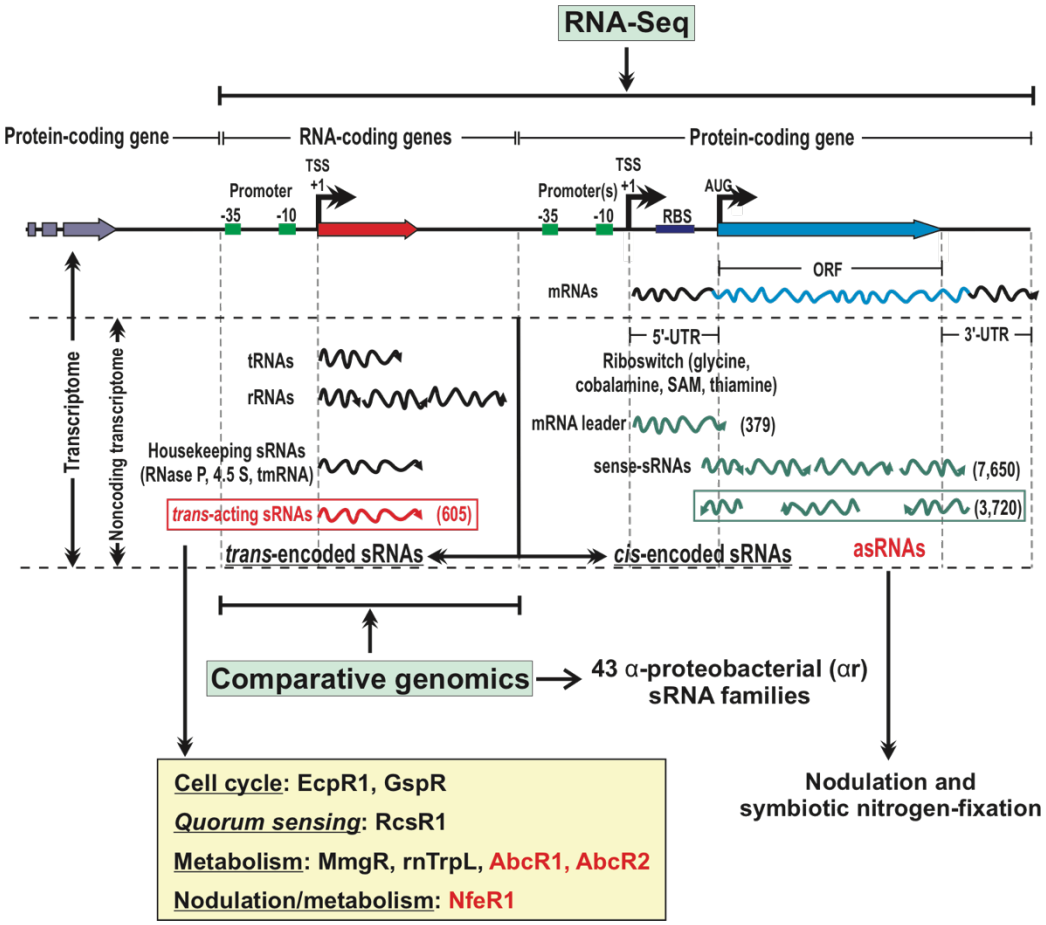


Figure 5. The *S. meliloti* non-coding transcriptome. Different non-coding RNAs revealed by comparative genomics and/or RNA-Seq in *S. meliloti* include: tRNAs, rRNAs, *trans*-sRNAs and *cis*-RNAs (i.e., riboswitches, asRNAs, sense-RNAs). The number of RNAs in each category is in brackets. Box shadowed yellow summarizes the functionally characterized *trans*-sRNAs. The *trans*-sRNAs studied in this Thesis are in red.

that occur ubiquitously and exert housekeeping functions in bacteria can be unequivocally identified by the sole comparison of the primary nucleotide sequence. This set includes:

- RNase P (Rfam family model RF00010), which is the ribozyme that cleaves off 5'-extra sequences on tRNA precursors [56].

- tmRNA or SsrA (RF00023), which has a dual function as transfer and mRNA, operating in *trans*-translation for stalled ribosome recycling [57]. Of note, tmRNA has been the only rhizobial sRNA identified experimentally by a random

mutagenesis screen for symbiotic genes in *B. japonicum* [58]. The *S. meliloti* tmRNA homolog is expressed in a growth- and stress-dependent manner as an unstable precursor, further processed into two readily detectable 214 and 82 nt-long RNA species likely derived from the mRNA and tRNA domains of the primary transcript [59].

- The 4.5S RNA (RF00169), which associates with the multidomain Ffh protein, in the bacterial SRP (small Signal Recognition Particle) ribonucleoprotein complex, universally required for co-translational protein targeting [60].

- The 6S RNA or SsrS (RF00013) RNA, which relies on an open promoter-like structure to counteract the activity of the σ^{70} RNA polymerase holoenzyme by a target mimicry mechanism, thus contributing to the transcriptional reprogramming during transition from exponential to stationary growth [61].

CM-based phylogenetic distribution of other 57 *S. meliloti trans*-sRNAs without associated functional evidence has been also addressed by several studies [62–65]. These analyses collected this set of *trans*-sRNA input sequences into 43 family models, indicating that several of these sRNA loci occur with different levels of paralogy in the *S. meliloti* genome. Distribution of the majority of these families was restricted to members of the family Rhizobiaceae within α-proteobacteria, including *Agrobacterium* species. Only a few chromosomally encoded sRNAs are conserved beyond Rhizobiaceae, occurring in representatives of Brucellaceae or even Bartonellaceae, Xanthobacteriaceae, Beijerinckaceae or Bradyrhizobiaceae. In contrast, sRNA loci mapping to *S. meliloti* symbiotic plasmids pSymB or pSymA are almost unique to *Sinorhizobium* spp. and *S. meliloti*, respectively, further supporting that these two megaplasmids mostly encode accessory functions, with pSymA being the most recently acquired replicon in the genus *Sinorhizobium*. Overall, the distribution patterns of these α–proteobacterial sRNA families hint at a major contribution of vertical inheritance

and frequent ancestral duplications events to the evolution of these class of riboregulators in α-rhizobia.

5. Regulation of symbiotic genes by asRNAs

Like in other bacterial species, pervasive antisense transcription also occurs in *S. meliloti* and several studies anticipate a broad impact of asRNAs on both the free-living and host-associated lifestyles of this bacterium [45, 46, 66, 67]. RNA-Seq surveys have revealed that *S. meliloti* expresses more than 3,000 asRNAs that are linked to ~35% of the predicted protein-coding genes [45, 46]. However, the functional significance of asRNA-mediated regulation has not been investigated further at genome-wide scale or for symbiotic genes.

The first characterized asRNA was the so-called IncA, which is conserved in related α-rhizobia and controls the replication gene *repC* of pSymA and pSymB, thus acting as an incompatibility determinant of the large *repABC* α–proteobacterial plasmids [68]. A systematic screening strategy to identify novel RNA regulators of N fixation revealed that both asRNAs and *trans*-sRNAs are overrepresented (41% and 24%, respectively) in *S. meliloti* pSymA symbiotic megaplasmid, with asRNAs particularly biased to nodulation (*nod*) and N fixation (*nif/fix*) genes. Further characterization of seven pSymA-borne asRNAs showed that most accumulate in response to different biological conditions (e.g., upon addition of the flavonoid luteolin, heat and cold shocks, microoxic conditions or into nodules) anticipating diverse functions [66]. Among them, SMa_asRNA_244, SMa_asRNA_277 and SMa_asRNA_279 were identified as antisense transcripts of the *nodD2* 3'-UTR, *nifE* CDS and *nifE* 3'-UTR, respectively. Remarkably, overexpression of another RNA termed SmelA031, which is antisense to the *nifK* CDS, has been shown to moderately affect N fixation in alfalfa nodules. These findings are only the first evidence of asRNA

regulation of symbiosis in rhizobia, but hundreds of rhizobial asRNAs await functional characterization.

6. Assigning functions to the *S. meliloti trans*-sRNAs

In prokaryotes, a large fraction of non-coding RNAs are catalogued as *trans*-sRNA regulators [21]. Biology of bacterial *trans*-sRNAs is essentially featured by their differential accumulation in response to environmental signals and target mRNA regulation by protein-assisted base-pairing mechanisms (Fig. 6). Therefore, primary studies to unravel *trans*-sRNA functions must necessarily address expression profiling, transcriptional regulation, gain- and loss-of-function phenotyping, identification of mRNA targets and targeting motifs, and characterization of the associated proteins (e.g., RNA chaperones or RNases).

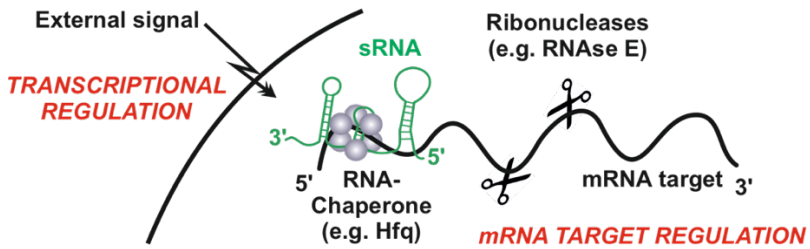


Figure 6. Key features of bacterial *trans*-sRNA biology. Differential regulation, target recognition and association with RNA chaperones and ribonucleases.

6.1. Experimental approaches

Expression of *trans*-sRNAs can be tracked by classical molecular-genetics methods that include the use of promoter-reporter fusions and probing of total RNA on Northern blot membranes [69]. On the other hand, target identification typically relies on comparative genomics-based predictions of most probable conserved sRNA-mRNA base-pairing interactions (e.g., CopraRNA, IntaRNA, Target RNA algorithms) followed by experimental validation of the candidates using a two-plasmid genetic reporter assay *in vivo* [37, 69–71]. The latter is based

on the constitutive co-expression from compatible plasmids in the same cell of the sRNA of interest and a translational fusion of the putative mRNA target to *eGFP* (Enhanced Green Fluorescence Protein), so that fluorescence of the reporter strains is related to sRNA-mediated translational repression or activation of the mRNA [37, 69]. For constitutive sRNA expression in *S. meliloti*, plasmid pSRK_C was generated in our laboratory by engineering the mid-copy pSRK [36, 38]. To overcome problems with pleiotropic effects or weak overexpression, we later combined the Isopropyl β-D-1-thiogalactopyranoside (IPTG)-inducible native system of pSRK vectors with the well-known *S. meliloti sinR-sinI* genes involved in QS [69]. This achieves strong pulse overexpression of the sRNA under study (Fig. 7). This inducible system has been proved to be more accurate for sRNA overexpression and target identification [72, 73]. Co-expression of relevant sRNA and mRNA mutant variants in the assays enables the precise mapping of the interacting motifs in both molecules.

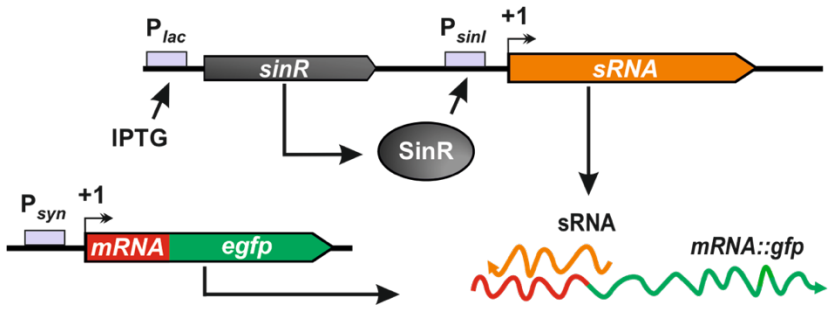


Figure 7. Reporter assay for the genetic dissection of regulatory sRNA-mRNA antisense interactions in *S. meliloti*. The sRNA and a translational fusion of its putative target mRNA to *eGFP* are co-expressed from compatible plasmids in the same bacterial cell lacking the sRNA locus. Fluorescence of the IPTG-induced and uninduced cultures is then scored quantitatively to assess sRNA-mediated post-transcriptional regulation of the target mRNA. P_{syn} is a constitutive promoter with a consensus σ^{70} signature.

Alternatively, the interactome of a *trans*-sRNA can be characterized on a genome-wide scale by profiling either transcriptome alterations upon pulse overexpression of the tested sRNA or the subset of mRNAs picked up by affinity chromatography using an aptamer-tagged (e.g., MS2) version of the sRNA as a

bait (Fig. 8) [74, 75]. Remarkably, this latter approach is also suitable to capture *in vivo* assembled sRNA-protein complexes, which can be further characterized by mass spectrometry analyses[69].

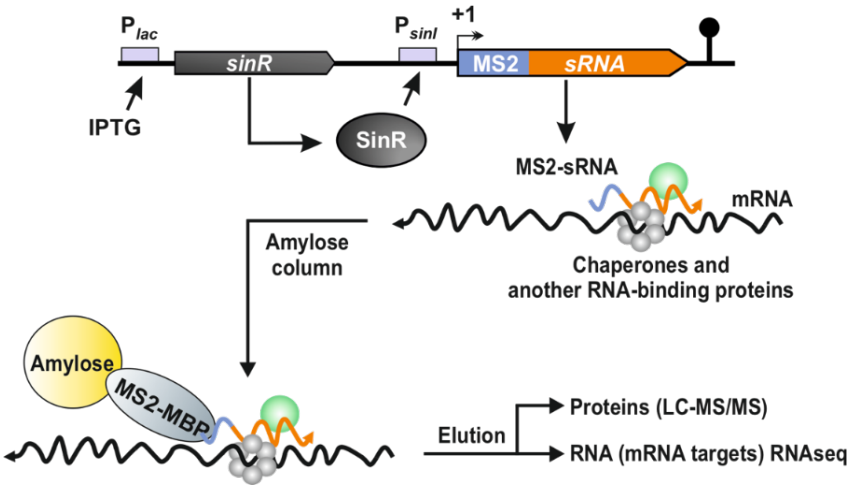


Figure 8. Characterization of the *trans*-sRNA interactome by MS2-affinity chromatography. MS2-MBP (Maltose Binding Protein) is immobilized in an amylose column. Lysates from bacteria expressing the tagged MS2-sRNA are applied to the amylose column. After several column washes, addition of maltose enables the elution of the bound RNA-protein complexes.

A combination of these strategies has been used to gain insights into the function and activity mechanisms of a handful of *S. meliloti trans*-sRNAs regulating free-living and symbiotic rhizobial traits such as cell-cycle, QS, metabolism, or nodule development.

6.2. sRNAs involved in regulation of cell cycle and QS

Mechanisms to control chromosome replication and cell division in fluctuating environments are of critical relevance for bacterial survival and establishment of symbiosis with leguminous plants (e.g., during bacteroid differentiation) [54, 76, 77]. The sRNAs EcpR1 (formerly SmelC291, sra33 or smrC10), and GspR (SmelC775) have been related to the adjustment of cell cycle progression under adverse conditions by silencing of some key genes involved in

cell cycle regulation [73, 78]. EcpR1 overexpression results into an Elongated Cell Phenotype, along with cell filamentation and genome endoduplication whilst over-accumulation of GspR hampers cell growth (Growth Stop Phenotype), consistently with cell cycle progression defects. EcpR1 and GspR knock-out mutants do not show clear growth or symbiotic phenotypes. However, both are outcompeted by their parent strains when co-cultured, suggesting that EcpR1 and GspR confer a fitness advantage to bacteria.

QS also plays an important role in regulating and coordinating the interaction between the symbiont and its host. The *sinR-sinI* system, one of the three QS systems present in *S. meliloti*, controls the production of the symbiotically active exopolysaccharide EPS II and the expression of genes for motility, chemotaxis, N fixation, and transport of metal and small molecules [79, 80]. One *S meliloti* sRNA has been reported to post-transcriptionally regulate *sinI*, the gene encoding the synthase of the QS autoinducer N-acyl-homoserine-lactone (AHL). It was called RcsR1 (initially referred to as SmelC587 or sm104) since its expression changes with temperature and salt stress (Rhizobial Cold and Salinity stress Riboregulator 1). RcsR1 impairs SinI translation by outcompeting RNase E at a cleavage site within the *sinI* 5'-UTR [81, 82]. RcsR1 was later shown to derive from transcriptional attenuation of one of the three tryptophan (*trp*) biosynthesis operons, *trpE(G)*, and to directly interact with the polycistronic *trpDC* mRNA to repress translation. Accordingly, it was re-named rnTrpL (trpL-derived attenuator RNA) [83]. RcsR1/rnTrpL co-regulation of QS and tryptophan biosynthesis is a proof of the *trans*-sRNA versatility linked to their capacity to target multiple mRNAs.

6.3. Riboregulation of metabolism and nodulation

Rhizobia must cope with a broad range of limiting nutrients in soil and specific nutritional needs throughout symbiosis with their cognate legume hosts.

Therefore, nodulation competitiveness first depends on the adequate regulation and coordination of efficient nutrient uptake and metabolism. Thus, it is not surprising the large number of ABC (ATP-Binding Cassette) transporters encoded by rhizobial genomes [39, 84, 85]. Nutrient uptake through ABC transporters primarily relies on periplasmic Solute Binding Proteins (SBP) that determine substrate-specificity and are induced by its own ligand, typically two or more molecules [86]. Acquisition of amino acids, peptides, metal ions or sugars through these systems impacts important cellular process such as differentiation, infection or conjugation [86]. Remarkably, a selective suite of ABC transporters is expressed in nodules induced in alfalfa by *S. meliloti*, thus supporting a specialized nutrient exchange between microsymbiont and host [87].

Pervasive RNA regulation of nutrient uptake was first described in *E. coli* and *Salmonella* by characterization of the GcvB sRNA, which regulon potentially includes ~1% of all mRNAs in these bacteria. The GcvB mRNA interactome has a prevalence of mRNAs from ABC transporters of amino acids and short peptides, but also contains mRNAs for amino acid biosynthesis and transcriptional regulation [88]. Functionally analogous riboregulators were later discovered in the α-proteobacterial species *A. tumefaciens* and *B. abortus* and were termed AbcR sRNAs (ABC Transporter Regulator) [89–93]. *S. meliloti* Sm1021 encodes three AbcR homologs that belong to the αr15 sRNA family, whose members have partial homology to the SuhB sRNA (Rfam model RF00519) [63]. However, expression of only two, named AbcR1 and AbcR2 (formerly referred to as either SmelC412/11, sm3'/3, sra41 or smrC16/15), has been readily detected by Northern blot hybridization. AbcR1 and AbcR2 are Hfq-dependent transcripts that exhibit growth- and stress-dependent expression in free-living *S. meliloti* bacteria, respectively. Interestingly, AbcR1 has been also detected in the so-called invasion zone of alfalfa nodules but both sRNAs are dispensable for a wild-type endosymbiosis. Computational predictions and preliminary experimental data

suggest that AbcR1/2 massively regulate nutrient uptake by inhibiting translation of large and overlapping arrays of mRNAs mostly encoding the periplasmic component of ABC transport systems [50, 91].

Successful bacteroid differentiation and symbiotic N fixation requires proper regulation of C and N metabolism in the rhizobial partner. C storage in nodules occurs in the form of polyhydroxybutyrate (PHB) granules, which are produced during the invasion process and degraded during bacteroid differentiation, being critical for symbiotic performance[94]. The Hfq-dependent sRNA MmgR (formerly SmelC689) is required to limit intracellular accumulation of PHB granules under conditions of C surplus (Makes More Granulates RNA) and within nodules. In fact, *mmgR* deletion results in accumulation of both PHB phasin proteins (PhaP1 and PhaP2). Interestingly, MmgR is induced upon N-limiting condition by the global N regulator NtrC while it is repressed by AniA, a C flow regulator, suggesting a role of this sRNA in sensing both the N and C status of the cell [94, 95].

Several studies addressing gene expression in nodules have identified highly expressed sRNAs in endosymbiotic rhizobia, thus anticipating riboregulation of bacterial symbiotic traits [47, 51, 54, 96]. However, in most cases expression in nodules concurs with the stress-induced accumulation of the sRNA in free-living rhizobia. This is the case of some *S. meliloti trans*-sRNAs mentioned above, i.e., AbcR1, MmgR, EcpR1 or GspR1, whose knock-out has none or modest impact in symbiosis [73, 78, 91, 94]. To date, phenotyping of knock-out mutants has revealed putative symbiotic functions for only one chromosomally encoded *S. meliloti* regulatory *trans*-sRNA, identified in early genome-wide screens as Smr14C2 or SmelC397, and later renamed NfeR1 (Nodule Formation Efficiency RNA) to reflect this fact [45, 63, 97]. This sRNA belongs to the αr14 sRNA family, and unlike MmgR and AbcR1/2, it is Hfq-independent. NfeR1 is upregulated in stressed and symbiotic bacteria and its loss-of-function

compromises nodule development and overall symbiotic efficiency of *S. meliloti* on alfalfa roots [97]. Interestingly, computational predictions suggest that NfeR1 might share activity mechanism and a set of ABC transporter target mRNAs with AbcR1/2, anticipating that all three sRNAs govern a dense overlapping regulon for the post-transcriptional regulation of *S. meliloti* metabolism during the symbiotic transition.

7. Proteins assisting sRNA activity.

Proteins involved in RNA activity and metabolism are universal players in riboregulation at different levels [98–101]. Cellular ribonucleases with diverse substrate preference are the effectors of sRNA-mediated post-transcriptional gene silencing [102, 103]. Decay of mRNA upon antisense interaction with regulatory RNAs commonly involves the prevalent prokaryotic endoribonucleases RNase E and/or RNase III, which are specific to single- and double-stranded RNAs (ssRNA and dsRNA), respectively [104]. Rhizobial genomes encode a set of more than 20 ribonucleases, but for the vast majority there is a lack of functional information[105]. Nonetheless, several reports have anticipated the involvement of *S. meliloti* RNase E in *trans*-sRNA stability and regulation [73, 82, 106, 107]. *S. meliloti* RNase III has been biochemically and genetically characterized [108]. Its catalytic features resemble those of the *E. coli* ortholog but with different requirements for optimal activity that could be related to the rhizobial lifestyle.

The gene annotated as *SMc01113* in the chromosome of the *S. meliloti* Sm1021 is almost ubiquitous in bacteria and has been included in the proposed minimal prokaryotic genome [109]. It encodes a putative metal-dependent hydrolase, later called YbeY, which in *E. coli* acts as specific ssRNA endoribonuclease. Biochemical characterization of *S. meliloti* YbeY also supports a universal role of this protein as RNase [110–112], which is strikingly competent for cleaving both ssRNA and dsRNA [111]. Transcriptome alterations of the *S.*

meliloti YbeY mutant, as revealed by microarray hybridization, predict a role of this enzyme in *trans*-sRNA and asRNA mediated silencing of genes involved in nutrient uptake and symbiotic N fixation, respectively [111, 113].

Some RNA-Binding Proteins (RBP) act as chaperones that promote unwinding of complex secondary structures, e.g. Cold-Shock Proteins (CSPs) [114] or base-pairing between regulatory sRNAs and their targets, e.g., the well-characterized Hfq (Fig. 6) [115] and the recently discovered ProQ [116]. Of note, Hfq is encoded by 55% of the bacterial genomes sequenced so far [117]. Besides acting as RNA matchmaker, Hfq has been shown to stabilize widely diverse RNA species, including asRNAs, *trans*-sRNAs, mRNAs and tRNAs. In *S. meliloti*, *trans*-sRNAs and asRNAs were underrepresented in the genome-wide profiling of RNA species co-immunoprecipitated (CoIP-RNA) with a functional FLAG epitope-tagged Hfq variant [50, 118], anticipating that in the legume symbionts other alternative RNA chaperones may also assist sRNA activity. In this regard, a recent affinity chromatography-based profiling of the proteomes associated to EcpR1, AbcR2 and NfeR1 identified the synthetase of the major methyl donor S-adenosylmethionine (SAM), MetK, as common binding partner of the three stress-induced sRNAs [119]. However, the role of this metabolic enzyme as novel RBP in riboregulation remains to be investigated.

.

Objectives

RNA-mediated regulation of gene expression is ubiquitous in all living organisms and contributes to fine-tune most, if not all, cellular processes in prokaryotes. Metabolic reprograming is a major bacterial adaptive trait required for efficient colonization of the soil, rhizosphere, and nodule environments by N-fixing legume symbionts. Primary characterization of AbcR1, AbcR2 and NfeR1, anticipate a major impact of these *trans*-sRNAs in the regulation of *S. meliloti* metabolism throughout symbiosis with its legume host, alfalfa. However, their functions are not yet delineated with detail.

In this Thesis, we have dealt with the characterization of the architecture of the AbcR1/2 and NfeR1 regulatory networks and their function in the *S. meliloti* adaptive metabolism with the following specific objectives:

O1. To decipher the transcriptional regulation and activity mechanisms of these *trans*-sRNAs, and dissect their respective mRNA interactomes at a genome-wide scale.

O2. To validate MetK as novel RNA-binding protein and address its contribution, and that of the endoribonucleases YbeY and RNase III, to regulation by AbcR1/2 and NfeR1.

Material and Methods

1. Microbiology techniques

1.1. Culture conditions.

Culture media. *E. coli* strains were routinely grown in Lysogeny Broth (LB) medium at 37 ºC [120], and rhizobia in either complex tryptone-yeast (TY) or defined mannitol/glutamate minimal medium (MM) media at 30 ºC [121, 122]. Media were prepared in deionized water as follows:

LB		**TY**	
NaCl	5 g/l	$CaCl_2 \cdot 2H_2O$	0.9 g/l
Tryptone	10 g/l	Tryptone	5 g/l
Yeast extract	5 g/l	Yeast extract	3 g/l

To solidify LB and TY, 1.6 or 1.5% agar (Panreac) was added, respectively. The media was autoclaved at 120 ºC for 20 min.

MM	
Potassium glutamate	1.1 g/l
Mannitol	10 g/l
K_2HPO_4	0.3 g/l
KH_2PO_4	0.3 g/l
$MgSO_4 \cdot 7H_2O$	0.15 g/l
$CaCl_2 \cdot 2H_2O$	0.05 g/l
NaCl	0.05 g/l
$FeCl_3 \cdot 6H_2O$	0.006 g/l
Biotin	0.2 mg/l
Calcium Pantothenate	0.1 mg/l

For preparation of MM, the pH was adjusted to 7.0 and the medium was autoclaved at 120 ºC for 20 min. Salts are added from 100-fold concentrated stock solutions, which are prepared in deionized water and autoclaved at 120 ºC for 20 min. Exceptionally, $FeCl_3$ is added from a 10000-fold concentrated stock solution that is sterilized by filtration using filters of 0.2 μm pore size. To solidify MM, 1.6%-purified agar (Oxoid) was added. Vitamins are added after MM autoclaving from a filtered 1000-fold concentrated stock solution, which is prepared in ultrapure water.

Antibiotics. The following antibiotics were added when required: streptomycin; Sm (600 μg/ml), tetracycline; Tc (10 μg/ml), kanamycin; Km (50 μg/ml for *E. coli* and 180 μg/ml for *Sinorhizobium* strains) and gentamycin; Gm (8 μg/ml for *E. coli* and 30 μg/ml for rhizobia). For growth in liquid media the antibiotic concentration was reduced to 50%.

Addition of antibiotics to the culture media was done from 100-fold concentrated stock solution (referred to rhizobia) that were prepared in ultrapure water or 50% absolute ethanol (Tc). Solutions were sterilized by filtration using filters of 0.2 μm pore size.

Growth conditions. Exponentially and stationary growing *S. meliloti* bacteria were obtained by incubation in the selected liquid medium to OD_{600} 0.6-0.8 and 2-2.4, respectively, in an orbital shaker (170 rpm). Expression of stress-dependent sRNAs was assessed in exponentially growing bacteria in MM further cultures for 1 h upon salt (400 mM NaCl) or heat (42 ºC) shocks.

To generate microaerobiosis, 30 ml cultures in 250 ml Erlenmeyer flasks grown in TY medium to OD_{600} 0.6 were fluxed with a mixture of 2% Oxygen/98% N for 10 min and incubated at 60 rpm for a further 4 h in that atmosphere [123, 124].

Conservation of bacterial cultures. Freezing was used for the long-term preservation of bacterial cultures. For maintenance of cell viability during the

storage period, 25% (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 ºC.

Bacterial strains. Wild-type bacterial strains used in this work are listed in Table 1.

Table 1. Wild-type bacterial strains used in this Thesis.

Strain	Relevant characteristics	Reference/Source
S. meliloti		
Sm2011	SU47 derivative; Sm^r	J. Denarie (C.N.R.A., Versailles)
Sm1021	SU47 derivative; Sm^r	[125]
Sm2B3001	Sm2011 *expR*$^+$ derivative; Nal^r, Sm^r	[126]
Sm2019	Sm2B3001 derivative; markerless Δ*expR,* Δ*sinRI*; Nal^r, Sm^r	M. McIntosh (SYNMIKRO, Marburg)
Sm2020	Sm2019 Δ*abcR1*/*abcR2*/*nfeR1* derivative	This work
E. coli		
DH5α	F$^-$*endA1glnV44thi 1recA1relA1gyrA96deoRnupGpurB20* φ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169, hsdR17($r_K^-m_K^+$), λ$^-$	Bethesda Research Lab
S17-1	*recA pro hsdR RP4-2-Tc::Mu-Km::Tn7*	[127]
BL21(DE3)	*E. coli str. B F– ompT gal dcmlonhsdS$_B$($r_B^-m_B^-$) λ(DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5]) [malB$^+$]$_{K-12}$(λ^S)*	Novagen

1.2. Methods for mobilization of exogenous DNA

Preparation of *E. coli* competent cells. To prepare *E. coli* DH5α and S17-1 cells to capture exogenous DNA, bacteria were subjected to a chemical treatment with RbCl as follows.

Bacteria were cultured in 100 ml of LB medium to OD_{600} 0.4. Growth was then stopped by incubating the culture on ice for 15 min. Bacteria were then pelleted by centrifugation (6000 rpm, 10 min, 4 ºC), and resuspended in 32 ml of sterile, pre-cooled RF1 solution (per 100 ml: 1.2 g RbCl; 0.99 g $MnCl_2 \cdot 4H_2O$;

0.294 g potassium acetate; 0.15 g $CaCl_2 \cdot 2H_2O$; 11.9 g glycerol; pH 5.8). Bacterial suspensions were incubated on ice for 15 min and then centrifuged at 6000 rpm for 10 min at 4ºC. The supernatant was removed and the cells were resuspended in 4 ml of pre-cooled RF2 solution (per 50 ml: 0.1046 g morpholino propanesulfonic acid (MOPS); 0.06 g RbCl; 0.55 g $CaCl_2 \cdot 2H_2O$; 5.95 ml glycerol; pH 6.8). The bacterial suspensions were aliquoted into 100 µl aliquots in pre-cooled tubes and immediately frozen in liquid N_2. Competent cells were kept for a limited period at -80 ºC. The transformation efficiency ranged from 1 to $5 \cdot 10^5$ cells/µg DNA.

In case of *E. coli* BL21(DE3), bacteria were subjected to physical treatment to achieve electrocompetent cells as follows.

Bacteria were cultured in 500 ml of LB medium to OD_{600} 0.5. Growth was then stopped by incubating the culture on ice for 20 min. After complete removal of the supernatant with a pipette, a cryoprotective agent (glycerol) was added. Cells were collected at 6000 rpm for 15 min at 4 ºC, resuspended in 100 ml of 10% cold glycerol and repeatedly centrifuged. Then, cells were resuspended in 20 ml of glycerol. Pellets were resuspended in 2 ml of 10% glycerol and cells were distributed into 50 µl aliquots that are immediately frozen in liquid N_2 and stored at -80 ºC. The transformation efficiency was approximately 10^6 cells/µg DNA.

***E. coli* transformation.** RbCl competent cells were thawed on ice for 20 min. DNA was added under sterile conditions and incubated with the cells for further 25 min. The cells were subjected to a thermal shock at 42 ºC for 90 s, and then returned to ice for 5 min. Fresh LB (900 µl) was added and cells were incubated for 1 h at 37 ºC. After this, the bacterial suspensions were dispensed on selective LB agar.

On the other hand, electrocompetent cells were thawed on ice for 20 min prior to addition of dialyzed DNA. Dialysis was achieved using 0.025 mm nitrocellulose filters. Cell suspensions were transferred to an electroporation

cuvette and subjected to an electrical pulse (1800 V for 3-5 ms) in an Eppendorf2510 electroporator. Fresh LB (900 µl) was added and cells were incubated for 1 h at 37 ºC. After this, the bacterial suspensions were dispensed on selective LB agar.

Transformation of *S. meliloti* with plasmid DNA by conjugation. The mobilization of plasmids to *S. meliloti* was achieved by biparental matting using *E. coli* S17-1 as donor strain. For that, the recipient *S. meliloti* strain was grown in TY agar for 48 h and *E. coli* S17-1 carrying the selected plasmid DNA was grown in LB agar for 24 h. Then, cells were mixed in a ratio 3:1 (*S. meliloti* : *E.coli*) in 300 µl sterile 0.9% NaCl solution. A drop of the cell suspension (10 µl) was incubated in TY agar for 16 h at 30 ºC. Transconjugants were selected on TY agar supplemented with the appropriate antibiotics. Visible colonies (typically after 3 days) were streaked in a new TY agar with antibiotics and, in parallel, in a solid Endo-Agar™ (Difco) for verification of possible contaminations by *E. coli.*

1.3. Construction of *S. meliloti* mutants

Selection of allelic exchange and co-integration. The introduction of genomic modifications in *S. meliloti* was performed by double recombination according to the method described for Gram negative bacteria [128]. For that, a 2-1.6 kb fragment of the modified/deleted gene was cloned into the suicide plasmid pK18*mobsacB* (see Molecular biology techniques), which was conjugated into *S. meliloti*. Transconjugantts carrying plasmid insertion into the genome were selected by growing on TY agar supplemented with both Km and 10% sucrose. To select double recombinants, colonies showing sucrose sensitivity were cultured in liquid TY medium to OD_{600} 1. Dilutions from this culture were dispensed on TY agar containing 10% sucrose. Visible colonies were replicated in TY with or without Km to further select Km-sensitive bacteria. Loss of the

wild-type gene in the selected colonies was subsequently confirmed by PCR and, in some cases, by whole genome sequencing.

1.4. Fluorescence reporter assays

pBB-*eGFP* and pR-*eGFP* plasmids were used to generate transcriptional or translational fusions to the *eGFP* reporter, respectively. For the transcriptional fusions, the promoter region of the gene of interest, extending up to its TSS, was cloned upstream to the 5'-UTR and CDS of *eGFP*. For the translational fusions, the 5'-region of the mRNA of interest, from its TSS to the first codons (typically 12-15), was cloned in frame to the truncated *eGFP* CDS. Plasmids were transferred by biparental conjugation to *S. meliloti* strains.

Cultures of the reporter strains (100 µl) were transferred to a 96 well microtiter plate for measurement of OD_{600} and fluorescence (excitation 485 nm, emission 520 nm) in a Thermo Scientific™ Varioskan™ LUX multimode microplate reader. Fluorescence values were normalized to the culture OD_{600} (Fl/OD_{600}).

In the two-plasmid genetic reporter assay was used for *in vivo* to validate target mRNAs candidates of a given sRNA, plasmids expressing the sRNA in an IPTG-dependent manner were mobilized to a *S. meliloti* strain harboring the mRNA translational fusion in pR-*eGFP*. Three double transconjugants were grown for 48 h to stationary phase in TY. Aliquots of 5 µl of each pre-cultures were inoculated in TY medium and grown to early exponential phase (OD_{600} 0.2-0.3). Then, cultures were divided into untreated and 0.5 mM IPTG treated cultures and incubated for 24 h before fluorescence measurement.

1.5. *S. meliloti* motility assays

Swarming motility assay. Bacteria were grown in liquid TY medium to OD_{600} 1 andwashed in 1 ml of MM medium twice by centrifugation at 4500 rpm for 1 min. Pellet was resuspended in 100 µl of MM to achieve 10-fold culture concentration (10^9 cells/ml). From here, drops of 2 µl were inoculated in 20 ml MM 0.6% Noble Agar (Difco) plates, previously dried for 15 min. Bacterial drop was dried for 10 min and then incubated for 24 and 48 h at 30 ºC.

Swimming motility assay. Bacteria were grown in liquid TY medium to OD_{600} 0.2-0.3 (1-3x10^8 cells/ml). Then, 3 µl were inoculated into plates containing 25 ml semi-solid Bromfield medium (0.04% Tryptone, 0.01% yeast extract, 0.01% $CaCl_2 \cdot 2H_2O$, 0.3% agar), which were solidified with the lid on. Plate were sealed and incubated with the lid up for 72 h at 30 ºC.

2. Molecular biology techniques

2.1. DNA handling

Isolation of total DNA. For total DNA extraction, the commercial kit Realpure Genomic DNA (Real) was employed. Alliquots of 1.5 ml from *S. meliloti* culture were collected by centrifugation at 5000 rpm for 2 min and washed with 0.2 ml of a 0.1% sarcosyl solution in TE [10 mM Tris-HCl, 1mM Ethylenediaminetetraacetic acid (EDTA); pH 8]. Cell lysis was performed by adding 0.6 ml of Lysis solution and incubation for 5 min at 80 ºC in the presence of Proteinase K. The cell lysate was cooled to room temperature and then 3 µl of RNase solution was added. The samples were mixed by inversion and then incubated for 1 h at 37 ºC. 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 13000 rpm for 5 min and the supernatant was collected in a new tube. The DNA was precipitated by adding 0.6 ml of cold isopropanol, invertion of the whole mixture and incubation for 10 min at -20 ºC.

After centrifugation for 10 min, the pellet was washed with 0.2 ml of 70% ethanol and resuspended in 100 µl of sterile ultrapure water. DNA was stored at -20 ºC.

Alternatively, colony lysis was used as a quick strategy to obtain total DNA in a sufficient quality for PCR amplification. For that, bacteria were picked with a sterile toothpick, stirred up in sterile ultrapure water (volume of water corresponding to PCR mixture without reagents) and incubated at 99 ºC for 5 min. After this, tubes were cooled on ice for 5 min and centrifuged at 5000 rpm for 2 min. PCR reagents were added to the supernadant, gently mixed and subjected to a standard PCR program.

Isolation of plasmid DNA. Plasmid isolation from *E. coli* and *S. meliloti* strains was achieved using the commercial kit mi-Plasmid Miniprep Kit (Metabion). 1.5 ml of culture was collected by centrifugation at 5000 rpm for 2 min. After removal of supernatant, 0.2 ml of MX1 Buffer (containing RNase A) was added and cell were resuspended by pipetting. Then, 0.25 ml of MX2 Buffer was added and samples were gently mixed for further incubation at room temperature for 5 min. After addition of 0.35 ml of MX3 Buffer and gently mix, samples were centrifuged at 13000 rpm for 7 min. Supernatants was loaded into a Mini Column and centrifuged at 9000 rpm for 60 s. The columns were then washed with 0.5 ml WN Buffer and 0.7 ml WS Buffer. Columns were centrifuged at 13000 rpm at 3 minutes and placed into a new tube and loaded with 50 µl of Elution Buffer. Finally, columns were centrifuged at 13000 rpm for 3 min to elute plasmid DNA, which was stored at -20 ºC.

Plasmid DNA from *E. coli* were also obtained by precipitation with magnesium salts [129]. 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% Sodium dodecyl sulfate (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, the mixture was incubated on ice for 2 min and centrifuged for 1 min at 13000 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 at 13000 rpm. The supernatant was transferred to a new tube and mixed with 0.6 ml of -20 ºC pre-cooled 100% ethanol to precipitate the plasmid DNA. The mixture was incubated at room temperature for 5 min and centrifuged for 5 min. The pellet was washed with 0.2 ml of 70% ethanol, dried in a vacuum concentrator and finally resuspended in deionized water containing 10 µg/ml RNase.

Polymerase chain reaction (PCR). Phusion™ High-Fidelity DNA Polymerase (ThermoFisher) or Q5® High-Fidelity DNA Polymerase (New England Biolabs) were used for high fidelity amplification of DNA fragments. OneTaq® DNA Polymerase (New England Biolabs) was used for routine PCR. The reactions were generally performed in a final volume of 25-50 µl following manufacture's instrcutions.

To visualize the PCR products, a horizontal electrophoresis was run in 0.8-2% denaturing agarose (VWR) gels prepared with TAE (40 mM TrisHCl, 2 mM Na_2EDTA, 0.11% (v/v) glacial acetic acid). Then, DNA staining was performed by immersing the gels in a solution containing 1 mg/ml GelRed® Nucleic Acid Gel Stain (Biosalab) in deionized water for 30 min – 1 h. To visualize DNA, gels were exposed to UV light with 365 nm of wavelength to visualization of DNA. DNA molecular weight markers employed were:

Marker III: phage λ DNA digested with *Hind*III and *Eco*RI enzymes (Roche).

Marker ϕ29: phage ϕ29 DNA with the same name digested with *Hind*III (Universidad Autónoma de Madrid).

Marker pGEM: pGEM-T plasmid DNA digested with *Hinf*I and *Eco*RI enzymes (Promega).

DNA cloning. PCR products were purified using the commercial kit mi-PCR Purification kit (Metabion). To isolate DNA from agarose gels, the commercial kit mi-Gel Extraction kit (Metabion) was used. Digestion of plasmid DNA and PCR products using restriction enzymes was carried out at optimal temperature and buffer conditions according to the manufacturer's guidelines (Roche or New England Biolabs). Digested DNA insert and vectors were also purified using mi-PCR Purification kit (Metabion) before ligation reactions. Ligation between insert and vector DNA was achieved by mixing them in a ratio 10:1 or 100:1 for mid and low-number of copy vectors, respectively. Mixtures were incubated in the presence of T4 ligase (Roche) at 4 ºC for 16 h. The ligation product was transformed into *E. coli* DH5α for further testing of colonies carrying positive recombinant plasmids by routine PCR from colony lysates.

In case of pGEM®-T Easy Vector Systems, PCR products must be subjected to an adenylation reaction catalyzed by OneTaq® DNA Polymerase in the presence of 0.3 mM dATP at 70 ºC for 30 min, prior to ligation. Then, adenylated DNA fragment was mixed with pGEM-T vector and a ligase provided with the kit.

Plasmids and universal oligonucleotides. Plasmids and sequences of the oligonucleotides used for DNA cloning are listed in Table 2.

Table 2. General plasmids and oligonucleotides used in DNA cloning.

PLASMIDS	Relevant characteristics	Reference/Source
pK18*mobsacB*	Suicide plasmid in *S. meliloti*, *sacB*, *oriV*, Km^r	[128]
pSRKKm	pBBR1MCS-2 derivative with a P_{lac} promoter, *lacIq, lacZa*$^+$, Km^r	[130]
pSRKGm	pBBR1MCS-2 derivative with a P_{lac} promoter, *lacIq, lacZa*$^+$, Gm^r	[130]
pSK_FLAG	pSRKKm carrying 3xFLAG for C-terminal protein tagging and IPTG-induced expression of tagged protein; Km^r	[119]
pSRK-MS2Term	pSRK derivative harboring the 43-nt MS2 aptamer sequence	[119]

pSKiMS2	pSRKKm carrying the MS2 aptamer sequence fused to *sinR*-P_{sinI}	This work
pBB-*eGFP*	pBBR1MCS-2 derivative for generation of promoter eGFP fusions; Km^r	[97]
pABCa	*oriVSm* (repABCpMlb); *oriVEc* (*E. coli oriVp15A*); AR (P_{min2}- *aacC1*); synTer-MCS (synTer1); Gm^r	[131]
pABCa::*GFP*	pABCa derivative carrying module for generation of promoter *eGFP* fusions from pBB-*eGFP*; Gm^r	This work
pR-*eGFP*	Vector for generation of target mRNA-*egfp* translational fusions; Ap^r, Tc^r	[91]
pET16b	Bacterial vector for inducible expression of N-terminally 10xHis-tagged proteins with a Factor Xa site	Novagen
pET29a	Bacterial vector for expression of N-terminally S-tagged proteins with a thrombin site	Novagen
Oligonucleotides	**5'-sequence-3'**	**Use**
PCR1	CGGGCCTCTTCGCTATT	Clonning in pSRK derivative and pK18*mobsacB* plasmids
PCR2	TTAGCTCACTCATTAGG	
SR_Fw	CTGATCGGCATCAGCGTCAC	Clonning in pBB-*eGFP*
GFP_Rv	GTTGGCCATGGAACAGGTAG	
pABC_Fw	CTGTTGTTTGTCGGTGAACG	Clonning in pABCa
pABC_Rv	GCCAGTTACCTCGGTTCAAA	
Egfp-139_rev	GATGAACTTCAGGGTCAGCTTG	Cloning in pR-*eGFP* (with PCR2)
SP6	GTATTCTATAGTGTCACCTAAATAGC	Cloning in pGEM-T Easy
T7	TAATACGACTCACTATAGGGCGA	
sinR_NdeIF	GCCACATATGGCTAATCAACAGGCTGTC	Construction of pSKi-sRNA, pSKiMS2 and pSKiMS2-sRNA
TSS3_28bp_b_sinIR	GTAGCGATGCTGTCAGGCTC	
MS2fusTSSI	GAGCCTGACAGCATCGCTACCGTACACCA TCAGGGTAC	
HindIIIVec	CGAGGTCGACGGTATCGATAAGCTTCGCC	
SecSRK	TTCCATTCGCCATTCAGGCT	

Sequencing and sequence analysis. Sanger sequencing was performed at the DNA sequencing (genomics service of Instituto de Parasitología y Biomedicina "López-Neyra" (CSIC, Granada) 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 visualizeed using BioEdit software.

2.2. RNA handling

Total RNA isolation. Total RNA was isolated from nodules and free-living bacteria cultured in the described conditions by acid phenol/chloroform extraction according to published protocols [40, 132]. For that, bacterial cultures (25 ml) were stopped by addition of 5 ml of STOP solution (90% ethanol, 10% phenol). Cultures were then centrifuged for 10 min at 3500 x*g* and 4 ºC. Pellets were frozen in liquid N_2.

Pellets obtained from exponentially growing cultures were resuspended in 1 ml of Lysis Solution (1.4% SDS, 4 mM EDTA, 500 µg Proteinase K) pre-heated at 65 ºC (3 ml for stationary growth phase cultures or mucoid strains). Bacterial suspensions were incubated at 65 ºC for 10 min (vortexed each 2 min) and then, placed on ice (mucoid strains were incubated at 65 ºC for 20 min). 500 µl of 5 M NaCl at 4 ºC was added for each ml of cell lysates. After incubation for 10 min on ice, the cell debris were pelleted by centrifugation for 20 min at 10000 rpm and 4 ºC. The supernatant was transferred to a new tube and RNA was precipitated with 3 volumes of cold 100% ethanol. After incubation at -80 ºC for at least 1 h, cell lysates were centrifuged for 30 min at 10000 rpm and 4 ºC. Ethanol was removed prior to Deoxyribonuclease I (DNase I) treatment.

RNA pellets were resuspended in 270 µl ultrapure water, transferred to 1.5 ml tubes and mixed with 25 µl of DNase 10-fold concentrated buffer (500 mM Tris-HCl 7.5, 10 mM $MgCl_2$), 3 µl of DNase I (Roche) and 2 µl RNase inhibitor (New England Biolabs). Reaction mixtures were incubated for 1 h at 37 ºC.

After the reaction was phenolized with 1 volume (250 µl) of phenol–chloroform–isoamylalcohol [25:24:1 (v/v)] mixture, the aqueous phase was transferred to a new tube. RNA precipitation was achieved by adding 20 µl of 3 M sodium acetate pH 5.2 and 4 volume (1 ml) of cold 100% ethanol. After

incubation at -80ºC for at least 1 h, the reaction mixtures were centrifuged for 30 min at 13000 rpm and 4ºC. The pellets were then washed with 0.2 ml of 70% ethanol. After centrifugation for 10 min, the pellets were dried in a vacuum concentrator and resuspended in ultrapure water. An Invitrogen™ Qubit™ 3 Fluorometer was used to determine the RNA concentration using the Qubit™ RNA High Sensitivity (HS) kit.

In order to assess quality, the total RNA was separated in 1.3% denaturing agarose gels in MOPS buffer (4-fold concentrated stock solution: 80 mM MOPS, 20 mM sodium acetate and 4 mM EDTA, pH 7) and 1.875% (v/v) formaldehyde. MOPS and formaldehyde were added after agarose melting in ultrapure water. Loading buffer (0.15% Orange G, 7.5% glycerol 10 mM EDTA) supplemented with GelRed was added to the RNA samples. Then, electrophoresis was run in MOPS buffer and with an usual voltage of 100 V.

Northern blot. RNA samples (15-30 µg) were resuspended in Loading Buffer (0.3% bromophenol blue, 0.3% xylene cyanol, 10 mM EDTA pH 7.5, 97.5% formamide) and denatured by incubation at 95ºC for 5 min, followed by a rapid transfer to ice. Smaples were subjected to vertical electrophoresis on 6% polyacrylamide/7 M urea [acrylamide/bis-acrylamide 29:1 (VWR)/ Urea (Sigma)] denaturing gels, which were solidified by addition to 1% (w/v) APS, and 0.5 µl per ml Tetrametiletilendiamina (TEMED), in TBE (0.089 M Tris-HCl pH8, 0.089 boric acid and 0.002 M EDTA) and with an usual voltage of 450 V.

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 in TBE. One of the 3MM papers was glued to the gel avoiding the formation of bubbles and two more 3MM were placed on it. With the help of the three papers, the gel was removed from the electrophoresis glass, and then covered with the membrane and three more 3 MM papers over 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 60 min. After the transfer was completed, the RNA was fixed to the membrane by exposure to UV light (120 mJ/cm^2).

Oligonucleotides labeled at their 5'-end with γ-^{32}P dATP were used as probe. 50 pmol of oligonucleotide were labeled with the Polynucleotide kinase (New England Biolabs) in the presence of 1 µl of γ-^{32}P dATP by incubation for 1 h at 37 ºC. 25-50 µl of labeling reaction were purified with G-25 columns (GE Healthcare).

Afterwards, membranes were pre-hybridized at 42ºC for 30 min with 20 ml of Pre-hybridization solution (0.5 M sodium phosphate buffer, pH 7.2, 7% SDS, 10 mM EDTA) in an hybridisation oven. The probe was denatured by heating at 95 ºC for 5 min. 50 pmol of radiolabeled probe was added to the solution for overnight hybridization at 42 ºC. Then, the probe was removed, and the membrane was washed successively with 2XSSC/0.1% (w/v) SDS for 1 min, 2XSSC/0.1%(w/v) SDS for 15 min, 1XSSD/0.1%(w/v) SDS for 15 min and 0.1XSSC0.1%(w/v) SDS for 15 min at 42ºC (20-fold concentrated SSC solution contains 3 M NaCl and 0.3 M sodium citrate)

The membrane was then covered with plastic and exposed over to a PhosforImager screen for 16 h**.** The images were acquired with a Personal molecular imager® FX scanner equipped with the Quantity One® program.

Quantitative reverse transcription PCR (RT-qPCR). RNA samples were additionally treated with Invitrogen™ DNase TURBO™ for 1 h at 37 ºC and further cleaned up with the RNeasy Mini Kit (Qiagen) following the manufacturer guidelines (except that, to improve sRNAs retention, RNA samples are loaded onto the columns mixed with 7 volumes of 100% ethanol). cDNA was synthesized with the Takara Prime Script RT Master Mix (Perfect Real Time) using 1 µg of total RNA. RT-qPCR was carried out in a QuantStudio 3 (Thermo Fisher Scientific) with the Takara TB Green Premix ExTaqII (Tli RNaseH Plus) using 0.5 µl of cDNA following manufacture's instructions. The ratios of transcript

abundance were calculated as the ΔΔCT mean average of three replicates on three independent RNA extracts. The constitutively expressed gene *SMc01852* was used to normalize gene expression [133]. Control reactions without reverse transcriptase (–RT) in the RNA samples were simultaneously performed to confirm an absence of DNA contamination.

MS2 affinity purification coupled with RNA-seq (MAPS). The affinity purification assays were performed following a previously described protocol adapted to *S. meliloti* [69, 119, 134–136]. For pSKMS2 construction, the intergenic region *sinR-*P_{sinI}*-*TSS_{sinI} was PCR amplified using genomic DNA as the template with the primers sinR_NdeIF/TSS3_28bp_b_sinIR, and the MS2-aptamer was amplified from pSRKMS2 using the MS2FusTSSI /HindIIIvec primer pair [119]. These two fragments overlap and were jointly used as template for amplification with sinR_NdeIF/HindIIIvec, and the resulting PCR product was restricted with *Nde*I and *Xba*I and inserted into pSRKKm to yield pSKiMS2. sRNAs were cloned downstream theMS2 aptamer using the *Xba*I site.

The MS2 coat protein, carrying a double mutation to avoid oligomerization [137], is N-terminally fused to the MBP yielding the recombinant MS2-MBP. Production of the fusion protein was IPTG-induced in *E. coli* DH5α and purified by FPLC over an amylose column (GE Healthcare) in a solution containing 20 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7.9, 200 mM KCl and 1 mM EDTA. The sample was subsequently purified over a heparin column (GE Healthcare) using the same buffer with a 20-400 mM KCl gradient to remove bound nucleic acids.

Cells equivalent to 200 OD_{600} were harvested by centrifugation 10 min at 3,500×*g* and 4 ºC. Pelleted cells were then washed once with 20 ml of 0.1% sarcosyl solution in TE buffer, centrifuged for 5 min at 6,000 rpm (4 °C) and frozen in liquid N_2. Then, pellets were resuspended in 4 ml of Buffer A [20 mM Tris–HCl pH 8.0, 150 mM KCl, 1 mM $MgCl_2$, 1 mM Ditiotreitol (DTT)]

supplemented with cOmplete™ Protease Inhibitor Cocktail (Roche) and RNase inhibitor (New England Biolabs). Cell lysis was performed by three consecutive passes through a French's Press 1000 PSIG (for RNA co-purification) or by disruption using a Branson Sonifier sonicator in three cycles of 10 s bursts at 32 W with a microprobe (for protein co-purification). The lysate was then centrifuged (15 min, 16,000×*g*, 4 °C) and the supernatant was incubated with 200 pmol of MS2-MBP bait protein for 30 min at 4ºC in Belly Dancer Shacker (VWR). The SigmaPrep™ spin column (Sigma) was prepared for affinity purification by 3 washes with 800 µl buffer A, and loading of 100 µl amylose resin (New England Biolabs) previously washed in buffer A twice. The mixture of cell lysate and the bait protein was then loaded into the amylose column, which interacts non-covalently with the MBP moiety. Unspecific-binding was removed by four column washes and then 800 µl buffer A containing 15 mM maltose were added to elute MS2-tagged sRNA-RNA/protein complexes.

RNA and proteins contained in the eluted fractions were separated by phenol–chloroform–isoamylalcohol [25:24:1 (v/v)] extraction and subsequent RNA precipitation of the aqueous phase with 4 volumes of ethanol supplemented with 20 µg of glycogen and protein precipitation of the organic lower phase with 3 volumes of acetone for 16 h at -20 °C. Protein pellets were washed with acetone and RNA pellets were washed with 70% ethanol. Then, samples were resuspended in ultrapure water.

To monitor the procedure, RNA was extracted from 50 µl of cleared lysate, flow through, and wash fractions by phenol–chloroform–isoamylalcohol [25:24:1 (v/v)] extraction. RNA preparations were probed with appropriate labelled oligonucleotides upon Northern blotting.

RNA-seq. For the profiling of RNA species bound to MS2-tagged sRNAs, RNA fractions eluted from columns were used. For transcriptome profiling, total RNA isolated from bacterial cultures was used. Strand-specific cDNA libraries

from RNA were generated and sequenced in the Illumina NextSeq Mid 150 platform. This process was performed in Instituto de Parasitología y Biomedicina "López-Neyra" (CSIC, Granada).

Demultiplexed sequencing reads were mapped with Bowtie2 v2.2.3 using standard parameters to the *S. meliloti* Sm1021 reference sequence downloaded from the Rhizogate portal [105, 138]. Uniquely mapped reads were assigned to protein coding genes or non-coding RNAs with Rsubread 3.12 [139]. Read counts for each genomic feature were normalized by coverage and the resulting RPKM values were the basis for fold change calculations [140]. The Integrative Genomics Viewer (IGV) software was used for data visualization [141].

***In vitro* synthesis and labelling of RNA.** The desired transcripts were amplified for *in vitro* transcription using DNA as template with the corresponding primer pairs. The forward primers incorporate the T7 promoter sequence. PCR products were further purified by phenol–chloroform–isoamylalcohol [25:24:1 (v/v)] extraction and ethanol precipitation. Then, 200 ng of PCR templates were transcribed with home-made T7 RNA polymerase by incubation at 37 ºC for 2 h in the presence of Reaction Buffer (0.15 M $MgCl_2$, 0.4 M Tris-HCl, 0.02 M spermidine, 0.05 M DTT), Invitrogen™ RNase OUT™ and 30 mM each rNTPs. Synthetized transcripts were DNase I-treated (Roche), phenolized, dephosphorylated with Antartic Phosphatase (New England Biolabs) for 1h at 37 ºC) and finally, phenolized again prior to purification on 6% polyacrylamide-7 M urea gel, which was stained with Sybr Gold. Purified trancripts were quantified with a Qubit3 Fluorometer (Thermo Fisher Scientific). For protein-binding assays, transcripts were radiolabelled at the 5'-end with γ-^{32}P-ATP using T4 Polinucleotide Kinase (New England Biolabs).

2.3. Proteins handling

Protein purification. The pET-16b vector (Novagen) was used to yield the desired proteins tagged at its N- termini with 10 histidine residues. For purification of the recombinant protein, the corresponding plasmids were mobilized to BL21 (DE3) strain (Novagen). Transconjugants were grown at 37 ºC in 1 l of LB medium supplemented with 100 μg/ml ampicillin and 50 μg/ml chloramphenicol to an OD_{600} of 0.4-0.6. Protein production was induced at 20ºC by addition of IPTG to a final concentration of 0.3 mM. After 18-20 h of incubation, cells were harvested by centrifugation at 5000 x*g* for 15 min at 4 ºC. Cells pellets were resuspended in 25 ml buffer B (50 mM Tris-HCl pH 8, 0.5 M NaCl) supplemented with cOmplete™ Protease Inhibitor Cocktail (Roche) prior to lysis by treatment with 20 μg/ml of lysozyme and three consecutive passes through a French's Press 1000 PSIG. The lysate was centrifuged at 18000x*g* for 1 h at 4ºC. The supernatant was loaded onto a HisTrapTM HP 5 ml column (GE Healthcare) equilibrated with buffer B, which was then washed with three column volumes of the same buffer. Then, column was washed with three column volumes of 5% buffer C (50 mM Tris-HCl buffer pH 8, 0.5 M NaCl and 1 M imidazole). Protein was eluted in a 5%-100% buffer C-gradient during 15 min. Fractions containing purified protein were dialyzed overnight against 50 mM Tris-HCl pH 8, 0.5 M NaCl, glycerol 10% at 4 ºC.

The pET29a vector (Novagen) was used to yield protein variant tagged at its N-termini with 15 residues of the S-peptide that include a thrombin site. In this case, buffer D (20 mM Tris-HCl pH 8.5, 300 mM NaCl, 0.1 mM EDTA) was used to resuspend the cells pellet. The supernatant was loaded onto a Hitrap Heparin HP column (5 ml, GE Healthcare) equilibrated with buffer D and eluted with a gradient of 0.4–1.5 M NaCl. Fractions containing the protein were dialyzed against buffer E (20 mM Tris-HCl pH 8.5, 500 mM NaCl, 0.1 mM EDTA, 10% glycerol).

Protein concentration was measured using Bradford assay (Bio-Rad). The purified protein was stored at -80 ºC.

Denaturing protein electrophoresis. To separate extracted proteins according to their molecular weight, proteins were loaded in vertical electrophoresis in denaturing acrylamide/bis-acrylamide gels containing SDS (SDS-PAGE). This gel was prepared in a Mini-ProteanII (Bio-Rad) electrophoresis cell and was composed of two different parts: separating gel [10-15% acrylamide/bis-acrylamide 29:1 (VWR), 0.375 M Tris-HCl pH 8.8, 1% (w/v) SDS, 1% (w/v) APS, and 0.5 µl per ml TEMED] and stacking upper gel [5% acrylamide/bis-acrylamide 29:1 (VWR), 0.125 M Tris-HCl pH 6.8, 1% (w/v) SDS, 1% (w/v) APS, and 0.5 µl per ml TEMED]. Protein samples were mixed with Protein Loading Buffer [0.01% (w/v) bromophenol blue, 12% (w/v) SDS, 300 mM Tris.HCl pH 6.8, 60% (v/v) glycerol] and 2% β-mercaptoethanol and then denatured by boiling for 3 min. After this, samples were cooled on ice for 5 min and loaded into the gel. The electrophoresis was run at 150 V in Running Buffer [0.124 M Tris-HCl, 1.252 M glycine, 5% (w/v) SDS]. For estimation of protein molecular weight of the proteins resolved on the gels, the Precision Plus Protein Dual Color Standards (Bio-Rad) was used.

Tagging and production of proteins with αFLAG. The pSK_FLAG vector was used to tag proteins at their C-termini with three consecutive units of the FLAG epitope (3xFLAG) [119]. The tagged protein is encoded downstream of an IPTG-dependent promoter. This allows western blotting using Mouse anti-FLAG® antibody (Sigma).

Protein immunoblot and gel staining. For Western blot, aliquots equivalent to 0.2 OD of cell lysates were resuspended in Protein Loading Buffer. Cell protein fractions were denatured by heating at 95°C for 5 min, resolved by SDS-PAGE and blotted onto a polyvinylidene difluoride membrane (PVDF; 0.45 µm pore, Amersham) similarly to described for Northern blotting. Prior to immobilization

of proteins on PVDF, the membrane was activated with 100% methanol for 90 s. The gel and twelve 3MM Whatman papers were equilibrated in Transfer Buffer (25 mM Tris-HCl pH 8.3, 192 mM glycine, 20% methanol). Protein transfer was carried out at 0.8 mA/cm^2 membrane for 50 min.

For immunological detection, membranes were subjected to blocking of non-specific binding sites by incubation with 1.5% (w/v) Blocking Reagent (GE Healthcare) in TBST Buffer (25 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% Tween20) for 1 h in a Belly Dancer Shacker. Then, membranes were probed with a monoclonal Mouse anti-FLAG® antibody (Sigma; 1:5,000) for 1 h. After six 5 min washing steps with TBST, the membrane was incubated with peroxidase-conjugated anti-mouse IgG antibody (Sigma; 1:20000) for 1 h. The secondary antibody was removed by six 5 min washing steps with TBST. Blots were developed by incubation for 5 min in Blotting Detection Reagent (GE Healthcare), which was prepared by adding solution A (luminol) and B (peroxide), and signals were detected with a ChemiDoc system (BioRad).

Alternatively, polyacrylamide gels were stained as appropriate either with 0.25% (w/v) Coomassie R-250 Brilliant blue (Bio-Rad) or the Bio-Rad silver staining kit according to the manufacturer's instructions. For Coomassie staining, gels were previously immersed in a (14:45:9) ethanol:acetic acid:water solution.

Liquid chromatography mass spectrometry (LC-MS/MS) and protein identification. MS analysis was performed at the Proteomics Service from Instituto de Parasitología y Biomedicina "López-Neyra" (CSIC, Granada). Protein samples equivalent to 120 OD_{600} were run 10 min in a 4% SDS-PAGE. The gel lane was cut into 10 slices and subjected to in-gel tryptic manual digestion. The resulting peptides were fractionated using an Easy n-LC II chromatography system (Proxeon) in line with an Amazon Speed ETD mass spectrometer (Bruker Daltonics).

2.4. Evaluation of nucleic acid and protein interactions

DNA-chromatography pull down assay. This protocol was adapted from Jutras *et al*. 2012. The DNA probes were generated by PCR amplification of either genomic or plasmid DNA using specific biotynilated forward plasmids. DNA fragments were concentrated and purified by phenol–chloroform–isoamylalcohol [25:24:1 (v/v)] extraction. DNA is contained in the upper aqueous phase, and was precipitated by addition to 20 µl of 3M sodium acetate (pH 5.2) and 1 ml of pre-cooled 100% ethanol. After incubation at -80ºC for at least 1 h, the reaction mixtures were centrifuged for 30 min at 13000 rpm and 4ºC. The pellets were then washed with 0.2 ml of 70% ethanol. After centrifugation for 10 min, the pellet was dried in a vacuum concentrator and resuspended in ultrapure water. A NanoDrop® ND-1000 spectrophotometer was used to determine the DNA concentration.

Cells equivalent to 400 OD_{600} were harvested by centrifugation 10 min at 3,500×*g* and 4 ºC. Pelleted cells were then washed once with 0.1% sarcosyl in TE buffer and frozen in liquid N_2. Cells were resuspended in 4 ml Protein Binding Buffer (20 mM Tris-HCl buffer pH 8, 1 mM EDTA, 10 mM HEPES, 10% glycerol, 100 mM NaCl and 0.05% Triton X-100) supplemented with cOmplete™ Protease Inhibitor Cocktail (Roche) and lysed by three consecutive passes through a French's Press 1000 PSIG. Cells were centrifuged at 12000x*g* and 4ºC for 10 min and supernatant was stored for the next step.

In parallel, 200 µl of Streptavidin Resin (GenScript) were washed four times in Wash Buffer (20 mM Tris-HCl pH 8, 1 mM EDTA, 150 mM NaCl) by centrifugation at 8200 x*g* for 15 s. Then, resin was resuspended in 0.6 ml of Wash Buffer containing 40 µg of biotinylated DNA in a 2 ml-tube and this mixture was incubated at 4 ºC for 1 h in a VWR® Rotisserie. After this, the mixture was centrifuged at 8200 for 1 min and the pellet was washed in 500 µl of Protein Binding Buffer three times. Bacterial lysates and bovine serum albumin (BSA) (5

μg/ml) were added to the streptavidin resin and DNA baits in a 15-ml falcon tubes, and incubated at 4 ºC for 16 h in a BellyDancer Shacker.

Afterwards, the mixture was centrifuged at 8200x*g* and 4ºC for 1 min and the supernadant was removed leaving 0.8 ml of buffer to resuspend the pellet. This mixture was loaded into a SigmaPrep™ spin column (Sigma) previously washed twice with Protein Binding Buffer. Then, resin was washed three times with Protein Binding Buffer supplemented with cOmplete™ Protease Inhibitor Cocktail (Roche) and 5 μg/ml BSA. Finally, the resin was incubated with 120 μl of Elution Buffer (50 mM Tris-HCl pH 8 and 2 M NaCl) for 10 min for the release of the proteins that bound to DNA upon centrifuged at 8200 x*g* for 2 min. The use of the Elution Buffer containing increasing concentrations of NaCl releases proteins with different DNA-affinities (protein eluted by low concentration of NaCl have low DNA affinity). To monitor the presence of DNA baits, DNA was released from streptavidin resin by addition to 60 μl of 8M guanidine·HCl (pH 1.5), centrifugation and stabilization with 60 μl of 1 M Tris-HCl pH 8.

Electrophoretic Mobility Shift Assays (EMSA). The promoter regions of the sRNAs of interest were amplified by PCR using Q5® High-Fidelity DNA Polymerase (New England Biolabs) and genomic DNA as the templates. The PCR products were isolated from the PCR reaction or an agarose gel by using the mi-PCR Purification or mi-Gel Extraction kits (Metabion), respectively and then labeled at its 5′-ends with γ-32P dATP and T4 polynucleotide kinase. The labeled probe (1-10 nM) was then incubated with different concentrations of the purified proteins in 20 μl of STAD [25 mM Tris-acetate pH 8.0, 8 mM Mg-acetate, 10 mM KCl, 3.5% (w/v) polyethylene glycol-8000 and 1 mM DTT] supplemented with 15 μg/ml of poly(dI-dC), and 200 μg/ml of BSA. The reaction mixtures were incubated for 30 min at 4°C, and samples were run on 4.5% (w/v) native polyacrylamide gels [4.5% acrylamide/bis-acrylamide 29:1 (VWR) 25 mM Tris, 200 mM glycine, 1% (w/v) APS and 0.5μl per ml TEMED] in Mini-Protean II

(Bio-Rad) for 2 h at 50 V at room temperature in Tris-glycine Buffer (25 mM Tris, 200 mM glycine). Loading Buffer composition was 40 mM Tris-HCl pH 8.0, 40% (w/v) glycerol, 100 mM NaCl, 0.2% (w/v) bromophenol blue, 0.1 mM EDTA. The results were analyzed with Personal FX equipment and Quantity One software (Bio-Rad).

DNA Footprint. The DNA probes were similarly generated by PCR. For the footprint on the top strand, the PCR was carried out with a 5'-end labeled with [γ-32P]-dATP forward primer. For the footprint on the bottom strand, the same primers were used, but in this case, reverse primer was 5' end-labeled. Purified labeled probe (20 nM) was incubated without or with protein in 50 µl reaction volume of STAD supplemented with 10 µg/ml of poly(dI-dC), and 200 µg/ml of BSA. Reaction mixtures were incubated for 30 min at 4 °C before being treated with 100 µl of 1:80000 DNase I dilution at 30ºC for 2 min. Then, 200 µl of phenol–chloroform–isoamylalcohol [25:24:1 (v/v)] mixture was added, mixed vigorously for 30 s and centrifuged at 13000 rpm for 5 min. Afterwards, DNA was precipitated from the aqueous phase by addition to 350 µl of pre-cooled 100% ethanol with 3M NaAc (pH 4.8). After a further washed in pre-cooled 70% ethanol and, finally resuspended in 9 µl of ultrapure water with 3 µl of STOP solution provided by the Thermo Sequenase Cycle sequencing kit (Appliedbiosystems). This kit was used for the generation of a sequencing ladder from the DNA probe. The results were analyzed with Personal FX equipment and Quantity One software (Bio-Rad).

Filter binding assays. Radiolabeled transcripts (1 nM) were incubated for 30 min at 30 ºC with different concentrations of the purified RNA-binding protein in a 10-µl reaction containing Binding Buffer (50 mM Tris-HCl, 0.5 mM NaCl, 1 mM DTT). Binding specificity was assessed by competition experiments in the presence of a molar excess (100 nM) of the corresponding unlabelled sRNA. BSA (2 µM) was also used instead of the tested RNA-binding protein as negative

binding control. Reactions were loaded into a dot-blot device (Bio-Dot; Bio-Rad) provided with nitrocellulose over a PVDF membranes (Amersham), which retain RNA-protein complexes and free RNA, respectively. Before and after sample blotting, wells were rinsed three times with 50 µl of DTT-less Binding Buffer. Membranes were finally exposed to a Phosphor Imager screen (Bio-Rand) and quantified with the Quantity One software (Rio-Rad).

CoIP-RNA assays with 3×FLAG tagged proteins. *S. meliloti* bacteria carrying the plasmid encoding FLAG-tagged protein (pSK_FLAG) was induced with 0.5 mM IPTG. Cells equivalent to 200 OD_{600} were harvested by centrifugation 10 min at 3,500×*g* and 4 ºC. Pelleted cells were then washed once with 20 ml of 0.1% sarcosyl solution in TE buffer, centrifuged for 5 min at 6,000 rpm (4 °C) and frozen in liquid N_2. Then, the pellet was resuspended in 4 ml of Lysis Buffer [50 mM Tris–HCl (pH 7.4), 150 mM NaCl, 1 mM EDTA and 1% TritonX-100] supplemented with cOmplete™ Protease Inhibitor Cocktail (Roche) and RNase inhibitor (New England Biolabs). Cell lysate was performed by three consecutive passes through a French's Press 1000 PSIG. The lysate fraction was then centrifuged (15 min, 16,000×*g*, 4 °C). The supernatant was collected and 100 µl of ANTI-FLAG M2 affinity gel (Sigma), previously washed in Wash buffer [50 mM Tris–HCl (pH 7.4), 150 mM NaCl] three times, were added. The mixture of cell lysate and resin was incubated overnight at 4ºC in Belly Dancer Shacker (VWR) and then centrifuged at 8200 for 1 min. After this, the supernadant was removed and the pellet was resuspended in 0.8 ml of Wash buffer and loaded into a SigmaPrep™ spin column (Sigma). This column was prepared for affinity purification by 3 washes with 0.8 ml of Wash buffer. Unspecific-binding was removed by four column washes with 0.8 ml of Wash buffer. 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 100 µl of Wash buffer prior to incubation for 30 min at 4 ºC with gentle agitation. Finally, the column was centrifuged for 30 s at 8200xg and 4 ºC.

RNA contained in the eluted fractions were separated by phenol–chloroform–isoamylalcohol [25:24:1 (v/v)] extraction and subsequent RNA precipitation of the aqueous phase with 4 volumes of ethanol supplemented with 20 µg of glycogen for 16 h at -20 °C. RNA pellets were washed with 70% ethanol. Then, samples were resuspended in ultrapure water and stored at -80 ºC.

3. Plant assays

Alfalfa (*Medicago sativa* L. 'Aragón') seeds were surface sterilized by immersion in a solution of 2.5% $HgCl_2$ for 9 min with agitation in a Rotisserie® (VWR). Under sterile conditions, seeds were washed six times and kept 2 h in sterile deionized water. Finally, seeds were washed three times with sterile deionized water and placed in Petri dishes with 1.5% agar-water. The stimulate germination, seeds were incubated with the lid up at 4 ºC in the dark for 24 h prior to incubation with the lid down at 30 ºC until the seedlings reached a length of about 10 mm (approximately after 24 h). Finally, seeds were transferred to sterile test tubes containing 10 ml of N-free nutrient Rigaud and Puppo solution (per 1 l: 0.2 g KH_2PO_4, 0.2 g $MgSO_4 \cdot 7H_2O$, 0.2 g K_2SO_4, 0.12 g $CaSO_4$, 0.05 g Ferric EDTA, 4 mg $Na_2MoO_4 \cdot 2H_2O$, 18.6 mg H_3BO_3, 22.6 mg $MnSO_4$, 29 mg $ZnSO_4 \cdot 4H_2O$, 24mg $CuSO_4 \cdot 5\ H_2O$, pH 7). Seeds were placed in filter paper stripes. The tubes with the plants were left in darkness for 1 day at room temperature and for 4 days in the growth chamber. To prevent roots from being directly affected exposed to light, the tubes were covered in their lower half with opaque paper prior to transport to the growth chamber, where plants were maintained under the following conditions: 500 µE 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.

Rhizobial strains were grown in liquid TY for 48 h at 170 rpm and 30 °C, 10 µl of pre-cultures were resuspended in TY broth to an OD_{600} 0.5, and then diluted 100-fold in sterile water to prepare an inoculum of approximately 10^6 bacterial cells/ml. Then, plants were inoculated with rhizobia strains in 1 ml inocula was added to each plant. Uninoculated plants served as a control.

Root colonization assays. Root colonization by bacteria was assessed by counting colony-forming units (CFUs). In this case, 5 seeds were placed in each test tube. At 2, 24, 48 and 72 h post-inoculation, 15 roots inoculated with each rhizobial strain were washed 3 times with 20 ml sterile water during 20 s to remove the loosely attached bacteria, and the roots were weighed in groups of five placed into 2-ml Eppendorf tubes (gently removing the leftover liquid with a pipette). Then, 1 ml of sterile TE buffer was added to each tube and the attached cells were released by two sonication pulses of 1 min each, with a pause time of 1 min between the pulses, in an Ultrasons sonicator bath at room temperature. Cells were subsequently quantified by counting CFUs (normalized to grams of root). For this, 100 µl of solutions of released bacteria were inoculated on TY plates from the following dilutions depending on the post-inoculation time: 2 h (0 and 10 times diluted); 24 h (10, 100 and 1000 times diluted); 48 and 72 h (100, 1000 and 10000 times diluted). Experiments were conducted in triplicate for each tested strain.

4. Computational methods

Promoter sequence alignments were generated with ClustalW implemented in BioEdit (http://www.mbio.ncsu.edu/BioEdit/bioedit.html) [142], and searches for conserved motifs were done with the MEME algorithm (http://meme-suite.org/index.html) [143]. The logo of the motif consensus sequence was generated at http://weblogo.berkeley.edu/logo.cgi.

Venn diagrams were generated with the Venny 2.0 tool (https://bioinfogp.cnb.csic.es/tools/venny/index2.0.2.html).

CopraRNA (v 2.1.2) and IntaRNA (v 3.2.0; http://rna.informatik.uni-freiburg.de:8080/) were used to predict sRNA-mRNA base-pairing interactions applying standard parameters [71, 144]. The full-length sRNA (IntaRNA), and its homologous sequence/s (CopraRNA), was/were used for a genome-wide scan of interactions with mRNAs involving their 5'-regions (from 50 nt upstream to 300 nt downstream of the annotated start codon), their 3'-regions (from 30 nt upstream to 50 nt downstream of the annotated stop codon) and their CDSs.

.

Chapter 1. Pervasive regulation of *S. meliloti* metabolism by AbcR1 and AbcR2

1. Background

AbcR1 and AbcR2 are Hfq-dependent 121-nt and 114-nt long homologous sRNAs tandemly encoded in the IGR flanked by the genes *SMc01226* and *lsrB*, which encode an ArsR-type and the LsrB (LysR-type symbiotic regulator B) transcriptional regulators, respectively [91]. *S. meliloti* AbcR1/2 are founding members of the so-called αr15 family of *trans*-sRNAs, which has representatives in most Rhizobiaceae and Brucellaceae species, including the mammal and plant pathogens *B. abortus* and *A. tumefaciens* [40, 63, 93]. αr15 members are similar in both nucleotide sequence (85% identity) and predicted secondary structure. *S. meliloti* AbcR1 and AbcR2 are predicted to fold into three stem loops (SL1-SL3) that expose two unpaired conserved anti-Shine Dalgarno (aSD) motifs, i.e., "CUCCUCCC" and "UUCCCCUC" in the loop of SL1 and between SL1 and SL2, respectively. Both structures also evidence two signatures reported as preferred binding sites for Hfq, namely the A/U rich single-stranded region that precedes the Rho-independent transcriptional terminator, which is well conserved in all αr15 relatives, and the terminal U residues, predicted to remain unpaired in both sRNAs [91, 93].

Despite their homology, these sRNAs are differentially expressed [91]. AbcR1 expression is induced during exponential growth and at early root infection, while it is downregulated in N-fixing endosymbiotic bacteroids. Conversely, the highest AbcR2 accumulation occurs at the onset of stationary phase growth and in response to abiotic stress, whereas it is barely detected in planta throughout symbiosis. Accordingly, only a lack of AbcR1 results in a growth delay in rich medium, and both AbcR1 and AbcR2 are dispensable for wild-type nodule organogenesis and N fixation in symbiosis with *M. sativa* plants [91]. Remarkably, *B. abortus* AbcR1 and AbcR2 are functionally redundant for virulence, i.e., only a double deletion prevents chronic infection of host

macrophages [90, 145] whereas in *A. tumefaciens* only AbcR1 seems to be competent for regulation [93]. Expression of *A. tumefaciens* and *B. abortus* AbcR1 is regulated by VtlR, the LysR-type transcriptional regulator (LTTR) encoded by the neighbouring gene [146, 147]. The AbcR1 genomic region occurs in high synteny across α-proteobacteria and therefore, the *S. meliloti* ortholog LsrB has been proposed as the most probable AbcR1 regulator in this bacterium. However, this hypothesis has not yet been experimentally demonstrated [148].

Transcriptomic and proteomic signatures of knock-out mutants have drafted the regulons of the *B. abortus* and *A. tumefaciens* AbcR1/2 sRNAs [90, 92, 145]. Although such profiling does not discriminate between directly and indirectly regulated mRNAs, differentially expressed transcripts were functionally enriched in nutrient uptake and virulence factors. Further experimental validation demonstrated that subsets of these target mRNA candidates are regulated through either of the two aSD motifs identifiable in the predicted secondary structure of αr15 sRNAs [92, 145]. In *S. meliloti*, genetic reporter assays have confirmed *livK, prbA* and *SMa0495*, all encoding the periplasmic component of ABC transporters foramino acid uptake, as targets of *S. meliloti* AbcR1/2 [50, 91]. These mRNAs are most probably regulated also by base-pairing of the aSD motifs to the RBS and flanking nucleotides, but these interactions have not yet been genetically dissected.

In this Chapter, we have explored the regulation of the *S. meliloti* AbcR1/2 sRNAs and uncovered their mRNA interactomes using MS2 affinity purification coupled with RNA-seq (MAPS) [134, 135]. Our data show that LsrB and the alternative σ factor RpoH1 are responsible for the differential transcription of AbcR1 and AbcR2, respectively. In turn, these sRNAs use their aSD motifs to downregulate translation of large and overlapping arrays of mRNAs encoding transport proteins and metabolic enzymes. Further, we show that AbcR1-mediated post-transcriptional fine-tuning of metabolism enhances the ability of *S.*

meliloti to colonize the root rhizoplane, a biotechnologically relevant symbiotic trait.

2. Results

2.1. Regulators of AbcR1/2 transcription

To identify putative functional motifs in the AbcR1/2 promoters (P_{abcR1} and P_{abcR2}), we compared the 100 nucleotides preceding the TSSs of these genes and their predicted homologs in several α-proteobacteria (Fig. 9A). P_{abcR1} alignment unveiled the -35/-10 core hexamers recognized by RpoD (σ^{70}) in α-proteobacteria (CTTGAC-N_{17}-CTATAT) [46]. Upstream this σ^{70} signature, we noticed the generic LTTR motif of prokaryotic promoters (T-N_{11}-A), which occurs in tandem in most sequences. Remarkably, a more defined motif perfectly matching the proposed LsrB-binding consensus GCAT-N_3-TG-N_3-T in *B. abortus* and *A. tumefaciens* was also evident between the -61 and -49 positions in P_{abcR1}. Comparison of the P_{abcR2} sequences revealed a -35/-10 box closely matching the *S. meliloti* RpoH1 consensus (CTTGAA-N_{16}-CCTATAT) but failed to detect additional conserved motifs [149].

To experimentally test these predictions, we first transcriptionally fused full-length (P_{abcR1}/P_{abcR2}) and truncated ($P_{abcR1\text{-}38}$/$P_{abcR2\text{-}38}$) versions of both promoters to *eGFP*. All four reporter constructs were independently introduced into *S. meliloti* wild-type (Sm2011 and Sm1021), and *lsrB* (SmΔ*lsrB*), *rpoH1*, and *rpoH2* knock-out mutant strains [149](Fig. 9B). Strains Sm2011 and Sm1021 derive from the same *S. meliloti* nodule isolate (SU47). Both are considered nearly isogenic (Table 1), but strikingly, we failed to generate a *lsrB* knock-out in Sm1021 using both single and double cross-over strategies, suggesting that the genomic background can affect *lsrB* mutant viability. The activity of the promoters was assessed in bacteria grown to exponential (P_{abcR1}) or stationary

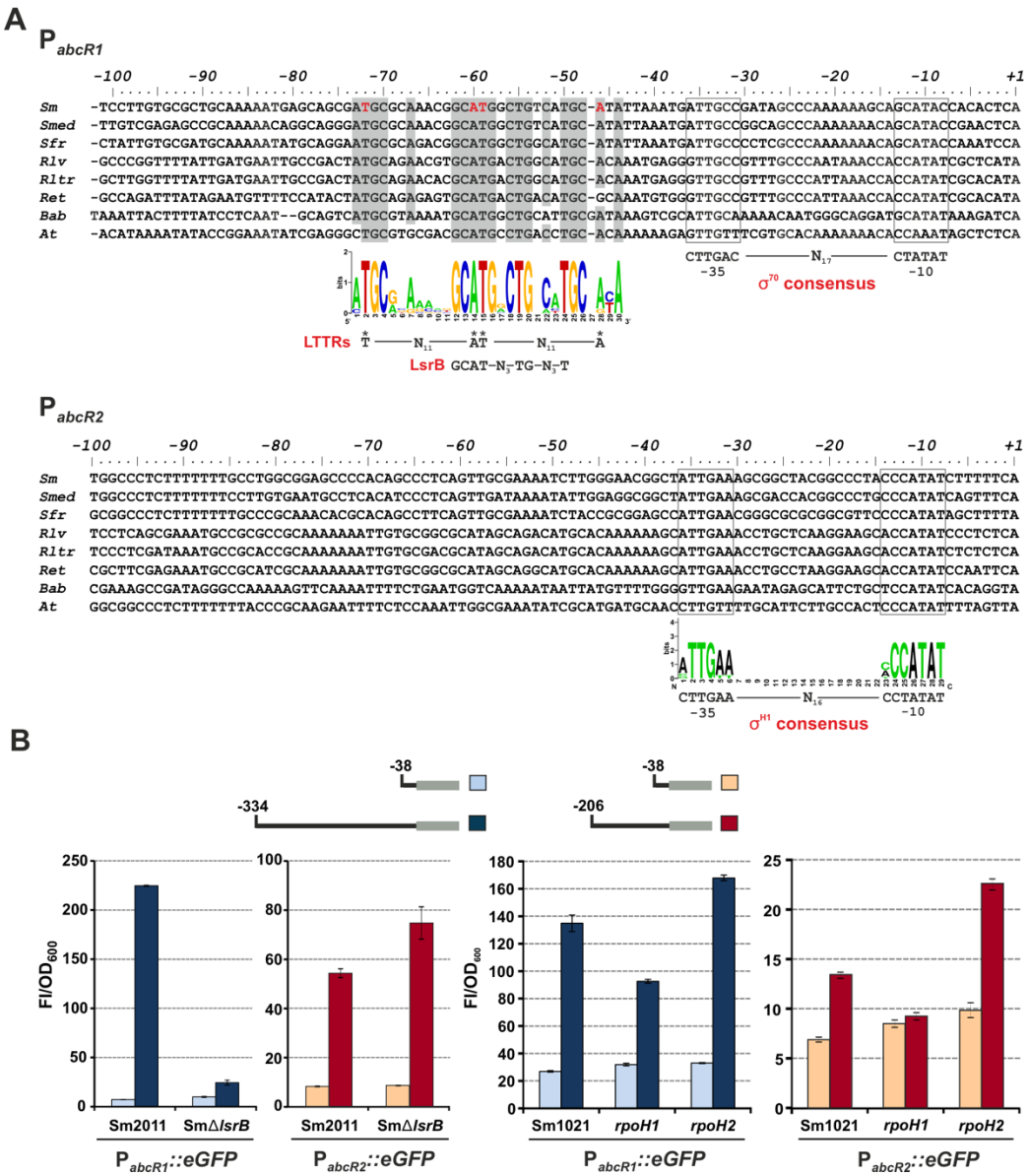


Figure 9. Regulation of AbcR1/2 expression. (A) Multiple sequence alignment of the promoters of AbcR1/2 homologs in α-proteobacteria. Consensus of conserved motifs are indicated below the alignments. *Sm*; *Sinorhizobium meliloti* Sm1021; *Smed*, *S. medicae* WSM419; *Sfr*, *S. fredii* HH103; *Rlv*, *Rizobium leguminosarum* bv. *viciae* 3841; *Rltr*, *R. leguminosarum* bv. *trifolii* WSM2304; *Ret*, *R. etli* CIAT652; *Bab*, *Brucella abortus* 2308; *At*, *A. tumefaciens* C58. **(B)** Fluorescence of promoter-*eGFP* fusions. Fluorescence derived from full-length and trimmed versions of P_{abcR1} and P_{abcR2}, as diagrammed above the bar graphs, were determined in wild-type (Sm2011 or Sm1021), and mutant *lsrB* (SmΔ*lsrB*), *rpoH1*, and *rpoH2* backgrounds. Reported values are means and Standard Deviation (SD) of nine fluorescence measurements normalized to the OD_{600} of exponential (P_{abcR1}) and stationary (P_{abcR2}) phase cultures, i.e., three replicates of three independent cultures of each reporter strain.

(P_{abcR2}) phase in complete TY medium, which induces AbcR1 and AbcR2 expression, respectively. Maximum fluorescence of the P_{abcR1}::*eGFP* reporter fusion was detected in Sm2011, decreasing by >22-fold upon removal of the predicted LsrB-binding motif. In the absence of LsrB, the activity of P_{abcR1} was merely double that of $P_{abcR1\text{-}38}$. Conversely, lack of LsrB did not significantly alter P_{abcR2} activity. Gel shift assays further demonstrated binding of LsrB to P_{abcR1} (Fig. 10A). In strain Sm1021 and its *rpoH* insertion mutants (34), P_{abcR1}-derived fluorescence was 3- to 6-fold higher than that of $P_{abcR1\text{-}38}$. In contrast, transcription from P_{abcR2} specifically decreased in the *rpoH1* mutant to the basal levels rendered by $P_{abcR2\text{-}38}$ (Fig. 9B). The strongly reduced activity of $P_{abcR2\text{-}38}$, relative to P_{abcR2}, in wild-type and mutant backgrounds is likely due to a deletion of transcriptional enhancers located upstream many RpoH boxes, which remain uncharacterized in *S. meliloti* [149].

In an independent series of experiments, we used the same set of *S. meliloti* strains to examine AbcR1 and AbcR2 accumulation in different media and growth conditions by Northern blot probing of total RNA with an oligonucleotide that cross-reacts with both transcripts (Fig. 10B and C). Hybridization of RNA from Sm2011 confirmed high levels of AbcR1 during exponential growth and increased abundance of AbcR2 upon an osmotic upshift in both complete TY and defined MM (Fig. 10C). In contrast, whereas AbcR2 retained its stress-induced expression in Sm*ΔlsrB*, AbcR1 was undetectable in this mutant. In *S. meliloti*, RpoH1 and RpoH2 are co-expressed at the onset of stationary phase growth in MM and in response to a salt shock [149]. Thus, we used these conditions to further assess RpoH-dependent expression of AbcR1/2 by Northern hybridization (Fig. 10C). Probing of RNA from strain Sm1021 confirmed AbcR2 expression during stationary phase and under salt stress. Unlike AbcR1, which accumulated in the absence of either σ^H factor, AbcR2 was not present in unstressed bacteria lacking RpoH1 and downregulated, but reliably detected, in both *rpoH* mutants

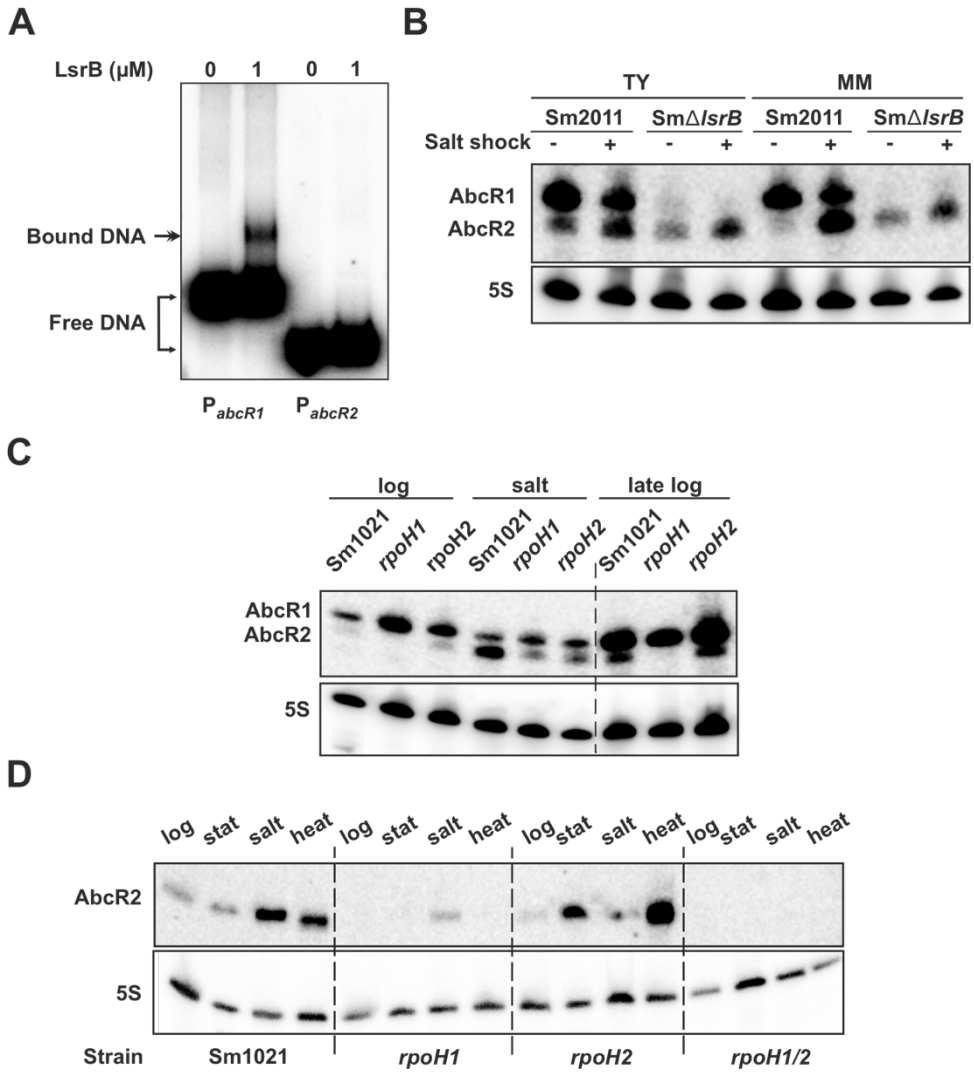


Figure 10. Transcriptional regulation of *S. meliloti* AbcR1 and AbcR2 sRNAs. (A) LsrB binds the AbcR1 promoter. Gel shift assays with radiolabelled P_{abcR1} (334 bp) and P_{abcR2} (206 bp) incubated with purified LsrB. **(B)** Northern blot analysis of LsrB-dependent AbcR1/2 expression. Total RNA was obtained from Sm2011 and Sm*Δlsr*B bacteria. **(C)** Northern blot analysis of RpoH1/2-dependent AbcR1/2 expression. Total RNA was extracted from Sm1021, as well as *rpoH1* and *rpoH2* knock-out mutants. **(D)** RpoH1/2 contribute to AbcR2 transcription upon an osmotic upshift. Northern blot probing with PbAbcR2 oligonucleotide of RNA from Sm1021 and its *rpoH1, rpoH2*, and *rpoH1/2* knock-out mutants cultured in conditions indicated along the top of the panel. The 5S rRNA was probed as an RNA loading control. Growth conditions are indicated on top of each panel. Membranes were probed with the PbAbcR1/2 radiolabelled oligonucleotide. The 5S rRNA was also probed as an RNA loading control. Hybridizations shown are representatives of at least three biological replicates per strain and growth condition.

upon a salt shock. Therefore, it is likely that RpoH1 and RpoH2 contribute additively to AbcR2 transcription under this specific condition. A previous series of Northern hybridizations probing RNA samples with an oligonucleotide specifically targeting AbcR2 anticipated this finding. In this experiment we also included the *rpoH1/2* double insertion mutant, as well as RNA from bacteria subject to a heat shock (40ºC), which also promotes the activity of both σ^H factors (Fig. 10D). Lack of RpoH1 was enough to render AbcR2 undetectable in all conditions except salt stress, in which complete inhibition of AbcR2 expression required the double knock-out. Together, these data revealed that the transcriptional regulation of AbcR1 and AbcR2 mostly depends on LsrB and RpoH1, respectively.

2.2. MAPS-based characterization of the AbcR1/2 targetomes.

We used MAPS to identify the AbcR1/2 mRNA partners at a genome-wide scale. For this, we fused the MS2 RNA aptamer to the 5'-ends of AbcR1/2, which allows for the specific capture of the tagged transcripts along with their interacting mRNAs by a MS2-MBP fusion protein immobilized on an amylose resin. Tagging at the 5'-end was previously shown to preserve the stable expression and functional secondary structure of the AbcR sRNAs [69, 119].

Given that the AbcR1/2 and NfeR1 sRNAs are predicted to share targets [97], wild-type AbcR1/2 (controls) and tagged AbcR1/2 were expressed from an IPTG-inducible promoter in the Δ*abcR1/2*Δ*nfeR1* strain Sm2020 (a Sm2011 derivative). The sRNAs were reliably detected as transcripts of the expected size in total RNA extractions following 15, 30, and 60 min IPTG addition (Fig. 11A). The chimeric transcripts also retained the ability to downregulate *prbA* as shown using a *prbA*::*eGFP* translational fusion (Fig. 11B). These data validated the tagged sRNAs as baits for affinity chromatography.

To maximize co-purification of target mRNAs and targetome coverage, we induced transcription of wild-type controls and MS2-sRNAs for a short time (15 min) in conditions that promote their endogenous expression, i.e., exponential growth in TY and MM for AbcR1, and stationary phase growth (TY and MM) as well as heat and salt shocks (both in TY) for AbcR2. Lysates from pools of AbcR1 and AbcR2 cultures were then subjected to affinity chromatography.

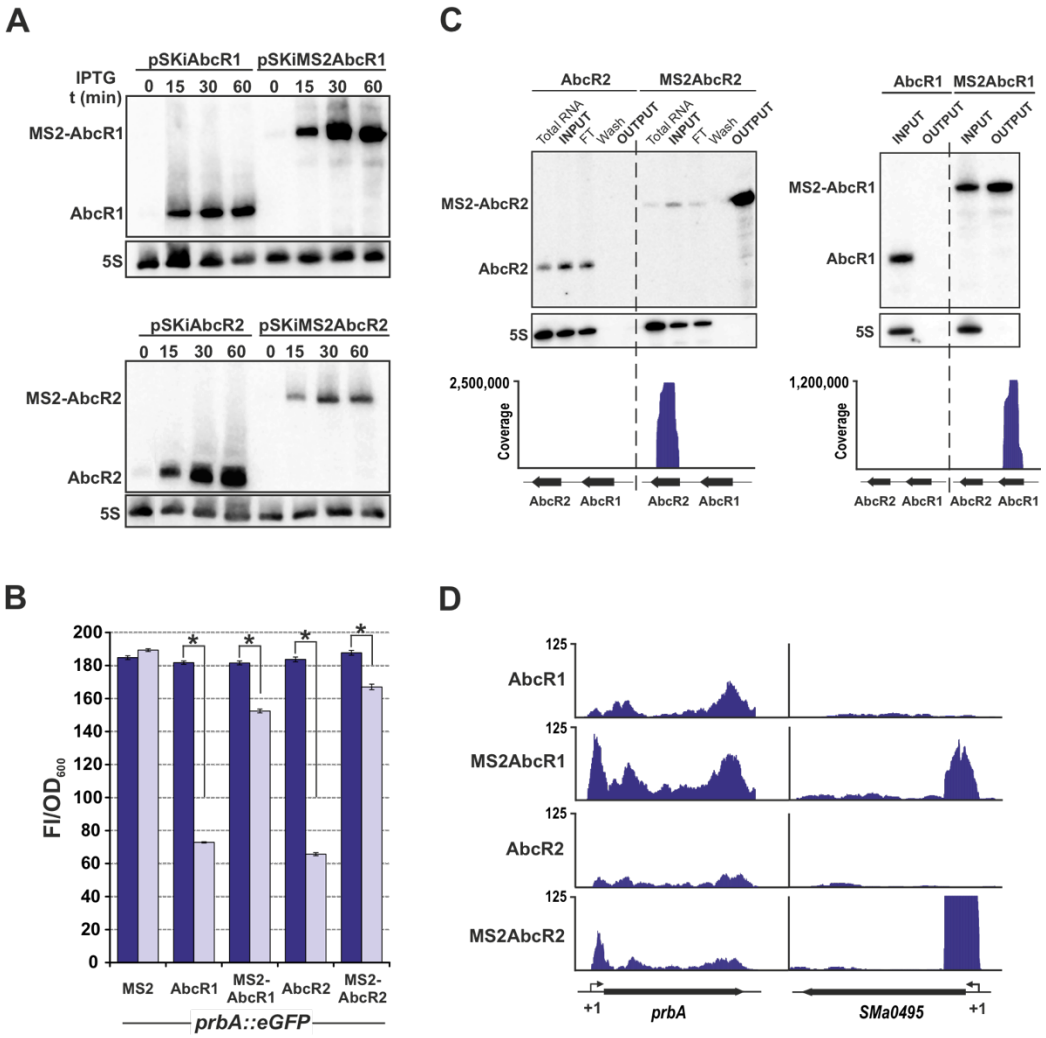


Figure 11. MPAS setup. Figure legend on the next page.

Figure 11. MAPS setup. (A) Northern blot detection of MS2-tagged AbcR1/2. RNA was obtained from *S. meliloti* Sm2020 transformed with pSKiAbcR1/2 or pSKiMS2AbcR1/2 0, 15, 30, and 60 min after addition of IPTG to cultures in MM, and probed with PbAbcR1/2. 5S rRNA was probed as an RNA loading control. **(B)** Tagged AbcR1/2 retained the ability to regulate *prbA*. Bar graphs represent fluorescence of reporter strains co-expressing the *prbA*::*eGFP* translational fusion, and AbcR1/2 or their tagged variants (plasmids pSKiAbcR1/2 or pSKiMS2AbcR1/2) upon IPTG induction (24 h). Plotted values correspond to means and SD of 18 fluorescence measurements normalized to the OD_{600} of the cultures (Fl/ OD_{600}). The asterisks over the bars indicate statistically significant differences at $p<0.05$. **(C)** Monitoring of affinity chromatography. Expression of wild-type and MS2AbcR1/2 was induced for 15 min with IPTG in *S. meliloti* Sm2020 transformed with pSKiAbcR1/2 or pSKiMS2AbcR1/2. RNA from input, flowthrough (FT), wash, and output chromatography fractions (as indicated on top of panels) was probed with PbAbcR1/2. 5S rRNA was probed as control. IGV plots of AbcR1/2 recovered in the elution (output) fractions are shown below. **(D)** Known AbcR1/2 target mRNAs co-purified efficiently with the tagged transcripts. IGV plots showing reads coverage and recovery profiles of *prbA* and *SMa0495* mRNAs upon affinity chromatography with wild-type and tagged AbcR1/2 as baits. The TSS of each mRNA is indicated (+1).

Hybridization of RNA from the input and output chromatography fractions showed that the baits were specifically retained by the MS2-MBP protein (Fig. 11C). Mapping of the sequencing reads from the eluted RNA to the *S. meliloti* reference genome (Sm1021) unequivocally demonstrated efficient recovery of the tagged sRNAs and co-purification of known targets (e.g., *prbA* and *SMa0495*), as expected (Fig. 11C and D).

Upon normalization by coverage, we compared read counts derived from controls and MS2AbcR1/2 mapping to four mRNA regions: *i*) the full-length mRNA, including the CDS and a virtual 5'-UTR of 50-nt, *ii*) a stretch of the 5'-region extending from nucleotide positions -50 to +100 relative to the translation start codon, *iii*) the CDS alone, and *iv*) the 3'-region encompassing 50 nt upstream and 30 nt downstream of the stop codon as a virtual 3'-UTR. We imposed a minimum of 50 (MS2AbcR1) or 25 (MS2AbcR2) mapped sequencing reads for an mRNA region to be considered in the comparisons. An mRNA was scored as a putative AbcR1/2 target if counts from tagged sRNA libraries exceeded a 3-fold difference (log2FC>1.5) with respect to the controls in at least one of the computed regions. IntaRNA-predicted antisense interactions using the tagged sRNAs as queries were then used to filter out those mRNAs likely captured by

unspecific binding to the MS2 aptamer. All three previously identified AbcR1/2 target mRNAs (*livK*, *prbA*, and *SMa0495*) passed the selection thresholds. All in all, MAPS identified 225 and 356 interacting mRNAs for AbcR1 and AbcR2, respectively, representing roughly 3%-6% of *S. meliloti* protein coding genes (URL).

2.3. AbcR1/2 broadly regulate core and accessory *S. meliloti* metabolism

According to the Sm1021 genomic sequence annotation [105], 70%-80% of the AbcR1/2 target mRNA candidates encode proteins with predicted function (Fig. 12A). Of those, 72% (AbcR1) and 55% (AbcR2) are most probably involved in the transport or metabolism of widely diverse substrates. An additional 8%-9% of both targetomes encode transcription factors, many of which are linked to metabolic operons. Collectively, the sets of metabolism-associated mRNAs represent 93%-99% of the AbcR1 and AbcR2 targets with functional homology in database entries. The relative distribution of both targetomes along the three *S. meliloti* replicons indicates that the impact of AbcR1/2-mediated post-transcriptional regulation is slightly biased towards the pSymB megaplasmid (Fig. 12A), which is enriched in metabolic genes. Furthermore, pangenome analysis of 23 complete *S. meliloti* genomes identified 153 (68% of the total) and 215 (60.4%) AbcR1 and AbcR2 targets, respectively, as belonging to the core genome (URL). Although the different experimental setups (i.e., growth conditions) preclude a rigorous comparison between the AbcR1 and AbcR2 targetomes, MAPS uncovered 96 common targets for both sRNAs (Fig. 12B). Interestingly, 221 (20%) of the 1,127 mRNAs previously identified as Hfq ligands were also scored as AbcR1/2 targets [50], which indicates a prominent role of AbcR1/2 in regulation of the extensive *S. meliloti* Hfq post-transcriptional network. The partial overlap between the AbcR1/2 targetomes and the Hfq partners suggests that most of these mRNAs are true sRNA targets rather than

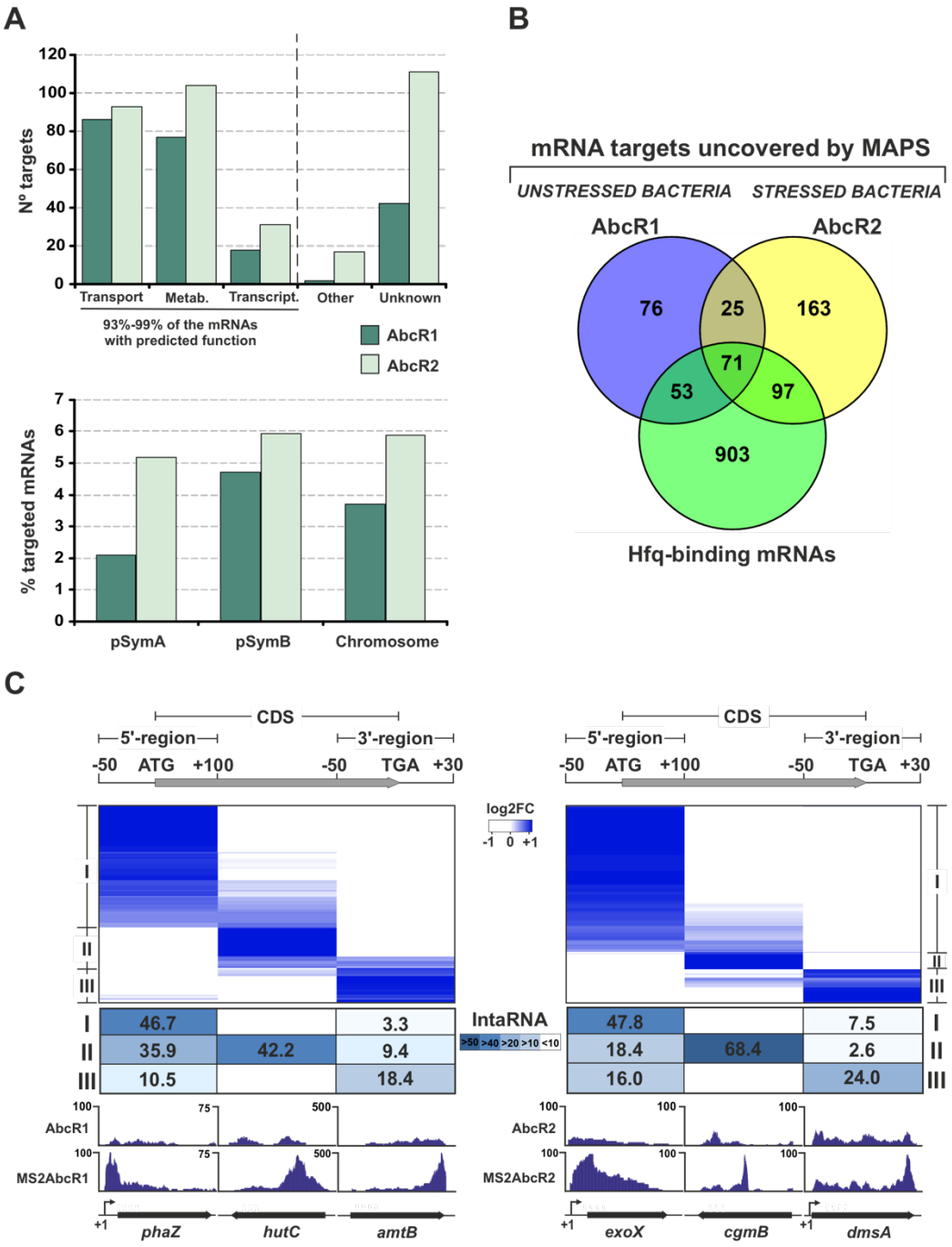


Figure 12. Overview of the AbcR1/2 mRNA interactomes determined by MAPS. Figure legend on the next page.

Figure 12. Overview of the AbcR1/2 mRNA interactomes determined by MAPS. (A) Functional categories of the captured mRNAs (top) and their distribution relative to the total number of protein-coding genes in each of the three *S. meliloti* replicons: the pSymA and pSymB megaplasmids, and the chromosome (bottom). **(B)** Venn diagram comparing the AbcR1/2 targetomes and the set of known Hfq-binding mRNAs. **(C)** Enrichment-based clustering of target mRNAs upon affinity chromatography with MS2-tagged AbcR1/2. Shown on the left is data for AbcR1, while data for AbcR2 is shown on the right. Heatmaps identify three groups of mRNAs enriched at the 5'-region (cluster I), CDS (cluster II) and 3'-region (cluster III) with respect to the control experiments with the wild-type sRNAs; these three regions are shown schematically at the top of the figure. The *S. meliloti* genome was interrogated with IntaRNA for thermodynamically favoured (E<-8) antisense interactions (minimum 7-nt seed) of AbcR1/2 in each of the mRNA regions diagrammed at the top of the figure. Numbers in the table indicate the % of mRNAs in each cluster that have a predicted antisense interaction within each region, and is presented as a heatmap (legend between tables). Interactions predicted in the CDS may overlap with those at the 5' or 3' mRNA regions and, therefore, numbers in columns and rows may add to more than 100%. IGV plots at the bottom show read coverage of target mRNAs representatives of each cluster.

transcripts recovered unspecifically by sole binding to Hfq. This analysis thus anticipates a pervasive AbcR1/2 regulation of *S. meliloti* adaptive metabolism.

2.4. AbcR1/2 use two distinct aSD motifs for regulation by modifiable base-pairing

Enrichment-based clustering of the mRNAs co-purified with tagged AbcR1/2 unveiled three groups of targets, which were characterized by sequencing coverage biases toward either the 5'-region (cluster I), the CDS (cluster II) or the 3'-region (cluster III) (Fig. 12C). Cluster I was the dominant group in both targetomes. IntaRNA predictions revealed a correlation between the enrichment of a specific mRNA region and the location of the expected antisense AbcR1/2 interaction sites (Fig. 12C). The target mRNAs *livK*, *prbA*, and *SMa0495* are representatives of the dominant cluster I. They were previously validated by means of translational fusions of their 5'-regions to *eGFP* as reporters of AbcR1/2-dependent regulation [50, 91]. Here, we used a similar genetic reporter assay to validate a new set of three target mRNA candidates within cluster I that encode transport proteins: *SMc02417*, *SMc03121*, and *SMa0392* (Fig. 13). IPTG-induced (over)expression of AbcR1/2 reduced fluorescence from the three

reporters significantly, indicating downregulation of translation and equivalent regulatory abilities of both sRNAs.

The overlapping targeting potential of AbcR1/2 likely relies on the identical pairs of 8-nt aSD motifs (aSD1 and aSD2) (Fig. 14A). To pinpoint the contributions of aSD1 and aSD2 to regulation, we first assessed the activity of two AbcR1/2 variants on the *SMc03121* and *SMa0495* target mRNAs using the translational reporter fusion assay. Specifically, we introduced nucleotide substitutions within either aSD1 (AbcR1a/AbcR2a) or aSD2 (AbcR1b/AbcR2b) that preserve the putative secondary structure of both transcripts while disrupting the predicted base-pairing at the translation initiation region of the target mRNAs (Fig. 14A). Interaction with *SMc03121* probably occurs via aSD1 whereas *SMa0495* is likely targeted by aSD2. Induced (over)expression of wild-type AbcR1/2 resulted in a decrease of reporter-derived fluorescence of both genes. Consistent with the predicted interactions, AbcR1/2a variants (aSD1 mutants) retained wild-type activity on *SMa0495* but lost the ability to repress *SMc03121*, whereas the regulatory effects were the opposite with mutants in aSD2. Regulation by the AbcR1/2 variants was rescued by making the corresponding nucleotide substitutions in the *SMc03121* and *SMa0495* interaction sites. Nucleotide changes in *SMc03121* and *SMa0495* leaders compensating mutations in AbcR1/2 that abrogated target regulation did not inhibit translation of the *eGFP* reporter (Fig. 14C; blue bars). However, *SMa0495-1* (complementary to AbcR1b) decreased basal fluorescence of the fusion to ~40% of the wild-type, most likely by interfering with the RBS. Remarkably, all these nucleotide substitutions fully abrogated activity of wild-type AbcR1/2, while restoring regulation by the complementary variants (Fig. 14C). A similar genetic dissection revealed *SMa0392*, *SMc02417*, and *prbA* regulation via aSD2 (Fig. 15). *SMa0392* leaders with nucleotide changes that restore pairing with AbcR1b/2b reduced basal level expression of the reporters by 75-50% and were not regulated by the wild-type

sRNAs. These mutations supported regulation by AbcR1b/2b, as expected. Consistent with disruption of the RBS, point mutations at the predicted AbcR1/2 binding sites in *SMc02417* inhibited translation (i.e., the basal activity of the reporters was scarcely 4% of the wild-type), thus precluding further unambiguous confirmation of aSD2 interaction at these sites. Finally, fluorescence patterns of wild-type and mutant *prbA* reporters suggest regulation by AbcR1/2 via aSD2 pairing at two contiguous sites in the mRNA leader. Wild-type sRNAs were fully active on *prbA* mutants disrupting pairing with aSD2 in only one of these sites. However, these mutations were sufficient to restore regulation by AbcR1b/2b. Collectively, these findings indicate that aSD1 and aSD2 are independent base-pairing targeting motifs that could be designed to regulate non-cognate AbcR1/2 target mRNAs.

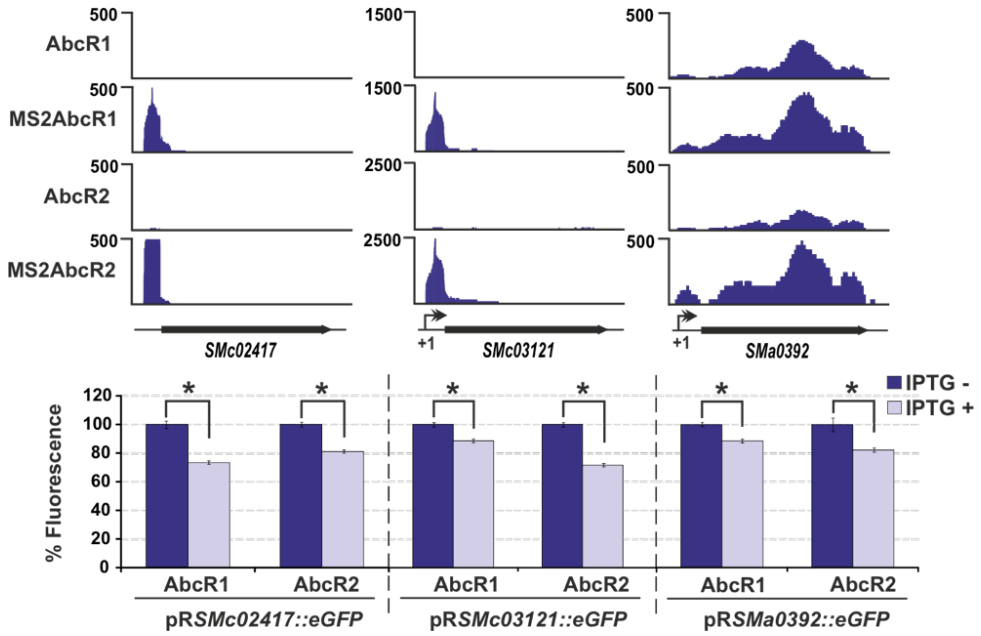


Figure 13. Experimental validation of newly identified target mRNAs. Co-purification of target mRNAs with AbcR1/2 and fluorescence reporter assays. IGV plots of *SMc02417*, *SMc03121*, and *SMa0392* coverage upon affinity chromatography with wild-type and MS2AbcR1/2. Annotated transcription start sites of mRNAs are indicated (+1). Bar graphs below represent fluorescence of strains co-expressing the target reporters indicated at the bottom, and AbcR1/2 in non-induced (100% fluorescence) and IPTG-induced (24 h) cells. Asterisks indicate statistically significant differences at $p<0.05$.

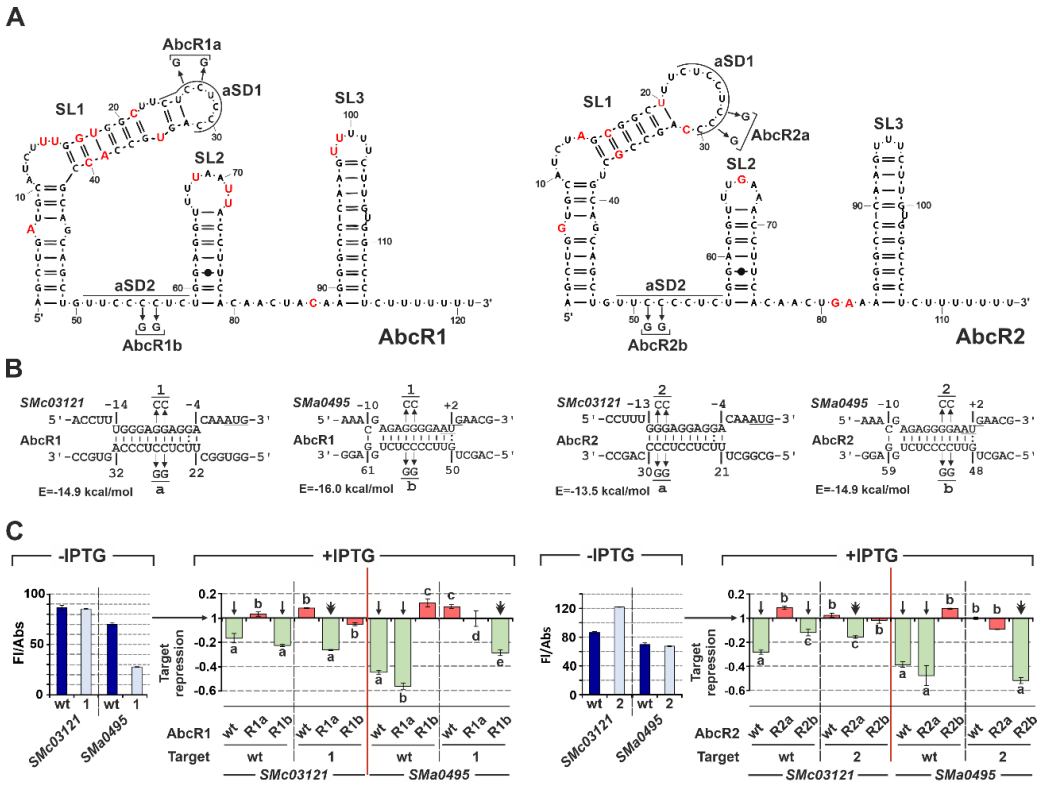


Figure 14. Genetic dissection of AbcR1/2-mRNA base-pairing interactions. (A) Predicted AbcR1 (left) and AbcR2 (right) secondary structures are shown at the top. Numbers indicate nucleotide positions relative to the 5'-end of each transcript. Stem-loops (SL) and the aSD targeting motifs are indicated. Nucleotides differing between the sRNAs are indicated in red, while the substitutions in aSD1/2 are indicated by arrows**. (B)** Diagrams shown below the secondary structures depict the predicted base-pairing interactions between AbcR1/2 and the *SMc03121* or *SMa0495* mRNAs, with indication of the hybridization energy (*E*). Numbers indicate nucleotide positions relative to the AUG start codon of the target mRNA or the 5'-end of the sRNA. Nucleotide substitutions in AbcR1/2 (a/b variants) and target mRNAs (1/2 variants) are indicated. **(C)** Fluorescence of each reporter (wild-type and variant 1 or 2) in non-induced bacteria (-IPTG) normalized by the OD_{600} of the cultures (Fl/ OD_{600}) is presented in the smaller bar graphs. Bar graphs to the right report the rates this basal fluorescence increased or decreased (target repression) upon IPTG induction of sRNA expression (24 h) in strains co-expressing the target reporters with wild-type AbcR1/2 or their mutant variants, as indicated at the bottom. Plotted values correspond to means and SD of 18 fluorescence measurements, i.e., three replicates of six double transconjugants for each reporter strain. Letters above/below bars indicate statistical groups among values from assays with each target mRNA (groups of compared values are demarcated by the red lines; ANOVA test, $p<0.05$). Arrows and the double arrowhead over the bars indicate wild-type and restored non-wild-type regulation, respectively. Red bars represent no regulation.

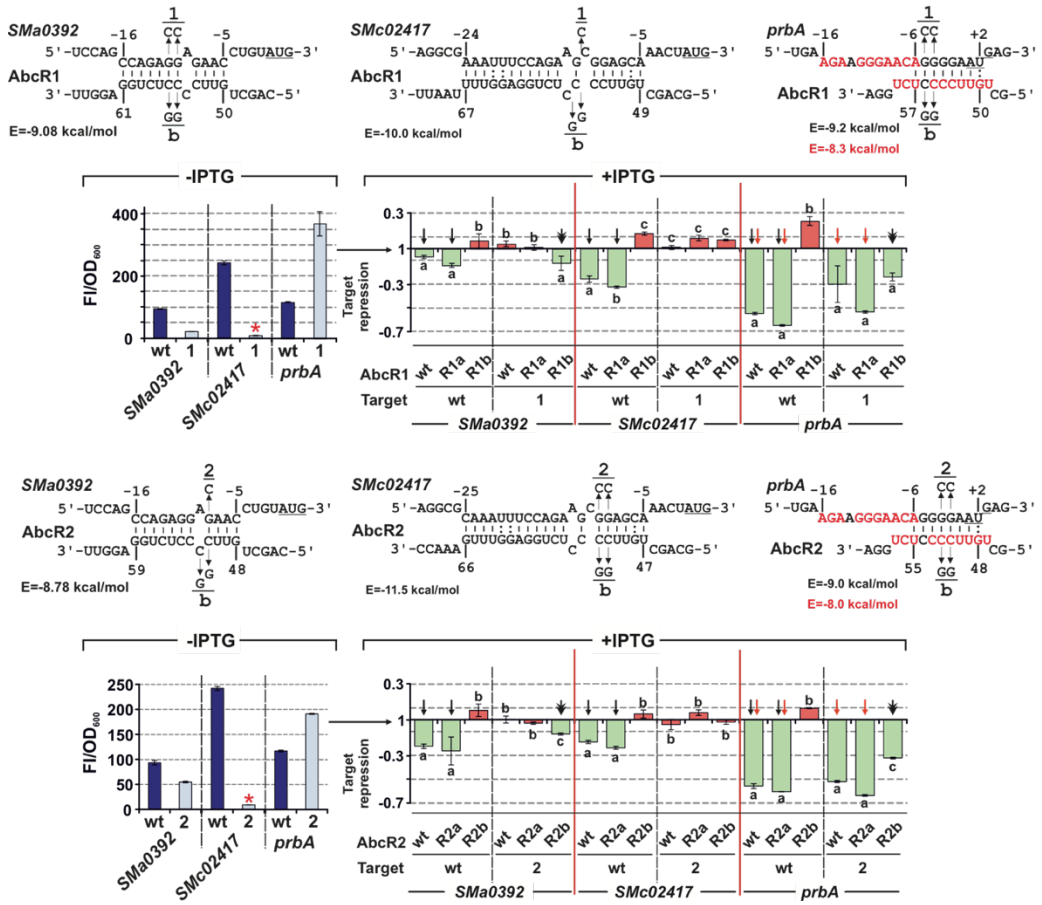


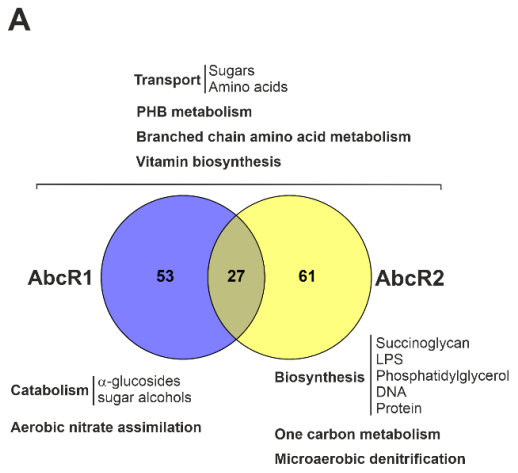
Figure 15. *SMa0392*, *SMc02417*, and *prbA* are regulated by aSD2. Diagrams depict the predicted base-pairing interactions between AbcR1/2 and each mRNA with an indication of the hybridization energy (*E*). Numbers indicate nucleotide positions relative to the AUG start codon of the target mRNA or the 5'-end of the sRNA. Nucleotide substitutions in AbcR1/2 (a/b variants) and target mRNAs (1/2 variants) are indicated. Nucleotides in red denote an alternative interaction between AbcR1/2 and *prbA*. Basal fluorescence of each reporter in non-induced bacteria (-IPTG) normalized by the OD_{600} of the cultures (Fl/ OD_{600}) is presented in the smaller bar graphs. Bar graphs to the right report the rates this basal fluorescence increased or decreased (target repression) upon IPTG induction of sRNA expression (24 h) in strains co-expressing the target reporters with wild-type AbcR1/2 or their mutant variants, as indicated at the bottom. Plotted values correspond to means and SD of 18 fluorescence measurements, i.e., three replicates in six double transconjugants for each reporter strain. Letters above/below the bars indicate statistical groups among values from assays with each target mRNA (groups of compared values are demarcated by the red lines; ANOVA test, $p<0.05$). Arrows and the double arrowhead over the bars stand for wild-type or restored non-wild-type regulation, respectively. Red bars indicate no regulation. Fluorescence patterns of the *prbA* reporters are compatible with target regulation by interaction of aSD2 at the two predicted sites (red and black arrows). Mutations in *SMc02417* abrogated *eGFP* translation, rendering reporters with basal fluorescence close to undetectable background levels (red asterisks). Consequently, assays with these reporters preclude reliable conclusions about the effect of these mutations in AbcR1/2 regulation.

2.5. Metabolic model-aided analysis of the AbcR1/2 targetomes

To further delineate AbcR1/2 function, we linked their regulons to the *S. meliloti* genome-scale metabolic model iGD1348, which combines core and accessory transport/metabolic reactions specified by 1,348 protein coding genes [150]. Eighty (35%) and 88 (25%) AbcR1 and AbcR2 targets, respectively, are represented in this model, with 27 belonging to both targetomes (Fig. 16A; URL). Traits regulated by both sRNAs are the uptake of diverse sugars and amino acids, PHB and branched chain amino acid (BCAAs) metabolism, and vitamin biosynthesis. AbcR1 seems to specifically regulate catabolism of α-glucosides and sugar alcohols, and aerobic assimilation of nitrate in rich media. One-carbon metabolism, microaerobic denitrification, and biosynthesis of succinoglycan (EPS), lipopolyssacharide (LPS), or phosphatidylglycerol are major pathways influenced by AbcR2 under abiotic stress. We next used Flux Balance Analysis (FBA) to predict the impact of AbcR1/2 target deletion on bacterial growth, and parsimonious FBA (pFBA) to predict the requirement of a particular gene for optimal flux patterns (i.e., the total metabolic flux rate). The consequences of gene deletion were examined in simulated defined media differing in the C substrate while keeping ammonia as the N source. A change of at least 10% in growth rate or total flux was considered significant (URL). Overall, these simulations predict that the combined AbcR1/2 regulon influences *S. meliloti* transport/metabolism during growth with 64 of the 83 (77%) tested C substrates.

Since AbcR1/2 promote post-transcriptional silencing, we expected downregulation of the sRNAs if one or more of their target mRNAs were predicted essential for optimal growth with a defined C substrate, e.g., mannitol, sucrose, glycerol, rhamnose, and glutamate (Fig. 16B). Probing of RNA from bacteria cultured in these media confirmed the predicted downregulation of AbcR1/2 with rhamnose and glutamate, but not with mannitol or sucrose; predictions support the observed AbcR1/2 expression with glycerol (Fig. 16C).

The apparent discrepancies between the *in silico* and experimental data are likely due to either the model assuming a complete loss rather than fine-tuning of target gene expression and/or model incompleteness, i.e., transcription factors are not included and genes specifying putative redundant transport/metabolic reactions could be missing. Although the model was not used to interrogate N metabolism, we found downregulation of both sRNAs under nitrate surplus. N stress imposed with either ammonia or nitrate prevented AbcR1 expression while promoting AbcR2 accumulation (Fig. 16C). Fluorescence of promoter-reporter fusions revealed an overall correlation between strength of transcription and AbcR1/2 steady-state levels in each growth condition (Fig. 16C). Exceptions were growth in glutamate/ammonia and N stress, thus hinting at post-transcriptional AbcR1/2 regulation in these conditions.



B

	Carbon source	AbcR1 Growth rate	AbcR1 Total flux	AbcR2 Growth rate	AbcR2 Total flux
	D-Mannitol	0.00	0.00	0.00	0.00
	Sucrose	0.00	0.00	0.00	0.00
*	Glycerol	1.00	1.04	0.89	1.13
*	L-rhamnose	0.00	0.00	0.98	0.94
*	L-Glutamate	0.00	0.00	0.00	0.00

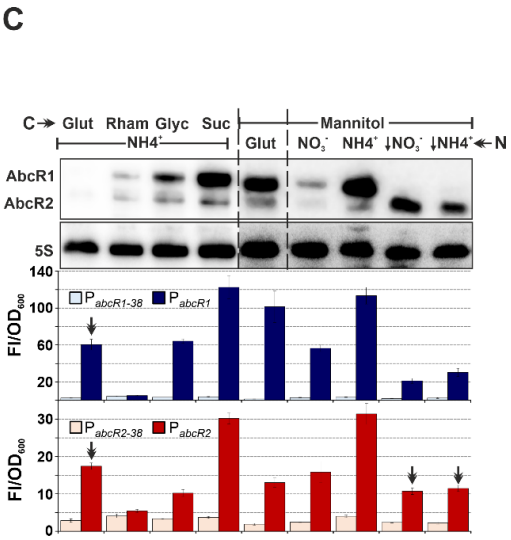


Figure 16. Metabolic model-assisted analysis of the AbcR/2 targetomes. Figure legend on the next page.

Figure 16. Metabolic model-assisted analysis of the AbcR1/2 targetomes. (A) Major transport/metabolic reactions likely regulated by AbcR1/2. The Venn diagram compares the number of AbcR1/2 target mRNAs represented in the model**. (B)** Predicted impact of target gene deletions in growth/metabolic flux in simulated defined media with the indicated C sources and ammonia as N source. Cells shaded red stand for the maximum impact and those in green indicate little to no impact. Asterisks indicate overlap between predictions and experimental results. **(C)** Changes in AbcR1/2 expression driven by shifts in C and N sources. Northern blot probing of total RNA from the *S. meliloti* Sm2B3001 strain (Sm2011 in which the *expR* gene was restored) upon growth to the onset of stationary phase in defined media with the C and N substrates indicated along the top. AbcR1/2 levels in mannitol/glutamate MM are considered the reference. Arrows indicate N stress imposed with 0.5 mM of either nitrate or ammonia. The 5S rRNA was probed as an RNA loading control. Shown is the hybridization corresponding to one of two biological replicates with identical results. Bar graphs below represent fluorescence values from promoter-*eGFP* fusions in each growth condition determined as described in Fig. 1A. Double arrowheads indicate conditions that presumably promote AbcR1/2 post-transcriptional regulation.

Modelling analysis predicts that AbcR1/2 expression may limit *S. meliloti* growth in glutamate/ammonia medium by silencing the L-amino acid ABC transporter AapJQMP (Fig. 17). Growth kinetics in this medium confirmed that the growth rate of strain Sm2020 was reduced upon IPTG-induced expression of AbcR1/2 or their AbcR1/2a variants, but not with AbcR1/2b, suggesting *aapJQMP* downregulation via the aSD2 motif (Fig. 17A). Scanning of *aapJQMP* with IntaRNA for base-pairing to AbcR1/2 unveiled a thermodynamically favoured interaction (E<-8 kcal/mol) with the aSD2 seed 12-nt upstream the start codon of *aapQ*, which encodes the permease of the system (Fig. 17B). This was consistent with the MAPS profiles, which suggested that this interaction might promote *aapQ* decay. Indeed, RT-qPCR on RNA extracts from a similar growth experiment confirmed AbcR1/2-dependent *aapQ* depletion through aSD2 (Fig. 17B). Altogether, these data support that AbcR1/2 selectively silence *S. meliloti* transport/metabolic mRNAs in response to shifts in both C and N substrates.

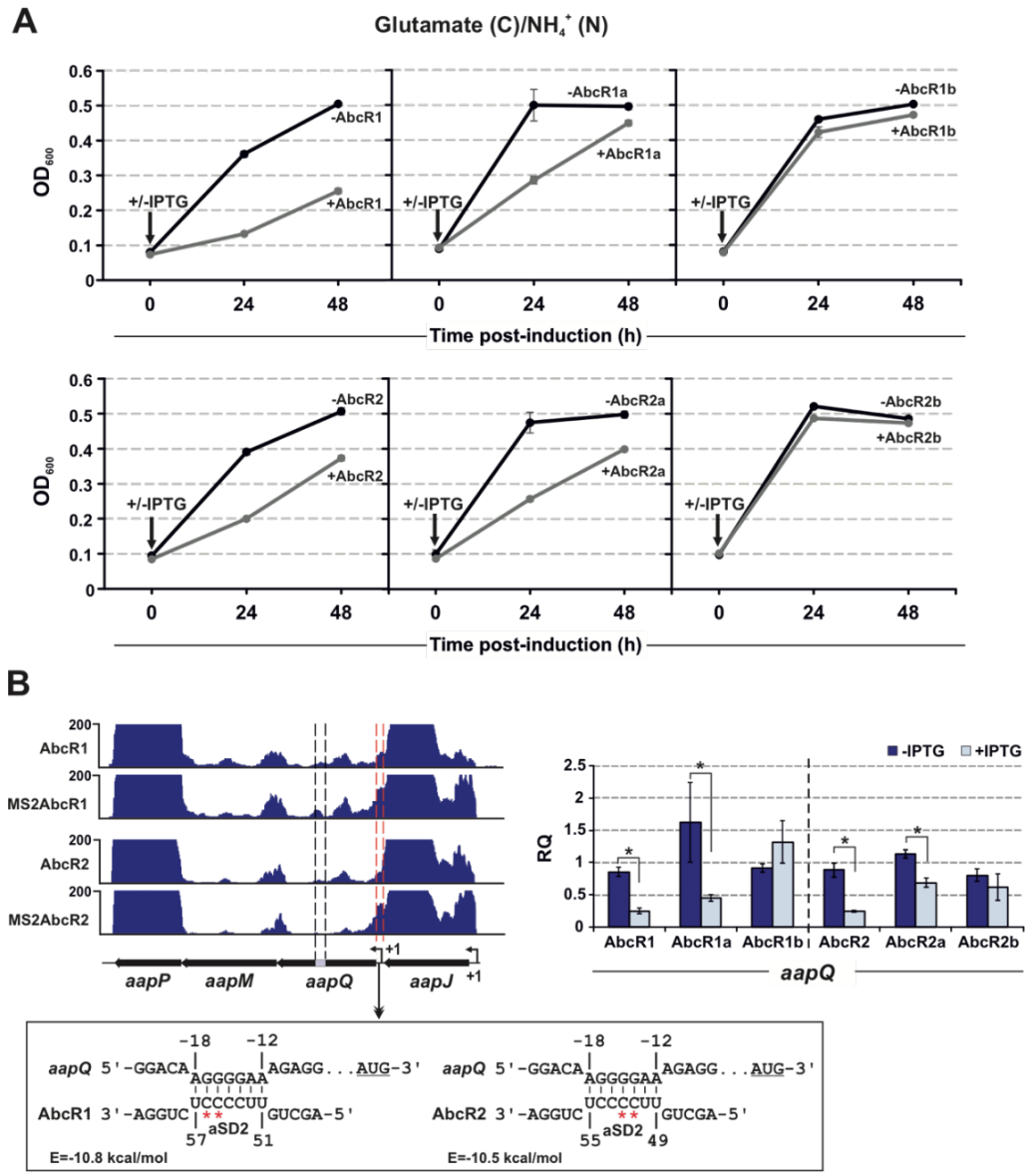


Figure 17. AbcR1/2 silence the mRNA coding for the L-amino acid permease AapQ. Figure legend on the next page.

Figure 17. AbcR1/2 silence the mRNA coding for the L-amino acid permease AapQ. (A) AbcR1/2 gain-of-function growth phenotype. Growth kinetics of the *S. meliloti* Sm2020 strain (a Sm2011 derivative) transformed with plasmids (over)expressing wild-type AbcR1/2 or their aSD1/2 mutants from an IPTG-inducible promoter. Bacteria were grown in defined glutamate/ammonia media to an OD_{600} of 0.1, at which point sRNA expression was either induced (IPTG addition) or left uninduced (no IPTG). OD_{600} was recorded in the induced and uninduced cultures 24 and 48 h following IPTG addition. Plotted values are means and SD of three independent experiments. **(B)** RT-qPCR analysis of *aapQ* regulation. Left, IGV images of the affinity purification recovery profiles of *aapJQMP*. The *aapQ* stretch enriched with respect to control experiments is demarcated by red dashed lines, while the region amplified using RT-qPCR is demarcated by black dashed lines. Diagrams below depict the predicted base-pairing interactions of AbcR1/2 at the translation initiation region of *aapQ* with an indication of the hybridization energy (*E*). Numbers indicate nucleotide positions relative to the *aapQ* AUG start codon or the AbcR1/2 5'-ends. The red asterisks denote the nucleotides mutated in the aSD2 AbcR1/2 motif (AbcR1/2b variants). Right, RT-qPCR analysis of *aapQ* abundance 30 min after inducing expression of wild-type AbcR1/2 or their mutant variants via the addition of IPTG to glutamate/ammonia cultures. Relative Quantification (RQ) values were normalized to *SMc01852* as a constitutive control. Values plotted in the bar graphs are means and SE of three replicates of three biological replicates. Asterisks above the bars indicate statistically significant differences at $p<0.05$.

2.6. AbcR1 is required for wild-type colonization of alfalfa roots

Expression profiles suggest a prevalent AbcR1/2 activity in free-living rhizobia colonizing bulk soil or the legume rhizosphere but not in endosymbiotic bacteroids [91]. Database searches identified several clusters of AbcR1/2 interacting mRNAs that are differentially expressed in rhizosphere-related conditions (i.e., exposure to alfalfa root exudates or to the nodulation genes inducer, luteolin) (Fig. 18; URL). These mRNAs specify well-recognized *S. meliloti* metabolic traits for efficient colonization of the alfalfa rhizosphere e.g., transport/metabolism of diverse amino acids and other complex N sources, and biosynthesis of the QS autoinducers AHLs (Fig. 19A). A CopraRNA-based survey of a set of phylogenetically related genomes predicts that regulation of orthologs of the AbcR1 target mRNAs belonging to the *S. meliloti* core genome is conserved across α-proteobacteria interacting with eukaryotic hosts (Fig. 19B).

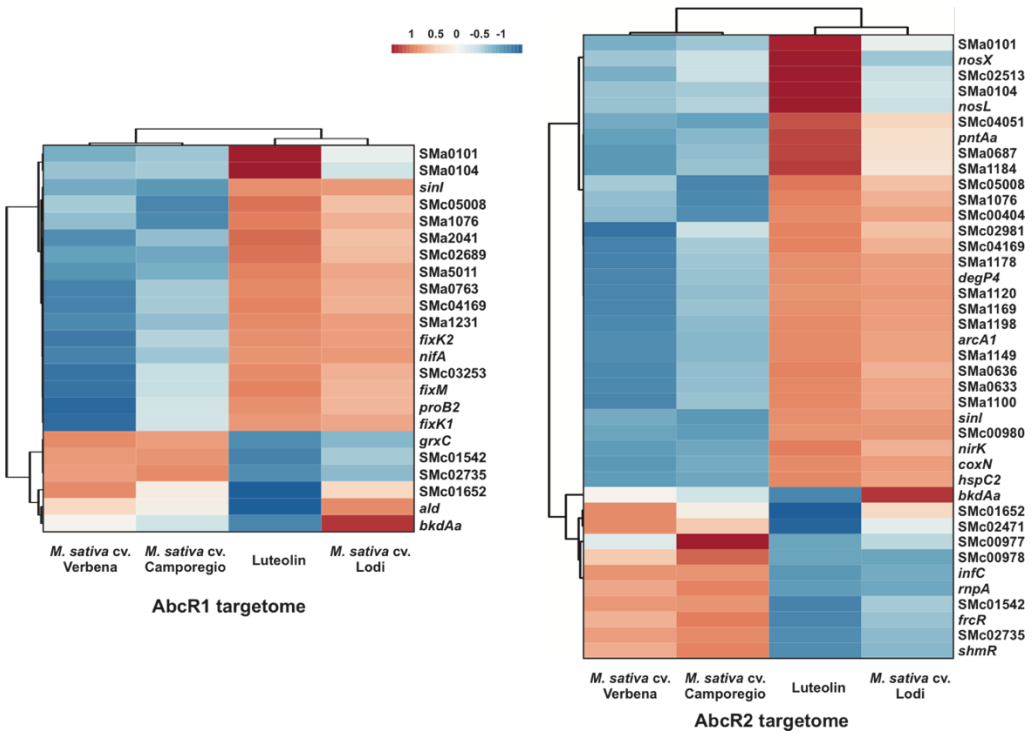


Figure 18. Occurrence of the AbcR1/2 target mRNAs in rhizosphere-related S. meliloti transcriptomic signatures. Heatmaps of expression data from *S. meliloti* Sm1021 cultures treated with luteolin and root exudates of three alfalfa varieties (Verbena, Camporegio and Lodi). Clustering is performed on Euclidean distances using the average distance method.

However, the occurrence of rhizosphere-related mRNA orthologues and their regulation by AbcR1 are limited to legume symbionts, and even more constrained to close *S. meliloti* relatives in the case of target genes belonging to the *S. meliloti* accessory genome. We obtained a similar picture when AbcR2 sequences were used as queries (not shown). This conservation pattern suggests that the AbcR1/2 regulon has evolved to help α-proteobacteria colonize the host-specific environment. These findings prompted us to investigate the impact of AbcR1/2 on the ability of *S. meliloti* to proliferate on the root rhizoplane during pre-infection stages of symbiosis with alfalfa.

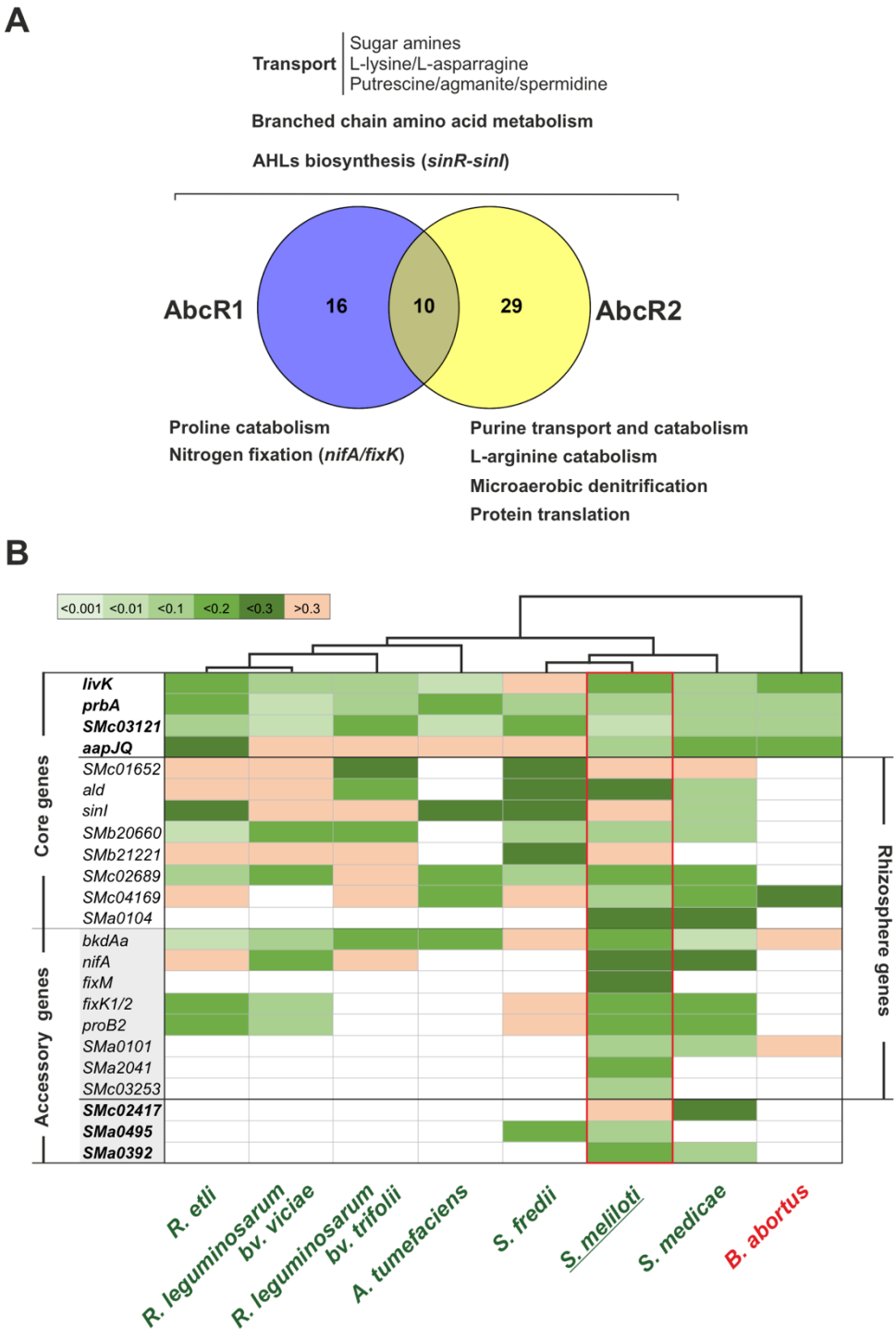


Figure 19. AbcR1/2 metabolic target mRNAs likely involved in rhizosphere colonization. Figure legend on the next page.

Figure 19. AbcR1/2 metabolic target mRNAs likely involved in rhizosphere colonization. mRNAs specify well-recognized *S. meliloti* metabolic traits for efficient colonization **(A)** Overlap between the AbcR1/2 targetomes and *S. meliloti* transcriptomic signatures in rhizosphere-like conditions. Major transport/metabolic reactions likely regulated by AbcR1/2 and relevant to rhizosphere colonization are indicated. **(B)** Conservation of AbcR1 regulation in α-proteobacteria. AbcR1 homologs in *Sinorhizobium meliloti* Sm1021, *S. medicae* WSM419, *S. fredii* HH103, *R. leguminosarum* bv. *viciae* 3841, *R. leguminosarum* bv. *trifolii* WSM2304, *R. etli* CIAT652, *A. tumefaciens* C58, and *B. abortus* 2308 were used as queries to interrogate the corresponding genome with CopraRNA for conserved base-pairing at the 5'-region of the annotated mRNAs (positions -100/+300 relatives to the annotated start codons). The heatmap represents conservation of the indicated set of core/accessory/rhizosphere *S. meliloti* mRNAs identified by MAPS (rows) and their base-pairing interactions with AbcR1. Columns are ordered according to an UPGMA tree based on homology of AbcR1 sequences. Cells are coloured based on the respective IntaRNA *p*-value as indicated, i.e, the probability of the predicted base-pairing in each individual genome. White cells indicate that the target mRNA is absent in the given bacterium according to the DomClust clustering of CopraRNA.

We inoculated sets of alfalfa plants grown hydroponically with equivalent cell densities (10^6 cells/ml rooting solution) of *S. meliloti* wild-type Sm2B3001 strain (Sm2011 in which the *expR* gene was restored), or single or double AbcR1/2 deletion mutants. Bacterial populations either attached to roots (cells/g root) or remaining in the rooting solution were then monitored by plate counting at 2, 24, and 72 h after plants inoculation (Fig. 20). Counts remained invariable and equivalent among strains in the rooting solution throughout the experiment, indicating that the rooting medium does not support bacterial growth. Conversely, bacterial density on the root surface increased exponentially, indicating active rhizoplane colonization supported by root exudates. Bacterial populations released from roots 24 and 72 h after inoculation were significantly lower in the AbcR1 and the AbcR1/2 deletion mutants compared to the wild-type strain, whereas the AbcR2 loss-of-function mutation did not influence colonization kinetics. FBA simulations with the *S. meliloti* metabolic model using a simulated rhizosphere environment similarly predicted that the AbcR1 targetome has a much greater influence on rhizosphere colonization than that of AbcR2 (URL). These data thus revealed a specific contribution of AbcR1 to alfalfa root colonization.

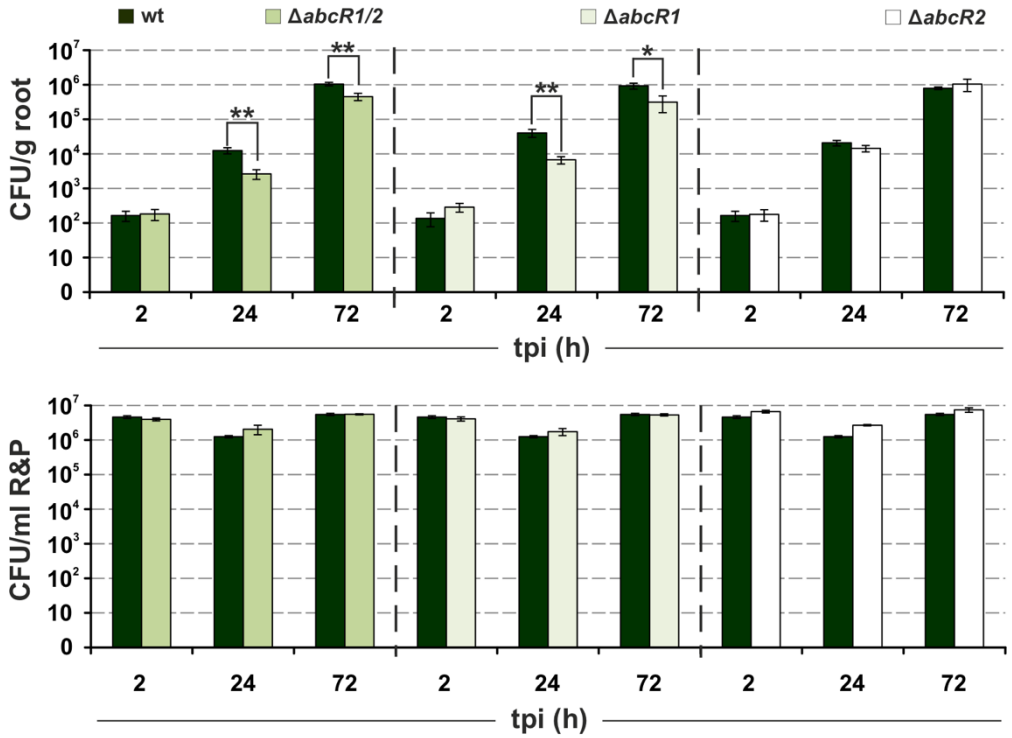


Figure 20. AbcR1 contributes to alfalfa root colonization. Root colonization assay. Alfalfa plants grown hydroponically were inoculated with either the wild-type Sm2B3001 strain, or its single or double *abcR1/2* deletion mutants as indicated along the top. Bar graphs represent the number of bacteria released from roots (CFU/g root; upper panel) or remaining in the R&P rooting solution (CFU/ml; lower panel) at different times post inoculation (tpi). Values are means and SD of counts on a total of 45 roots from plants inoculated with each strain (three sets of 15 plants per strain treated independently). * and ** above the bars indicate statistically significant differences at $p<0.05$ and $p<0.005$, respectively.

3. Discussion

Here, we showed that the *S. meliloti* homologous *trans*-sRNAs AbcR1 and AbcR2 respond to metabolic and stress signals transduced via the LTTR LsrB and the alternative σ factor RpoH1, respectively, to silence large and overlapping arrays of mRNAs related to nutrient uptake and metabolism (Fig. 21). Remarkably, AbcR1-dependent metabolic rewiring optimizes the ability of rhizobia to colonize complex nutrient-rich niches such as the root rhizoplane during early stages of symbiosis with legumes.

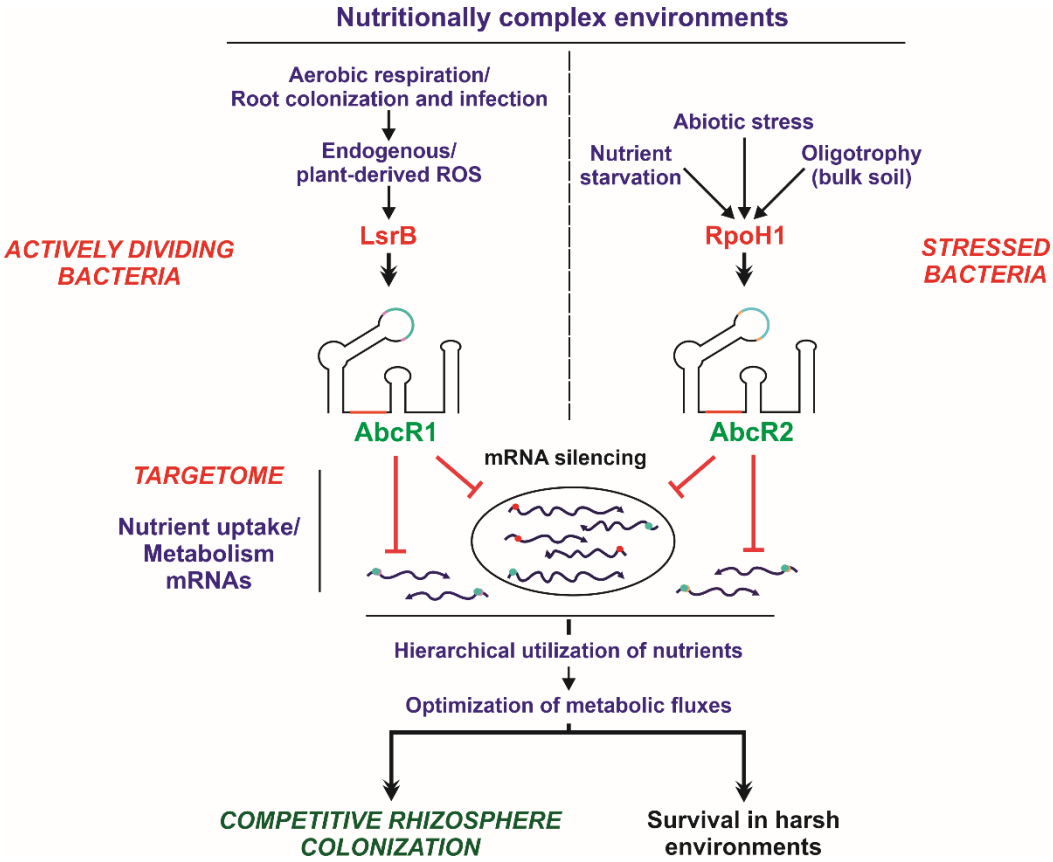


Figure 21. The *S. meliloti* AbcR1/2 post-transcriptional regulatory network. Graphical summary of data. Details in the text.

3.1. LsrB and σ^{H1} trace independent input modules for the AbcR1/2 regulatory network

Supporting *in silico* predictions, genetic and biochemical approaches unequivocally identified LsrB with the σ factor RpoH1 as the major regulators of *S. meliloti* AbcR1 and AbcR2 transcription, respectively. These findings confirm that regulation of AbcR1 is conserved in α-proteobacteria. *S. meliloti* LsrB senses the concentration of reactive oxygen species (ROS) that may derive either from aerobic respiration (endogenous ROS), redox-cycling compounds secreted by neighbouring soil organisms, or the oxidative bursts of plant defence responses during symbiosis [151–153]. Redox signal transduction by LsrB boosts

transcription of genes for the biosynthesis of LPS and ROS scavenging systems (e.g., glutathione), thereby preventing cell damage, root infection arrest, and premature nodule senescence [154, 155]. Thus, plant-derived ROS is likely the biotic signal that drives AbcR1 transcription in undifferentiated rhizobia at early symbiotic stages [40, 91]. Our data further suggest that LsrB might also transduce shifts in C and N metabolism in free-living *S. meliloti* bacteria, possibly by sensing differential accumulation of endogenous ROS. Unlike that of AbcR1, AbcR2 regulation had not previously been investigated in α-proteobacteria. A microarray-based transcriptome profiling of *S. meliloti rpoH* mutants revealed differential expression of a few annotated *trans*-sRNAs, but overlooked downregulation of AbcR2 [149]. RpoH1 recognizes gene promoters that respond to diverse stressors, such as heat shock, salinity, nutrient starvation, or the plant intracellular milieu [149], which is consistent with the stress-induced transcription of AbcR2 in free-living bacteria.

In *S. meliloti*, lack of either LsrB or RpoH1 results in severe growth and endosymbiotic phenotypes, hinting at their seemingly constitutive activity that only partially explains the differential AbcR1/2 expression (e.g., AbcR2 is not detected in nodules) [54, 149, 156, 157]. Thus, post-transcriptional and/or post-translational modifications of LsrB and RpoH1 might be further determinants of AbcR1/2 transcription rates in free-living and symbiotic rhizobia [158]. In this regard, it is known that the strength of LsrB regulation depends on oxidation of two cysteine residues that promotes protein dimerization via disulfide bonds [151]. The extent to which this post-translational modification influences LsrB affinity for promoter binding and AbcR1 transcription must be investigated.

3.2. MAPS-derived insights into the AbcR1/2 network

MAPS allowed tackling the comprehensive genome-wide profiling of the *S. meliloti* AbcR1/2 mRNA interactomes in growth conditions that stimulate

endogenous upregulation of each sRNA [159]. Remarkably, roughly 6% of *S. meliloti* mRNAs were identified in the combined AbcR1/2 targetome, most of which encoding nutrient uptake, catabolism, or biosynthesis functions. AbcR1/2 thus resemble *E. coli* GcvB sRNA in regulating an exceptionally large number of metabolic genes [160, 161]. AbcR1/2 are core components of the *S. meliloti* pangenome [63], but many of their putative mRNA targets are Hfq partners, belong to the accessory genome, and/or are encoded in the pSymB megaplasmid. This suggests a major impact of these sRNAs in the effective integration of acquired adaptive metabolism into core regulatory networks [50, 162].

The markedly uneven distribution of sequencing reads over large sets of mRNAs co-purified with tagged AbcR1/2 together with further genetic approaches suggest that AbcR1/2 act predominantly by a canonical Hfq-dependent mechanism relying on base-pairing at the RBS leading to translation inhibition [163]. Nonetheless, our data also envisage minor but plausible alternative modes of action independent of Hfq or involving interactions into the CDS of the target mRNAs. The latter has already been shown for the *A. tumefaciens* AbcR1/2 homologs [92]. Reporter assays confirmed AbcR1/2 regulation of a set of three newly identified targets (*SMc02417*, *SMc03121*, *SMa0392*) that all code for ABC transport proteins. These experiments provided further evidence that the aSD seeds (aSD1/2) are major motifs involved in mRNA targeting. aSD1 and aSD2 differ slightly in their nucleotide sequences, and they were genetically shown to fulfil independent targeting roles. A few of the nucleotides that flank aSD1 differ between AbcR1 and AbcR2, which might provide specificity for targeting. Indeed, we previously showed specific AbcR1-mediated silencing of *livK* most likely through aSD1 [91]. Conversely, aSD2 is embedded within an ultraconserved nucleotide stretch and presumably supports regulation of common target sets. Therefore, it seems likely that the functional

specificity of AbcR1 and AbcR2 is largely conferred by their differential expression rather than by their targeting potential.

3.3. A metabolic model delineated the adaptive functions of AbcR1/2

In *S. meliloti*, many predicted transport or metabolic reactions have scarce experimental support. A metabolic model-assisted analysis of the targetomes charted by MAPS suggested that AbcR1/2 impact extends beyond primary C/N energy pathways to the regulation of biosynthesis of the major C storage polymer PHB or cell envelope components (e.g., EPS, LPS, phosphatidylglycerol). In *S. meliloti*, PHB biosynthesis is negatively regulated by the MmgR sRNA under C surplus conditions [94]. MmgR is repressed by the global C flow regulator AniA [95, 164], whose coding mRNA was identified as an AbcR1/2 target. Thus, AniA may serve as a connection node of the MmgR and AbcR1/2 regulatory networks for the robust control of C homeostasis. On the other hand, mRNAs encoding proteins involved in synthesis of cell wall components were identified as AbcR2 specific targets. This leads us to speculate that AbcR2 may play a role in regulation of cell envelope remodelling in response to different stresses [165–167]. The model also linked AbcR1/2 targets to nitrate assimilation and denitrification pathways, which was further supported by profound changes in AbcR1/2 expression upon shifts in the quality and quantity of the N source. RNA regulation of N metabolism has been reported in free-living N fixers but is unprecedented in *S. meliloti* [168–170].

A major outcome of this analysis was the prediction of AbcR1/2 gain-of-function growth phenotypes linked to downregulation of the uptake of specific C substrates. As a proof of principle, we demonstrated AbcR1/2-mediated silencing of the permease component of the L-amino acid transporter AapJQMP. Together with the BraDEFGC transport system, AapJQMP rescues the symbiotic autotrophy for BCAAs of *Rhizobium leguminosarum* bv. *viciae* bacteroids within

indeterminate pea nodules, which is not a feature of alfalfa nodules [171, 172]. Regulation of *aapQ* by AbcR1/2 is predicted to be conserved in *R. leguminosarum* and may have a specific impact on pea nodule metabolism that merits further investigation.

3.4. AbcR1/2 help colonization of nutritionally complex environments by *S. meliloti*

Root exudates make the rhizosphere and rhizoplane nutrient-rich but strongly selective environments for the root microbiome [24, 172, 173]. The AbcR1 loss-of-function phenotype thus suggest that RNA regulation of metabolism provides *S. meliloti* with a competitive advantage for root colonization and saprophytic long-term survival in the rhizosphere. Both the reported specific contribution of AbcR1 to *S. meliloti* growth in complete media and this novel phenotype are consistent with AbcR1 far outweighing AbcR2 levels when both sRNAs are co-transcribed in rhizobia actively dividing under nutrient surplus or at pre-infection stages of legume symbiosis [91]. Expression and targetome profiles of AbcR2 predict a similar impact of this sRNA in colonization of bulk soil or the rhizosphere under harsh environmental conditions.

The AbcR1/2 interactomes are well-represented in the transcriptomic signatures of rhizospheric *S. meliloti* bacteria, which are enriched in transport/metabolic genes for the utilization of amino acids, sugar amines, and polyamines [174]. Remarkably, mariner-based transposon insertion sequencing has recently uncovered that knock-out of genes encoding uptake systems for quaternary amines, BCAAs, L-amino acids (e.g., *aapJQMP*), opines, and polyamines enhances *R. leguminosarum* bv. *viciae* fitness in pea rhizosphere [175]. Similarly, downregulation of the α-glucoside/trehalose/maltose transporter *aglE*, also identified as AbcR1/2 target, switches metabolism of *Ensifer* spp. to the utilization of plant-derived dicarboxylic acids in a disaccharide-rich bulk soil

to favour nodulation of pigeon pea [176]. Competitive colonization of rhizosphere and other nutritionally complex environments presumably demands optimization of metabolic fluxes in rhizobia through the hierarchical utilization of available substrates. Massive but controlled silencing of metabolic mRNAs from the LsrB and RpoH1 regulons would help prevent energy-expensive uptake, catabolism, and biosynthesis of non-priority compounds.

Overall, our findings depict an unprecedentedly large RNA network that governs metabolic adaptations of *S. meliloti* during colonization of the nutritionally complex and selective alfalfa rhizosphere. Similar networks have likely diverged to help α–proteobacteria adapt to their specific host-associated soil environments. As AbcR1/2 can be retargeted to regulate non-cognate mRNAs, this network might be rewired to engineer highly competitive biofertilizers.

4. Experimental setup

Bacterial strains and growth conditions. Bacterial strains and plasmids specifically used in this Chapter, along with their relevant characteristics, are listed in Table 3. To assess the stress-dependent expression of AbcR1/2, exponentially growing bacteria in MM were cultured for a further 1 h upon salt (400 mM NaCl) and heat (40ºC) shocks. To test the effect of shifts in N metabolism on AbcR1/2 accumulation, the L-glutamate (6.5 mM) of the standard MM was replaced by NH_4Cl (10 or 0.5 mM) or KNO_3 (10 or 0.5 mM). Similarly, the impact of different C sources in AbcR1/2 expression was assessed in MM media with NH_4Cl (10 mM) as the N source and either mannitol (54 mM), sucrose (10 mM), glycerol (15 mM), glutamate (6.5 mM), or rhamnose (15 mM) as the C substrate.

Table 3. Bacterial strains and plasmids used in Chapter 1.

Strain/Plasmid	Relevant characteristics	Reference/Source
STRAIN		

S. meliloti		
Sm*Δlsr*B	Sm2011 Δ*lsrB* derivative; Sm^r	This work
VO3128	Sm1021 *rpoH1::aadA* derivative; Sm^r, Sp^r	[149]
AB3	Sm1021 *rpoH2:: aacCI* derivative; Sm^r, Gm^r	[149]
AB9	Sm1021 *rpoH1::aadA rpoH2::aacCI* derivative; Sm^r, Gm^r, Sp^r	[149]
Sm*ΔabcR1*	Sm2B3001 Δ*abcR1* derivative	This work
Sm*ΔabcR2*	Sm2B3001 Δ*abcR2* derivative	This work
Sm*ΔabcR1/2*	Sm2B3001 Δ*abcR1/2* derivative	This work
PLASMIDS		
pK18Δ*lsrB*	Suicide plasmid for *lsrB* deletion; Km^r	This work
pK18Δ*abcR1*	Suicide plasmid for *abcR1* deletion; Er^r, Km^r	[91]
pK18Δ*abcR2*	Suicide plasmid for *abcR2* deletion; Er^r, Km^r	[91]
pK18Δ*abcR1R2*	Suicide plasmid for *abcR1/abcR2* deletion; Er^r, Km^r	[91]
pK18Δ*nfeR1*	Suicide plasmid for *nfeR1* deletion; Er^r, Km^r	[97]
pSRK-R1	pSRK derivative constitutively expressing AbcR1	[91]
pSRK-R2	pSRK derivative constitutively expressing AbcR2	[91]
pSRKMS2AbcR1	pSRK derivative constitutively expressing MS2AbcR1	[119]
pSRKMS2AbcR2	pSRK derivative constitutively expressing MS2AbcR2	[119]
pSKiAbcR1	pSRKKm carrying the AbcR1 coding sequence fused to *sinR*-P_{sinI}	This work
pSKiAbcR2	pSRKKm carrying the AbcR2 coding sequence fused to *sinR*-P_{sinI}	This work
pSKiMS2AbcR1	pSKMS2 derivative expressing MS2AbcR1	This work
pSKiMS2AbcR2	pSKMS2 derivative expressing MS2AbcR2	This work
pSKiAbcR1a	pSKiAbcR1 derivative expressing AbcR1 carrying G26G27 substitutions	This work
pSKiAbcR1b	pSKiAbcR1 derivative expressing AbcR1 carrying G55G56 substitutions	This work
pSKiAbcR2a	pSKiAbcR2 derivative expressing AbcR2 carrying G28G29 substitutions	This work
pSKiAbcR2b	pSKiAbcR2 derivative expressing AbcR2 carrying G51G52 substitutions	This work
pBBAbcR1-38*::eGFP*	pBBR1MCS-2 derivative expressing a transcriptional fusion of a truncated *abcR1* promoter (38-bp) to *egfp*; Km^r	This work
pBBAbcR1*::eGFP*	pBBR1MCS-2 derivative expressing a transcriptional fusion of the full-length *abcR1* promoter (334 bp) to *egfp*; Km^r	This work
pBBAbcR2-38*::eGFP*	pBBR1MCS-2 derivative expressing a transcriptional fusion of a truncated *abcR2* promoter (38-bp) to *egfp*; Km^r	This work

pBBAbcR2*::eGFP*	pBBR1MCS-2 derivative expressing a transcriptional fusion of the full-lenght *abcR2* promoter (206 bp) to *egfp*; Km^r	This work
pR*SMc03121::eGFP*	*SMc03121::eGFP* translational fusion (-156/+36 relative to *SMc03121* AUG) ; Ap^r, Tc^r	[97]
pR*SMa0495::eGFP*	*SMa0495::eGFP* translational fusion (-78/+54 relative to *SMa0495* AUG) ; Ap^r, Tc^r	[50]
pR*prbA::eGFP*	*prbA::egfp* translational fusion (-157/+48 relative to *prbA* AUG) ; Ap^r, Tc^r	[50]
pR*SMc02417::eGFP*	*SMc02417::egfp* translational fusion (-372/+78 relative to *SMc02417* AUG) ; Ap^r, Tc^r	This work
pR*SMa0392::eGFP*	*SMa0392::eGFP* translational fusion (-168/+231 relative to *SMa0392* AUG) ; Ap^r, Tc^r	This work
p16*lsrB*	pET16b derivate carrying N-terminally 10xHis-tagged LsrB	This work

Oligonucleotides. Sequences of the oligonucleotides used as probes for Northern hybridization, or as amplification primers for cloning and RT-qPCR are listed in Table 4.

Table 4. Oligonucleotides specifically used in Chapter 1. Restriction sites are underlined.

Oligonucleotides	**5'-sequence-3'**	**Use**
PbAbcR2	GAGGAGAAAGCCGCTAGATGCACCA	Northern blot probing
PbAbcR1/2	AACCTCCAGAGGGGAACAGCTGCTG	
Pb5S	TACTCTCCCGCGTCTTAAGACGAA	
EcoR1uplsrB	CCAGGAATTCGGATCAAGACGGATAGCG	Construction and PCR verification of Sm*ΔlsrB* mutant
BamHIATGlsrB	GAGAGGATCCCCATAAAGGCTCAGCCGGA	
BamHITGAlsrB	CTCTGGATCCACTTCTGACGATCCGTTC	
XbaIdownlsrB	CGACTCTAGAACTGCCCTGCGCAGACG	
checklsrBin	GAACATGCATCGGTCGTCTC	
checklsrBout	CGGTAGAAGACCCATTGGC	
AbcR1OexfusTSSI	GAGCCTGACAGCATCGCTACAGCTGATGCATCTTTGGTG	Construction of pSKiAbcR1,

AbcR2OexfusTSSI	GAGCCTGACAGCATCGCTACAGCT GGTGCATCTAGCGG	pSKiAbcR2, pSKiMS2AbcR1 and pSKiMS2AbcR2
R1G26G27_R	CCAGAAGCCACCAAAGATGC	Construction of pSKiAbcR1a combined with sinR_NdeIF (R) and SecSRK (F)
R1G26G27_F	GCATCTTTGGTGGCTTCTGGTCCC AGTGCCACCGCAGC	
R1G55G56_R	CCGGAACAGCTGCTGCGGTGG	Construction of pSKiAbcR1b combined with sinR_NdeIF (R) and SecSRK (F)
R1G55G56_F	CCACCGCAGCAGCTGTTCCGGTCT GGAGGTTTTAATTACC	
R2G28G29_R	CCAGGAGAAAGCCGCTAGATG	Construction of pSKiAbcR2a combined with sinR_NdeIF (R) and SecSRK (F)
R2G28G29_F	CATCTAGCGGCTTTCTCCTGGCCA GCCGCTGCAGCAGCTGT	
R2G51G52_R	CCAACAGCTGCTGCAGCGGC	Construction of pSKiAbcR2b combined with sinR_NdeIF (R) and SecSRK (F)
R2G51G52_F	GCCGCTGCAGCAGCTGTTGGCCTC TGGAGGTTTGAAACCTT	
XbaIAbcR1	GCCGTCTAGAGCTGATGCATCTTT GGTGGC	P_{abcR1} amplification
PC15Rv	TCTAGAAGCCGCTAGATGCACCTG CT	
EcoRIPC16	CGTCGAATTCTGCCGATAAGCGCC GATA	P_{abcR2} amplification
PC16Rv	TCTAGAGATGCATCAGCTGAGTGT GG	
PR1_50i	CTAGAGATGCATCAGCTGAGTGTG GTATGCTGCTTTTTTGGGCTATCGG CAATCAA	Generation of $P_{abcR1\text{-}38}$ by anneali
PR1_50	CTAGTTGATTGCCGATAGCCCAAA AAAGCAGCATACCACACTCAGCTG ATGCATCT	
PR2_58i	CTAGAAGCCGCTAGATGCACCAGC TGAAAAAGATATGGGTAGGGCCG TAGCCGCTTTCAATAGA	Generation of $P_{abcR2\text{-}38}$ by annealing
PR2_58	CTAGTCTATTGAAAGCGGCTACGG CCCTACCCATATCTTTTTCAGCTGG TGCATCTAGCGGCTT	
SMc02417_F	GCTAGCATCGTTTATGGATTCCAT CC	Amplification of the *SMc02417* 5'-region fused to *eGFP*
SMc02417_R	GGATCCCTCAAGAGCACGCAATTT CG	
a0392_F	GGATCCATCCGGGTTCCGGATCTG	Amplification of the *SMa0392* 5'-region fused to *eGFP*
a0392_R	GCTAGCGACGGCAATCCCGGTCAT	
a0392R1R	GGTCTGGCTGGACCTTCTGG	

a0392R1F	CCAGAAGGTCCAGCCAGACCAGA ACCTGTAATGTCGC	Compensatory mutations in *SMa0392* 5'-region for targeting with AbcR1b or AbcR2b
a0392R2R	GTCCTCTGGCTGGACCTTC	
a0392R2F	GAAGGTCCAGCCAGAGGACAACC TGTAATGTCGC	
a0495R1R	GGTCTCGTTTTTTCTGGTACC	Compensatory mutations in *SMa0495* 5'-region for targeting with AbcR1b or AbcR2b
a0495R1F	GGTACCAGAAAAAACGAGACCGG AATGAACGCAATGAAAAAC	
a0495R2R	GGCCTCTCGTTTTTTCTGGTACC	
a0495R2F	GGTACCAGAAAAAACGAGAGGCC AATGAACGCAATGAAAAAC	
prbAR1R	GGTGTTCCCTTCTTCAGCCG	Compensatory mutations in *prbA* 5'-region for targeting with AbcR1b or AbcR2b
prbAR1F	CGGCTGAAGAAGGGAACACCGGA ATGAGCGATTACAAAG	
prbAR2R	GGCCTGTTCCCTTCTTCAGC	
prbAR2F	GCTGAAGAAGGGAACAGGCCAAT GAGCGATTACAAAGAC	
c02417R1R	GTTCTGGAAATTTGCGCCTC	Compensatory mutations in *SMc02417* 5'-region for targeting with AbcR1b or AbcR2b
c02417R1F	GAGGCGCAAATTTCCAGAACCGG AGCAAACTTATGATG	
c02417R2R	GGGCTTCTGGAAATTTGCGCC	
c02417R2F	GGCGCAAATTTCCAGAAGCCCAGC AAACTTATGATGAAG	
c03121R1R	GGTCCCAAAGGTTCTATGTTC	Compensatory mutations in *SMc03121* 5'-region for targeting with AbcR1a or AbcR2a
c03121R1F	GAACATAGAACCTTTGGGACCAGG ACAAATGCACAAGAAAC	
c03121R2R	GGCAAAGGTTCTATGTTCTAAAG	
c03121R2F	CTTTAGAACATAGAACCTTTGCCA GGAGGACAAATGCACAAG	
LsrB_Fw_NdeI	ATATCATATGGGGGATTCTATGTC GCT	Amplification of the LsrB CDS
LsrB_Rv_BamHI	ATGGATCCTTATCAGAAGTTCCAG TTTCTCG	
aapQ594F	CGCCGCAAGTGTGTTCTTTG	qRT-PCR of *aapQ*
aapQ712R	GAAATGTGACCAGCGGCAGA	
Smc01852F	TCACCAACACTGCCGACTGC	
Smc01852R	TCGTGTGCAGGATGCTGATG	

Construction of *S. meliloti* mutants. Knock-out mutants were generated using the suicide plasmid pK18*mobsacB* as described in Material and methods.

Sm*ΔlsrB* was generated in Sm2011 by a markerless in-frame deletion of the *lsrB* CDS using pK18Δ*lsrB*. To construct pK18Δ*lsrB,* 822-nt and 814-nt DNA fragments flanking the *lsrB* ORF were amplified from genomic DNA with the EcoRIuplsrB/BamHIATGlsrB and BamHITGAlsrB/XbaIdownlsrB primer pairs. PCR fragments were digested with *Eco*RI/*Bam*HI and *Bam*HI/*Xba*I, respectively, and ligated to the pK18*mobsacB Eco*RI and *Xba*I restriction sites, leading to insertion of the tandem fragments via their common *Bam*HI site. Sm2020 (triple *abcR1*, *abcR2*, and *ŋfeR1* deletion mutant) was generated in Sm2019 (derived from Sm2011; Table 1) by successive replacement of the three sRNA loci by a 135-bp erythromycin resistance cassette (SSDUT1) using plasmids pK18Δ*ŋfeR1* and pK18Δ*abcR1R2* [91, 97]. Similarly, Sm*ΔabcR1,* Sm*ΔabcR2* and Sm*ΔabcR1R2* were generated in the parent strain Sm2B3001 (derived also from Sm2011; Table 1) using plasmids pK18Δ*abcR1,* pK18Δ*abcR2*, and pK18Δ*abcR1R2* [91].

EMSA with LsrB. The LsrB CDS was PCR amplified from genomic DNA using the primers LsrB_Fw_ndeI/LsrB_Rv_BamHI and cloned into the vector pET-16b (Novagen) between the *Nde*I/*Bam*HI restriction enzymes sites, yielding p16LsrB encoding a His-tagged LsrB. Recombinant LsrB was produced and purified as described in Material and Methods. The EcoRIPC16/PC16Rv and XbaIAbcR1/PC15Rv primer pairs were used to amplify P_{abcR1} (334-nt) and P_{abcR2} (206-nt), respectively, which were further purified from agarose gels with the GFX™ PCR DNA and Gel Band Purification Kit (GE Healthcare). Binding reactions were performed with 100 nM radiolabeled probes in the absence or presence (1 μM) of purified LsrB, which were then subjected to electrophoresis and analyzed with the Personal FX equipment and Quantity One software (Bio-Rad) as described in Material and Methods.

Construction of plasmids for induced AbcR1/2 expression and tagging. For the IPTG-induced expression of wild-type and MS2 aptamer-tagged

AbcR1/2, we constructed plasmids pSKiAbcR1, pSKiAbcR2, pSKiMS2AbcR1, and pSKiMS2AbcR2. AbcR1 and AbcR2 were amplified from pSRKMS2AbcR1 or pSRKMS2AbcR2 (constitutively expressing tagged AbcR1 or AbcR2) using the PCR1/PCR2 primers [119] (Table 2). PCR products were digested with *Xba*I and *Xho*I and inserted into pSKiMS2 to generate pSKiMS2AbcR1 and pSKiMS2AbcR2. Alternatively, AbcR1 and AbcR2 were amplified from pSRK-R1 or pSRK-R2 (constitutively expressing the wild-type transcripts) using the AbcR1OexfusTSSI/secSRK or AbcR2OexfusTSSI/secSRK primer pairs, respectively [91]. Both forward primers contain a complementary sequence to TSS3_28bp_b_sinIR, which is used together with sinR_NdeIF primer to *sinR*-P_{sinI} amplification (Table 2). The first PCR products were used as template for a second PCR using the primer pairs sinR_NdeIF/secSRK (Table 2). The resulting fragments were restricted with *Nde*I and *Xba*I and inserted into pSRKKm to generate pSKiAbcR1 and pSKiAbcR2.

Replacements of specific nucleotides within aSD1/2 were introduced using a two-step PCR strategy based on overlapping fragments using pSKiAbcR1 or pSKiAbcR2 as the template. The first round of PCR amplifications was performed with sinR_NdeIF or secSRK (both hybridizing to all plasmid templates) and their respective primer pair carrying the desired mutations (Table 4). Each pair of complementary PCR products was used as the template in the second PCR with sinR_NdeIF/secSRK. The resulting products were digested with *Nde*I/*Xba*I and ligated to pSRKKm to yield plasmids pSKiAbcR1a, pSKAbcR1b, pSKAbcR2a, and pSKAbcR2b that were mobilized to *S. meliloti* strain Sm2020 by biparental matings. The affinity purification assays were performed following a previously described protocol (Material and methods).

Sm2020 cells carrying pSKiMS2AbcR1 or pSKiAbcR1 (control of column-binding specificity) were grown in TY and MM media to exponential phase. Bacteria carrying pSKiMS2AbcR2 or the control pSKiAbcR2 were cultured in

TY and MM media to stationary phase, or subjected to temperature and salt upshifts upon growth in MM to exponential phase. RNA preparations were probed with PbAbcR1/2 upon Northern blotting.

Fluorescence reporter assays. The transcriptional fusions reporting promoter activity were generated in the promoterless vector pBB-*eGFP*. Full-length AbcR1 (334-nt) and AbcR2 (206-nt) promoters were amplified with the primer pairs XbaIAbcR1/PC15Rv and EcoRIPC16/PC16Rv, respectively. The PCR products were digested with *Xba*I (P_{abcR1}) or *Hin*dIII/*Xba*I (P_{abcR2}) and cloned into pBB-*eGFP* to generate pBBAbcR1::*eGFP* and pBBAbcR2::*eGFP*. Shorter versions of both promoters (38-nt) were generated by annealing the oligonucleotides PR1_50i/PR1_50 ($P_{abcR1\text{-}38}$) and PR2_58i/PR2_58 ($P_{abcR2\text{-}38}$), and cloning the products into pGEM-T. $P_{abcR1\text{-}38}$ and $P_{abcR2\text{-}38}$ were retrieved from pGEM-T by *Spe*I-*Xba*I restriction, and finally inserted in pBB-*eGFP* to yield pBBAbcR1-38::*eGFP* and pBBAbcR2-38::*eGFP*.

Reporter fusions of *SMc02417* and *SMa0392* to *eGFP* were generated in plasmid pR-*eGFP*. For this, genomic regions of *SMc02417* and *SMa0392* from their respective transcription start sites to the 12th or 77th codons were amplified with the a0392F/a0392R and c02417F/c02417R primer pairs, respectively. The resulting PCR products were digested with *Bam*HI/*Nhe*I and cloned into pR-*eGFP* to yield pR*SMc02417*::*eGFP* and pR*SMa0392*::*eGFP*. Compensatory nucleotide substitutions in all tested target mRNAs (i.e., *SMc03121*, *SMa0495*, *prbA*, *SMc02417*, and *SMa0392*) for regulation by the corresponding AbcR1/2 variants were introduced by a two-step PCR using the respective wild-type reporter fusion as the template. The first round of PCR amplifications were performed with PCR2 or Egfp-139_rev (both hybridizing to the plasmid templates) and their respective primer pair carrying the specific mutations (Table 4). Each pair of complementary PCR products was used as the template in the second PCR with PCR2/Egfp-139. The resulting products were digested with

*Bam*HI/*Nde*HI and ligated to pR-*eGFP* to generate the new set of reporters (the 1 and 2 variants of each wild-type reporter). All reporter plasmids were transferred by biparental conjugation to Sm2020 harboring plasmids expressing either wild-type AbcR1/2, or their a/b variants. Transconjugants for each RNA-target fusion combination were evaluated using a two-plasmid genetic reporter assay *in vivo* described in Material and Methods.

Computational methods. Flux Balance Analysis (FBA) simulations were performed by George C. diCenzo in MATLAB R2019a (mathworks.com) using the SBMLToolbox version 4.1.0 [177], libSBML version 5.17 [178], scripts from the COBRA Toolbox commit 6a99a1e [179], and the iLOG CPLEX Studio 12.9.0 solver (ibm.com). All analyses were performed on the *S. meliloti* metabolic model iGD1348 [150]. For each nutrient condition that was tested, the maximal growth rate and the overall metabolic flux rates of iGD1348 were determined using the 'optimizeCbModel' and 'pFBA' functions of the COBRA Toolbox. Then, each gene belonging to the sRNAs targetome was either individually or simultaneously deleted, and the maximal growth rate and the overall metabolic flux rates determined. sRNA was predicted to regulate transport/metabolism during growth with a C source when deletion of any individual gene or the entire targetome resulted in a growth rate <90% of the wild-type model, or an overall flux rate >110% that of the wild-type model. Analyses in simulated bulk soil and rhizosphere conditions were performed using previously defined nutritional conditions [180].

The 23 complete *S. meliloti* genomes present in the ftp NCBI folder on 2021/02/17 were downloaded and the pangenome was computed by Alessio Mengoni with the 'Pan/Core-Genome' and the 'Gene Phyloprofile' tools of the MicroScope platform [181], with thresholds of 80% aa identity and 80% alignment coverage, as previously reported [182]. The list of sRNA target genes was used as a query to investigate the pattern of expression in *S. meliloti*

transcriptomic data from cultures treated with alfalfa root exudates and luteolin [174] as a proxy of rhizospheric conditions.

Data availability. Raw RNAseq data can be accessed at the URL

Chapter 2. Regulation of nitrogen metabolism by NfeR1

1. Background

NfeR1 along with its five additional putative homologs predicted in the *S. meliloti* reference genome belongs to the so-called αr14 family of α-proteobacterial sRNAs [63, 64]. However, the 123-nt long NfeR1 transcript was the only reliably detected by Northern hybridization of *S. meliloti* RNA extracts [40, 63]. Members of this family share a predicted secondary structure consisting of three hairpins, each carrying identical and ultraconserved "CCUCCUCCC" aSD motifs in their unpaired regions, whilst their respective stems differ highly in the primary nucleotide sequences [63, 64, 183]. *In silico* predictions suggest that all three aSD motifs are equally competent to interact at the translation initiation region of the putative mRNA targets.

The highest NfeR1 levels have been detected in bacteria growing exponentially in glutamate/mannitol medium (MM) and upon a salt shock. Of note, NfeR1 is also highly expressed in all steps of the symbiotic interaction, i.e., rhizoplane colonization, root hair infection, bacteroid differentiation and N fixation [40, 54, 63, 97]. A first alignment of the promoter regions of NfeR1 homologs encoded by diverse *Sinorhizobium*, *Rhizobium*, *Agrobacterium* and *Brucella* representatives evidenced a recognizable RpoD (σ^{70}) -35/-10 signature (CTTAGAC-N_{17}-CTATAT), which is widely conserved in the α-subgroup of proteobacteria [46]. A search for putative additional promoter motifs revealed a conserved upstream 29-nt stretch that is likely the major determinant of the high *nfeR1* transcription rates under stress and symbiotic conditions [97].

Consistently with this accumulation profile, NfeR1 loss-of-function compromises osmoadaptation of free-living bacteria, and nodulation kinetics, nodule development and symbiotic efficiency of *S. meliloti* on alfalfa roots. A large proportion of the nodules induced by a *S. meliloti nfeR1* deletion mutant were apparently colonized by terminally differentiated bacteroids, but looked

round shaped rather than elongated, and smaller than wild-type indeterminate mature nodules [97]. CopraRNA and IntaRNA mostly predict mRNAs encoding ABC transporters as NfeR1 targets, which is reminiscent of AbcR1/2 regulation of nutrient uptake. However, these predictions have not been experimentally verified. Microarray-based transcriptomics upon an osmotic upshift revealed that a lack of NfeR1 alters expression of a wide set of genes functionally linked to cell processes underlying an adequate response to high salinity[97]. In MM broth, genes for the aerobic nitrate assimilation were also misregulated by the *nfeR1* knock-out (e.g. *glnII, narB, narK, nirB, nirD*). Nevertheless, IntaRNA does not predict thermodynamically favored NfeR1 base-pairing to the vast majority of these differentially accumulated mRNAs, and therefore, they must be regarded as secondary indirect targets of NfeR1 regulation. The search for mRNAs directly regulated by antisense interactions to NfeR1 thus remains opened.

It is known that glutamate is a poor N source and, therefore, MM formulated with this amino acid imposes a N stress to bacteria [184]. In *S. meliloti*, regulation of the Nitrogen Stress Response (NSR) involves a sensor protein GlnD, two PII proteins (GlnK and GlnB), and the NtrBC two component regulatory system. Under N surplus, deuridylylated GlnB interacts with the C-terminal domain of the histidine kinase NtrB to inhibit its autophosphorylation while activating its phosphatase activity, leading to dephosphorylation of NtrC. When α-ketoglutarate levels are high because of N deficiency, GlnD uridylylates GlnB, thereby promoting NtrB-mediated NtrC phosphorylation [185–187]. NtrC is a ubiquitous and well-characterized bacterial regulator, whose phosphorylated form (NtrC-P) is active in binding upstream enhancers at RpoN (σ^{54})-dependent promoters, thereby inducing transcription of large sets of N assimilation genes [188, 189]. NtrC-induced genes include those encoding the ammonium transporter AmtB, diverse amino acid transporters, and the gutamine synthetases (GS) GSI and GSII (*glnA* and *glnII* genes, respectively). NtrC-dependent

transcriptional repression has been more rarely described but it likely occurs as a mechanism of negative autoregulation of the *S. meliloti dusBntrBC* operon [184, 190, 191].

In this Chapter, we have investigated the transcriptional regulation of *S. meliloti* NfeR1 and characterized its mRNA interactome using MAPS. Our data indicate that LsrB promotes a seemingly constitutive NfeR1 transcription, which is downregulated by NtrC-mediated repression under N surplus conditions. NfeR1 accumulation in N deficient media likely strengthens the *S. meliloti* NSR by alleviating the (auto)repression of the *ntrBC* bicistronic mRNA.

2. Results

2.1. Determinants of NfeR1 transcription

As a first approach to identify putative regulatory proteins that bind the NfeR1 promoter (P_{nfeR1}), we carried out a DNA affinity chromatography-pulldown assay with biotinylated DNA fragments that are retained in streptavidin columns. As the DNA bait, a full-length P_{nfeR1} ($P_{nfeR1\text{-}213}$) was amplified by PCR from -213 to +14 positions (relative to the annotated NfeR1 TSS) using a biotinylated forward primer. As control, we used a DNA fragment in which a 60-nt stretch (positions -40 to -100 in $P_{nfeR1\text{-}213}$) containing the conserved motif unveiled by a previous promoter alignment was deleted ($P_{nfeR1\Delta}$) (Fig. 22A). Biotinylated $P_{nfeR1\text{-}213}$ and $P_{nfeR1\Delta}$ probes were mixed with cell lysates from wild-type strain SmB23001 growing exponentially in MM and upon a salt shock, both promoting endogenous upregulation of NfeR1, and complete TY, in which NfeR1 is barely detected (Fig. 22B). Proteins with different binding affinities for the DNA baits were sequentially eluted from columns by the addition of increasing concentrations of NaCl. Aliquots of the eluted fractions were first fractionated by

SDS-PAGE and proteins visualized by silver staining. (Fig. 22B). Preparative SDS-PAGE was then conducted and readily detectable bands were sliced out from

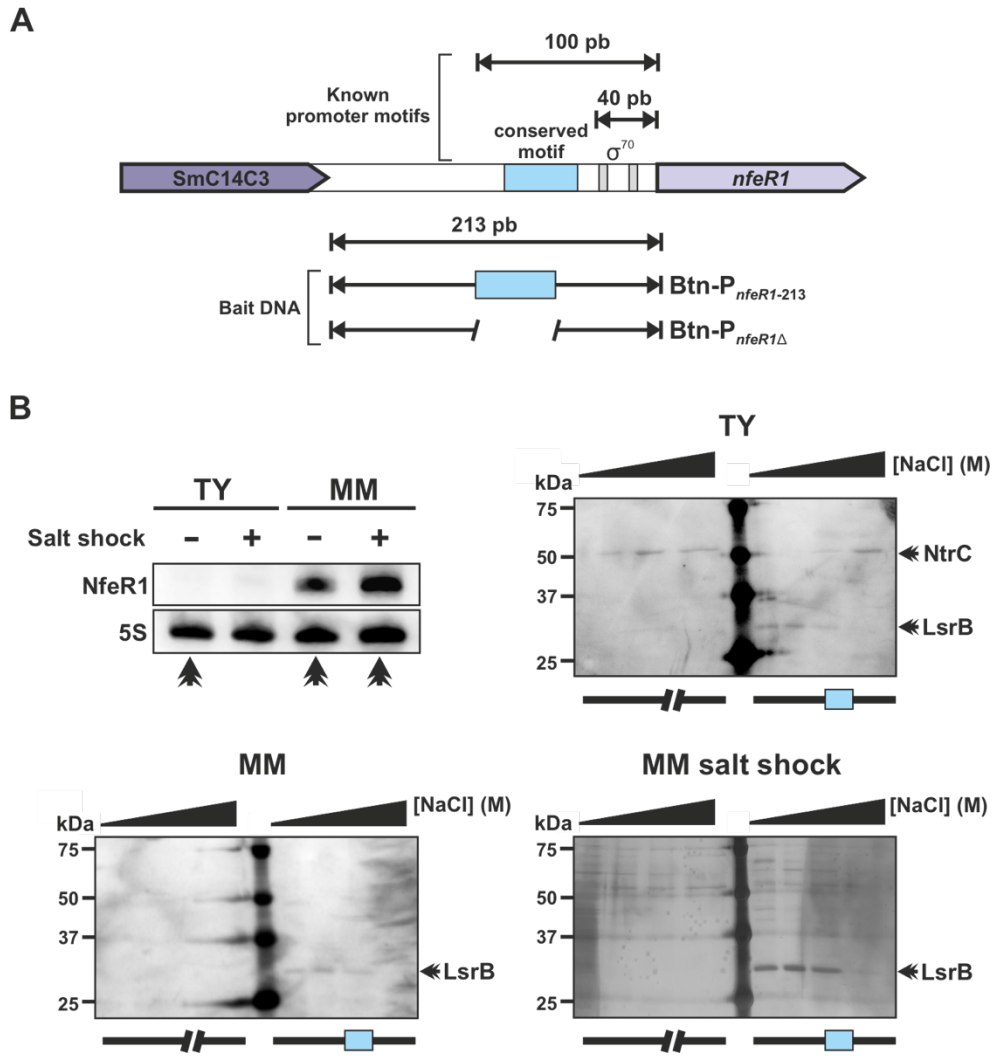


Figure 22. Affinity chromatography pull down assay with P_{nfeR1} (A) Scheme of the NfeR1 promoter region containing the conserved σ^{70} binding site (grey boxes) and the conserved motif responsible of P_{nfeR1} induction (blue box). The biotinylated DNA probe for the pull down assay contained the 213-nt long intergenic region SmC14C3-*nfeR1* (Btn-$P_{nfeR1-213}$). The control bait DNA lacks the conserved motif. **(B)** Top left panel: Northern blot probing of NfeR1. Total RNA was obtained from Sm2B3001 cultured in the conditions indicated along the top of the panel. The 5S rRNA was probed as an RNA loading control. Lysates for pull down assays were obtained from cultures in the conditions indicated by arrows i.e., TY, MM, MM upon salt shock (400 mM for 1 h). Affinity chromatography eluted fractions by different salt concentrations (0.3, 0.8, 1.5, 2 mM NaCl) were analyzed by SDS-PAGE followed by silver staining. Precision Plus Protein Dual Color Standard was run in the middle lane.

the Coomassie stained gel, and further analyzed by liquid chromatography mass spectrometry (LC-MS/MS). This analysis identified LsrB (36 kDa) as the protein bound to $P_{nfeR1\text{-}213}$ but not to the control $P_{nfeR1\Delta}$ in all culture conditions. Additionally, the 54 kDa transcriptional regulator NtrC was identified in the analysis. Of note, NtrC was recovered as partner of both $P_{nfeR1\text{-}213}$ and $P_{nfeR1\Delta}$ specifically in lysates from TY cultures, suggesting binding either upstream or downstream of the identified conserved promoter motif.

Therefore, we performed a new alignment with sequences extending up to 213-nt upstream the TSS of several NfeR1 homologs in *Sinorhizobium* and *Rhizobium* species (Fig. 23). A more detailed inspection of the conserved -72 to -60-nt stretch revealed matching to the LTTRs box, T-N_{11}-A, which resembles the LsrB binding site identified in the AbcR1 promoter. This new alignment also

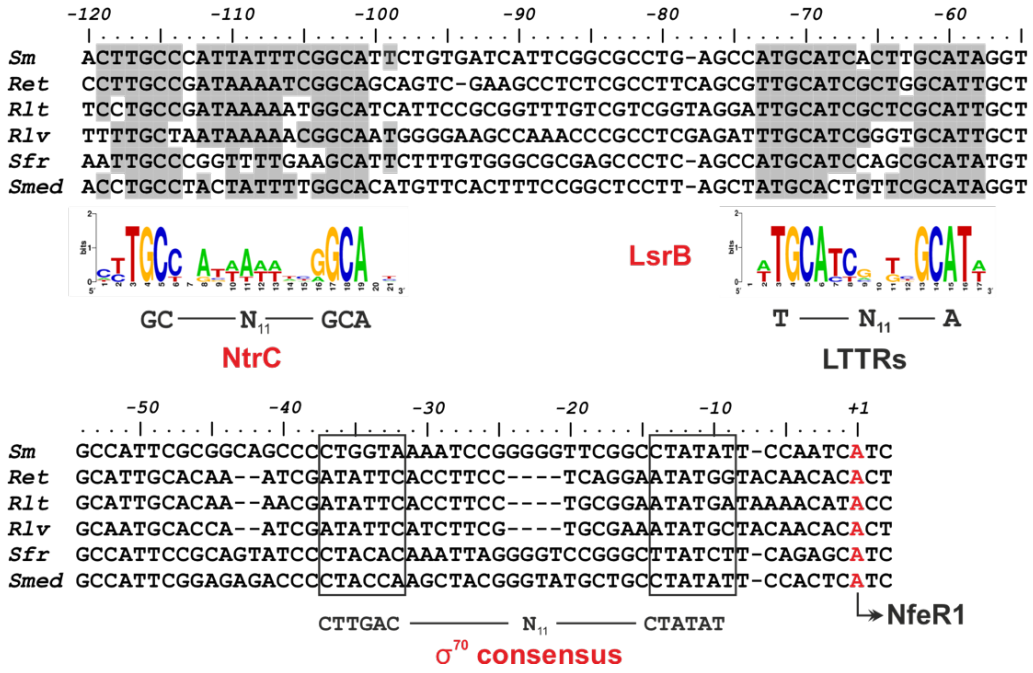


Figure 23. Multiple sequence alignment of NfeR1 promoter homologs in α-proteobacteria. Concensus sequences og the conserved motifs are indicated below the alignments. *Sm*; *Sinorhizobium meliloti* Sm1021; *Ret*, *R. etli* CIAT652; *Rltr*, *R. leguminosarum* bv. *trifolii*

WSM2304; *Rlv*, *Rizobium leguminosarum* bv. *viciae* 3841; *Sfr*, *S. fredii* HH103; *Smed*, *S. medicae* WSM419.

evidenced a putative NtrC binding motif GC-N_{11}-GCA between positions -101 to -116 in *S. meliloti* P_{nfeR1}, which has been described in both NtrC activated and repressed promoters [192–194]. Besides, the conserved A/T-rich region within the putative NtrC binding site has been also reported in the promoter of genes regulated by NtrC in *Salmonella* sp. [192].

To further assess LsrB and NtrC binding to P_{nfeR1}, we performed EMSA with different radiolabeled versions of the promoter (Fig. 24). These experiments unambiguously revealed LsrB binding to a 100-nt long promoter ($P_{nfeR1\text{-}100}$) (Fig. 24A). Conversely, both trimming of ($P_{nfeR1\text{-}40}$) or point mutations (P_{nfeR1*}) within the proposed LsrB-binding motif abrogated LsrB binding to P_{nfeR1}. Incubation with NtrC resulted in an electrophoretic mobility shift of $P_{nfeR1\text{-}213}$ but not of $P_{nfeR1\text{-}100}$, consistently with the location of the predicted NtrC binding site. The formation of the NtrC-P_{nfeR1} complex was also evident upon incubation of NtrC with lithium carbamoyl phosphate, suggesting that binding is independent of the protein phosphorylation status (Fig. 25B). DNA footprinting unveiled a nucleotide stretch protected against DNase I digestion that maps the NtrC interaction site between positions -95 and -121 at P_{nfeR1}, as expected (Fig. 25C). Collectively, these data anticipate that NtrC and LsrB regulate NfeR1 transcription by binding at almost contiguous sites within the sRNA promoter.

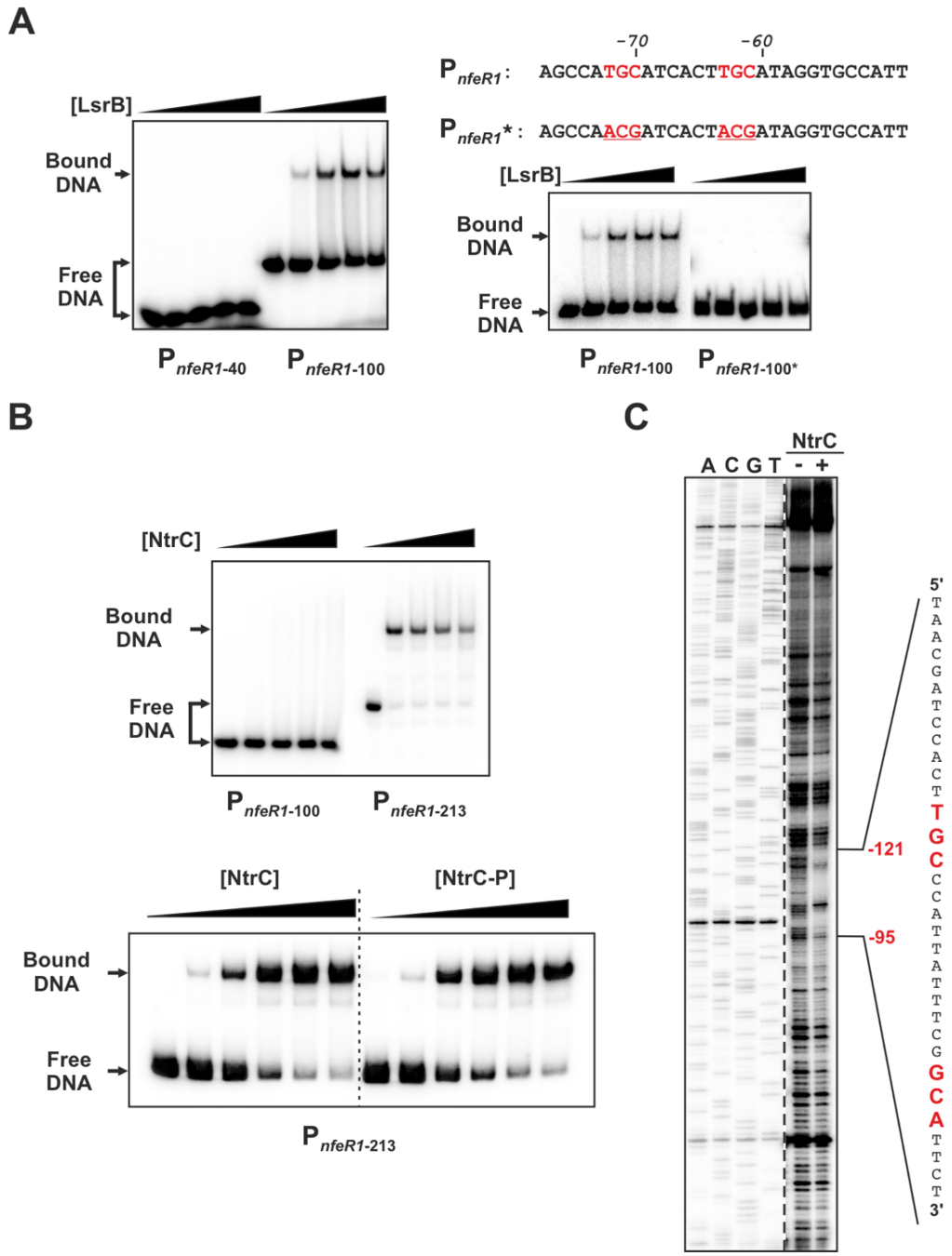


Figure 25. LsrB and NtrC *in vitro* binding to NfeR1 promoter. Figure legend on the next page.

Figure 25. LsrB and NtrC binding *in vitro* to NfeR1 promoter. (A) LsrB binds the *nfeR1* promoter at the conserved motif responsible for NfeR1 induction. Left, gel shift assays with radiolabelled $P_{nfeR1\text{-}100}$ (100 bp) and $P_{nfeR\text{-}40}$ (40 bp) incubated with increasing concentrations of purified LsrB (0, 0.5, 1, 2, 4 μM). Right, gel shift assays with $P_{nfeR1\text{-}100}$ mutant variants. Nucleotide substitutions in $P_{nfeR1\text{-}100}$ to generate $P_{nfeR1\text{-}100*}$ are marked in red**. (B)** NtrC binds the NfeR1 promoter upstream the LsrB-binding site. Gel shift assays with radiolabelled $P_{nfeR1\text{-}100}$ (100 bp) and $P_{nfeR1\text{-}213}$ (213 bp) incubated with increasing concentrations of purified NtrC (0, 0.25, 0.5, 0.75, 1 μM, upper panel; 0, 10, 50, 100, 200, 300 nM, lower panel). NtrC-P, phosphorylated NtrC upon incubation with lithium carbamoyl phosphate. **(C)** Footprinting of NtrC. $P_{nfeR1\text{-}213}$, 5′ end-labeled on the bottom strand, was incubated without (-) and with NtrC (1 μM; +) After partial digestion with DNase I, the reactions were subjected to urea-PAGE. The nucleotide sequence protected by NtrC, as inferred by the sequencing ladder from the $P_{nfeR1\text{-}213}$ fragment (A, C, G, T lanes), is indicated to the right.

2.2. LsrB induces whereas NtrC represses NfeR1 transcription

We therefore investigated the effects of LsrB and NtrC on NfeR1 transcription *in vivo*. For that, we fused DNA fragments encompassing 213, 100, and 40-nt upstream the NfeR1 TSS to a promoterless *eGFP*, i.e., $P_{nfeR1\text{-}213}$*::eGFP*, $P_{nfeR1\text{-}100}$*::eGFP*, and $P_{nfeR1\text{-}40}$*::eGFP,* respectively. The latter two fragments were generated by sequential trimming of the NtrC and LsrB binding sites in $P_{nfeR1\text{-}213}$ (Fig. 26A). These reporter transcriptional fusions were independently mobilized to the wild-type *S. meliloti* Sm2B3001 strain, and fluorescence was scored in TY and MM cultures. As expected, the transconjugants carrying $P_{nfeR1\text{-}40}$*::eGFP* (i.e., lacking both the NtrC and LsrB binding motifs) yielded hardly perceptible fluorescence, whereas the highest *eGFP* expression was evident in bacteria harbouring $P_{nfeR1\text{-}100}$*::eGFP* in both culture conditions (Fig. 26A). The NfeR1 transcript is hardly detectable in bacteria grown in TY medium (Fig. 22B; upper left panel). However, $P_{nfeR1\text{-}100}$ activity was far above that of $P_{nfeR1\text{-}40}$ in this condition, suggesting active LsrB-dependent NfeR1 transcription. Fuorescence measured in the reporter strain expressing *eGFP* from $P_{nfeR1\text{-}213}$ decreased by ~10-fold with respect to that of $P_{nfeR1\text{-}100}$, being close to the background derived from the promoterless construct (P-*eGFP*), which evidences a marked transcriptional repression of the promoter by NtrC. In MM, this relative repression of the promoter (i.e., $P_{nfeR1\text{-}100}$ vs. $P_{nfeR1\text{-}230}$-derived fluorescence) was scarcely ~2.4-fold.

However, $P_{nfeR1\text{-}230}$ activity was 8-fold higher in bacteria grown in MM than in TY medium, which fairly correlates with the steady-state levels of NfeR1 detected by Northern blot probing.

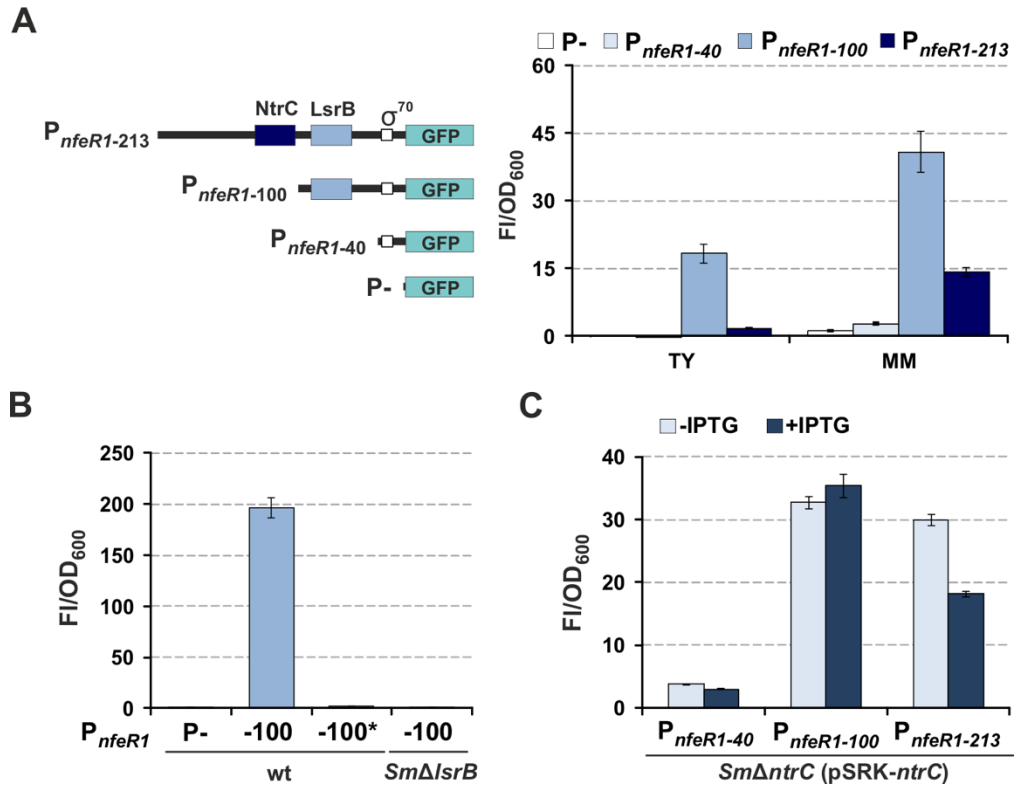


Figure 26. Transcriptional regulation of NfeR1 by LsrB and NtrC. **(A)** Fluorescence of promoter-*eGFP* fusions in the pABCa plasmid. Fluorescence derived from full-length and trimmed versions of P_{nfeR1} (pABCa plasmid), as diagrammed, were determined in Sm2B3001 growing in TY or MM. **(B)** LsrB-dependent P_{nfeR1} activity. pBB-$P_{nfeR1\text{-}100}$-*eGFP*-derived fluorescence was measured in wild-type and LsrB mutant strain (Sm2011 and Sm*Δlsr*B, respectively). *, indicates $P_{nfeR1\text{-}100}$ containing the mutations as described in Fig. 25. **(C)** NtrC-dependent P_{nfeR1} activity. Fluorescence derived from full-length and trimmed versions of P_{nfeR1}-*eGFP* fusions in the pABCa vector were determined in Sm2B3001Δ*ntrC* transformed with pSRK-NtrC. Reporter bacteria were cultured in MM with or without IPTG. Reported values are means and SD of nine fluorescence measurements normalized to the OD_{600} of cultures, i.e., three replicates of three independent cultures of each reporter strain.

We next addressed the transcriptional regulation of NfeR1 genetically. For that, we transformed the deletion mutant derivatives Sm*Δlsr*B and Sm*Δntr*C with the reporter transcriptional fusions. Both, *lsrB* knock-out and point mutations at

the LsrB binding site, abrogated transcription from $P_{nfeR1\text{-}100}$ (Fig. 26B). In SmΔ*ntrC*, *ntrC* induced expression from pSRK-NtrC decreased $P_{nfeR1\text{-}213}$-derived fluorescence by 1.5-fold in MM cultures whilst $P_{nfeR1\text{-}100}$ (i.e, lacking the NtrC binding motif) activity was independent of NtrC (Fig. 26C). Together, these data indicate that LsrB is absolutely required for NfeR1 expression, whereas NtrC acts as transcriptional repressor.

2.3. NfeR1 is upregulated under N stress conditions

Since NtrC is a well-known transcriptional activator of a large set of genes involved in the NSR, we reasoned that NfeR1 might respond to the N status in *S. meliloti*. To text this hypothesis, the wild-type strain and its *ntrB* deletion mutant were transformed with $P_{nfeR1\text{-}213}$*::eGFP*, and the transconjugants were grown in TY and MM, which was formulated with different N sources at likely limiting or excess concentrations. As N stress reporter, we used the transcriptional fusion of the NtrBC-dependent *glnII* promoter to *eGFP*, which was mobilized to the same strains (Fig. 27A). P_{glnII}*::eGFP*-derived fluorescence in the mutant lacking NtrB was thus considered background of the NSR outcome in each culture condition. Fluorescence from $P_{nfeR1\text{-}213}$*::eGFP* was more than 7-fold higher in bacteria grown in the classical MM (formulated with 5.4 mM glutamate as the N source) than in complete TY, and increased as the N concentration decreased in MM-NO_3 and MM-NH_4, as expected. Remarkably, the highest transcription levels of *eGFP* from $P_{nfeR1\text{-}213}$ were measured in conditions of N stress as reported by P_{glnII}. Of note, $P_{nfeR1\text{-}213}$-derived fluorescence was not influenced by the *ntrB* knock-out, further suggesting that repression of NfeR1 transcription does not require NtrB-dependent phosphorylation of NtrC.

Northern blot probing of RNA extracts from wild-type bacteria grown in the same conditions evidenced a correlation between the NfeR1 accumulation pattern and the reporter activities, further confirming upregulation of NfeR1 under N

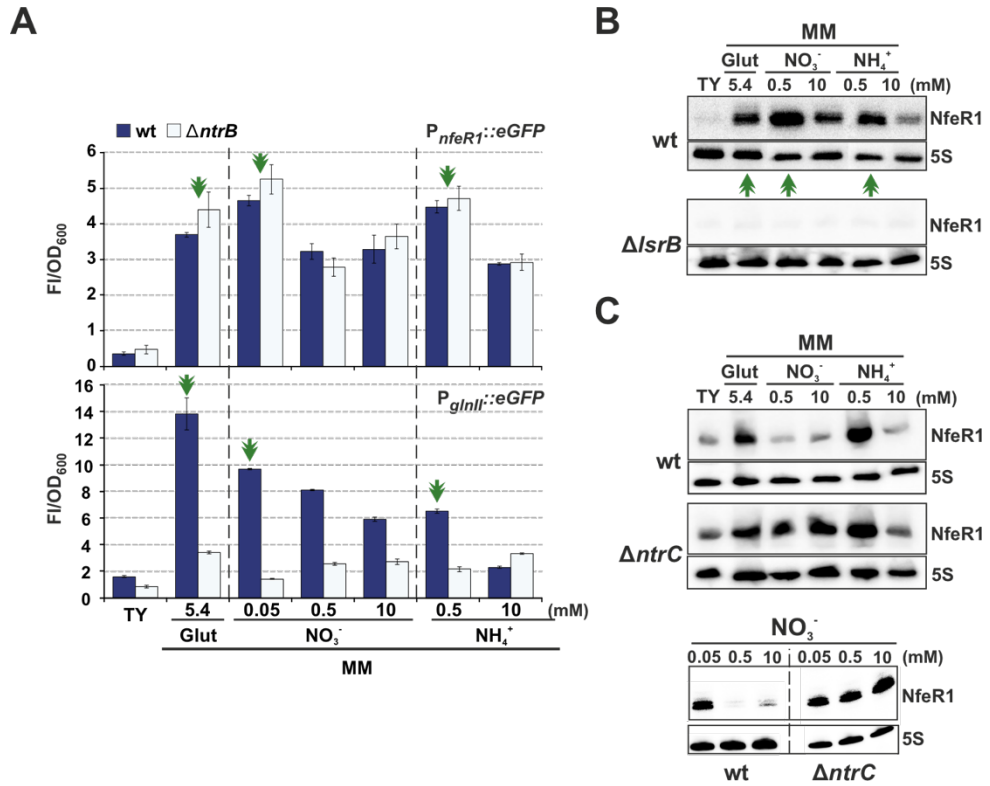


Figure 27. NfeR1 is a N stress-induced sRNA. (A) Activity of NfeR1 and *glnII* promoters under N stress. Fluorescence derived from $P_{nfeR1\text{-}213}$ and P_{glnII}-*eGFP* fusions were determined in Sm2B3001 (wild-type) and Sm2B3001Δ*ntrB* growing in MM supplemented with different N sources and concentrations as indicated. TY is regarded as rich-medium (N surplus). **(B)** Northern blot analysis of N-dependent NfeR1 expression. Total RNA was obtained from Sm2B3001 or Sm2B3001Δ*lsrB* cultured in conditions indicated along the top of the panel. Green double arrowsheads indicate conditions that promote highest NfeR1 accumulation **(C)** Northern blot analysis of NtrC regulation of NfeR1. For total RNA extraction, Sm2B3001 or Sm2B3001Δ*ntrC* were grown in TY to exponential phase, washed in PBS and cultured in indicated conditions for 4 h. The 5S rRNA was probed as an RNA loading control.

stress (Fig. 27B). Moreover, NfeR1 was undetectable in the SmΔ*lsrB* strain indicating again that LsrB is indispensable for the expression of the sRNA in all conditions. Since a lack of NtrC compromises *S. meliloti* growth under N stress, to assess the impact of NtrC on NfeR1 steady-state levels, wild-type and SmΔ*ntrC* cells were first cultured in TY until exponential phase (OD_{600} 0.8), washed in PBS solution, resuspended in different modified MM and grown for further 4 h in the fresh media (Fig. 27C). Probing of total RNA from the cultures confirmed NfeR1

wild-type upregulation under N stress, and seemingly constitutive accumulation of the transcript in Sm*ΔntrC*. It is worth noting that in this series of experiments nitrate concentration had to be reduced to 0.05 mM to detect NfeR1 expression.

We therefore conclude that LsrB operates as an apparently constitutive NfeR1 activator, whereas NtrC constraints transcription of the sRNA under N surplus.

2.4. NfeR1 is required for a wild-type NSR

To assess the impact of NfeR1 in the NSR, we first used RT-qPCR to measure *ntrB*, *ntrC* and *glnII* levels in MM log cultures of Sm2020 (i.e., triple AbcR1/2 and NfeR1 knock-out) harboring pSKiNfeR1 (Fig. 28A). IPTG-induced expression of NfeR1 during 1 h provoked upregulation of the three mRNAs, which was particularly evident in the case of *ntrB*. We further examined expression of these key NSR genes in bacteria lacking only NfeR1 (Sm*ΔnfeR1*) grown in MM-NH_4 to mimic either N stress (0.5 mM NH_4Cl) or N excess (10 mM NH_4Cl) (Fig. 28B). Since NtrC is necessary for *glnII* induction, and it has been reported that negatively regulates its own operon *dusBntrBntrC* [184, 193, 195], we also conducted RT-qPCR in the Sm*ΔntrC* mutant. Wild-type and mutant bacteria precultured in TY were washed, resuspended in MM-NH_4 at 0.5 mM or 10 mM and grown for further 4 h before RNA extraction. As expected, the three genes were strongly upregulated under N stress in wild-type bacteria, and *glnII* was almost undetectable in bacteria lacking NtrC at both ammonia concentrations. Additionally, *ntrB* levels were markedly upregulated in Sm*ΔntrC* under N excess conditions, consistent with NtrC acting as repressor of the operon. Confirming data described above (Fig. 28B), the absence of NfeR1 resulted in downregulation of *ntrB*, *glnII* and, more subtly, of *ntrC*. Taken together, these findings indicate that NfeR1 contributes positively to the NSR in *S. meliloti*.

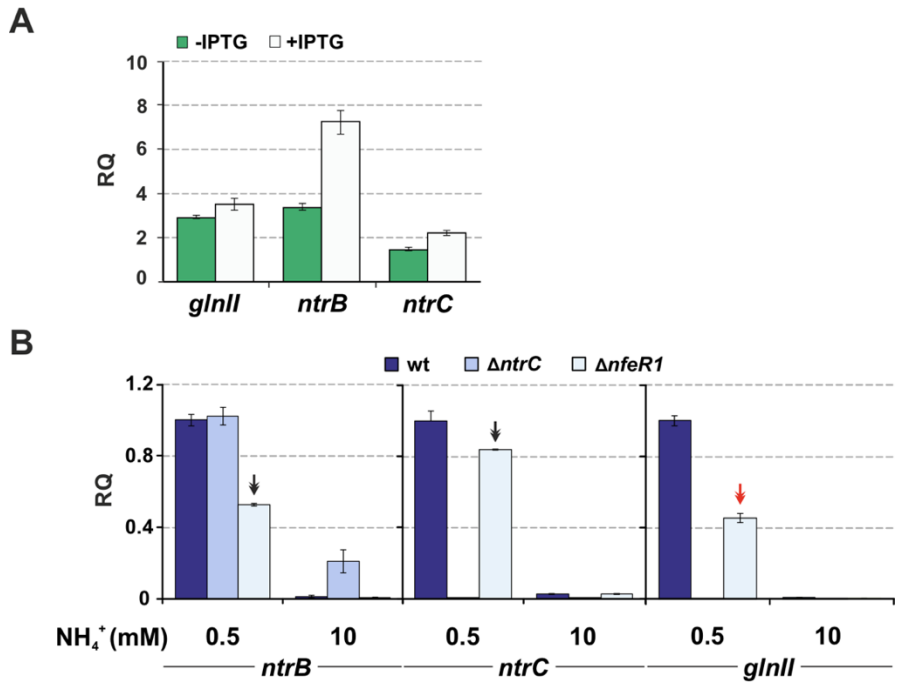


Figure 28. NfeR1 is required for full expression of *ntrBC* and *glnII*. (A) RT-qPCR analysis of NfeR1-dependent *ntrBC* and *glnII* mRNA abundance. RNA was xtracted from Sm2020 transformed with pSKiNfeR1 1 h after IPTG-induction of NfeR1 in MM broth. **(B)** RT-qPCR analysis of *ntrBC* and *glnII* abundance in wild-type Sm2B3001 and derived mutants lacking NtrC or NfeR1. For total RNA extraction, bacteria were grown in TY to exponential phase, washed in PBS and cultured in MM supplemented with 0.5 or 10 mM ammonia for 4 h. Arrows indicate abundance of mRNAs of the Ntr system (black) and of the NSR marker *glnII* (red) in the NfeR1 mutant. Relative Quantification (RQ) values were normalized to *SMc01852* as a constitutive control. Values plotted in the bar graphs are means and SE of three replicates of two independent cultures.

2.5. The NfeR1 aSD motifs have redundant function in target mRNA regulation

To identify putative mRNA partners of NfeR1, we first used the CopraRNA algorithm. 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 initial predictions. NfeR1 is expected to interact at the RBS of its mRNA targets through the three ultraconserved aSD motifs that remain single-stranded within each stem

loop (aSDa-c) of the αr14 family members (Fig. 29). Indeed, CopraRNA returned a large list of putative mRNAs probably targeted by the NfeR1 aSD seeds, further revealing a functional enrichment in mRNAs encoding ABC transport proteins.

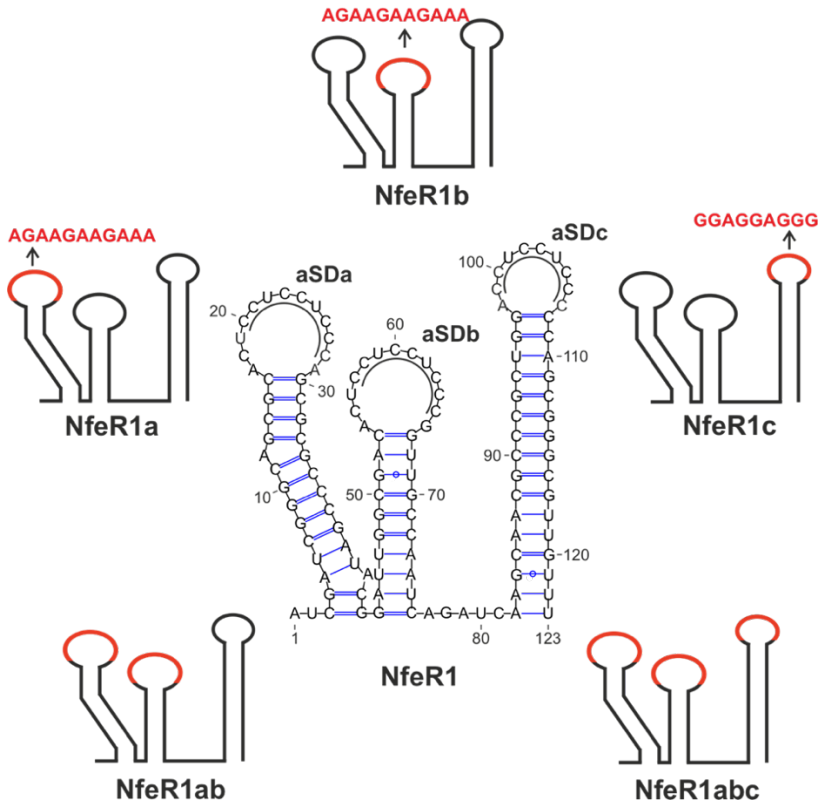


Figure 29. Predicted secondary structure of wild-type NfeR1 sRNA and its mutant variants. The three NfeR1 stem-loops carry identical aSD motifs ("CCUCCUCCC"; aSDa-c) within their unpaired region (underlined). Mutation of aSDs were generated by nucleotide replacements that are indicated in red. Single (NfeR1a/b/c), double (NfeR1ab) and triple (NfeR1abc) mutations of the aSD motifs are not predicted to affect NfeR1 folding.

We selected two out of the 20 top ranked mRNA targets, *SMb20442* and *SMc03121*, both encoding the periplasmic component of ABC transport systems, for further experimental validation by the double-plasmid reporter assay. The *SMc03121* mRNA, which encodes a yet uncharacterized periplasmic transport protein, was already validated as AbcR1/2 target in this work, whereas regulation of *SMb20442*, whose protein product likely binds monosaccharides, had not been

investigated. *SMc03121::eGFP* and *SMb20442::eGFP* reporters were mobilized to a Sm2020 s train harboring pSKiNfeR1. The activity of both translational

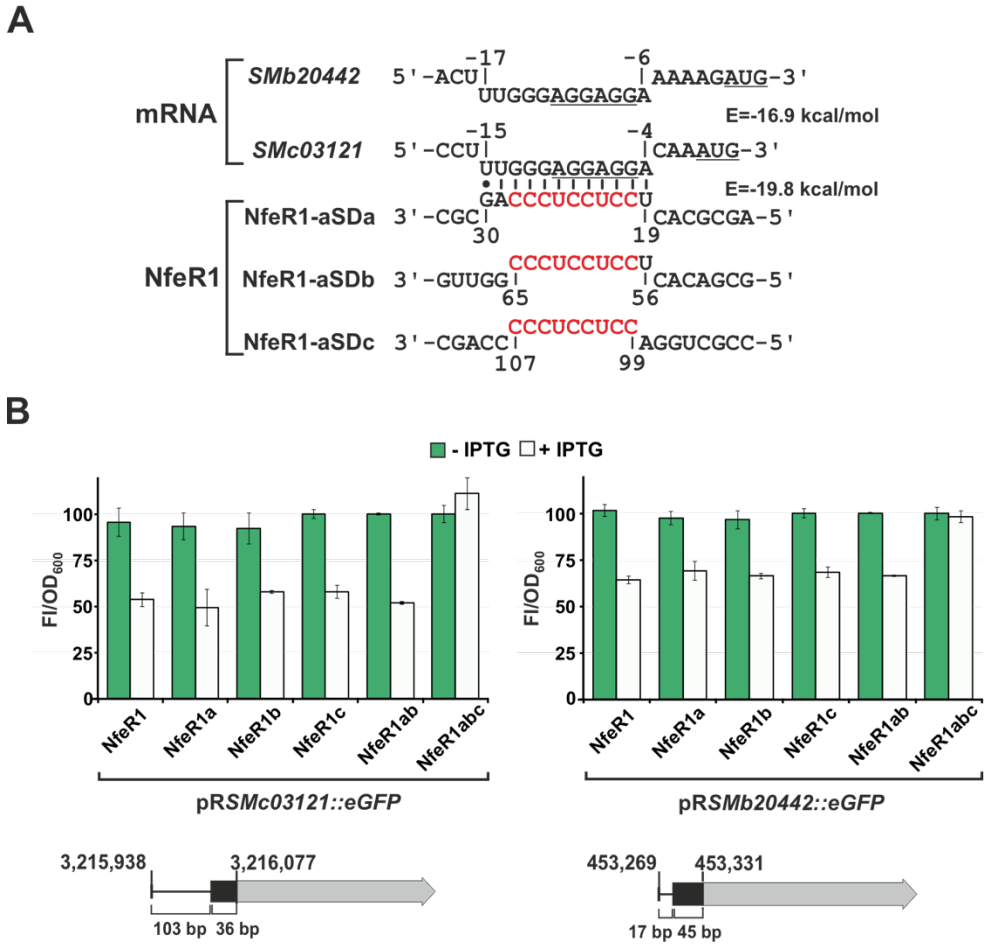


Figure 30. NfeR1-mediated regulation of mRNAs encoding periplasmic ABC transport proteins. (A) IntaRNA predicted base-pairing interactions between NfeR1 aSDa-c region (in red), and the *SMb20442* and *SMc03121* mRNAs. The SD sequence and AUG start codon of each mRNA are underlined. Numbers denote nucleotide positions relative to the start codon of the mRNAs or the NfeR1 TSS. The predicted minimum hybridization energy (*E*) between NfeR1-aSDa and the mRNAs are indicated. Energy values for interactions involving aSDb and aSDc are similar. **(B)** Fluorescence of reporter strains co-transformed with the translational fusions diagramed on the bottom and plasmids overexpressing the wild-type NfeR1 or its mutant variants (NfeR1a-c, NfeR1ab or NfeR1abc) upon IPTG-mediated induction. Numbers in the diagrams stand for coordinates within the Rm1021 genome, and base pairs of the 5'-UTR (lines) and CDS (black boxes) of each mRNA fused to *eGFP*. Values plotted in the histogram correspond to the means and SD of 27 fluorescence measurements normalized to the OD_{600} of the cultures, i.e. three determinations of three independent cultures of three independent double transconjugants for each reporter strain.

fusions upon IPTG-induced expression of NfeR1 decreased by 46% and 38%, respectively, with respect to uninduced cultures. These results indicate post-transcriptional downregulation of *SMc03121* and *SMb20442* mRNAs by NfeR1, most likely by interfering with translation. In both cases, the putative interaction region of NfeR1 mapped specifically to either of the 11-nt, 10-nt or 9-nt aSD motifs (aSDa-c), (Fig. 30A). To test the contribution of each interaction site to target regulation, we generated single (NfeR1a/b/c), double (NfeR1ab) and triple (NfeR1abc) NfeR1 mutant variants, in which the three aSD seeds were individually or simultaneously replaced by non-wild type sequences (Fig. 29). These nucleotide substitutions most likely disrupt base-pairing with *SMc03121* and *SMb20442* mRNAs, while preserving the predicted secondary structure of NfeR1. The single and double mutants retained the NfeR1 ability to reduce *SMc03121::eGFP* and *SMb20442::eGFP*-derived fluorescence from both translational fusions upon IPTG addition (Fig. 30B). However, the simultaneous mutation of all three interaction sites (NfeR1abc) disabled NfeR1 for target mRNA regulation.

Altogether, these data validate *SMc03121* and *SMb20442* mRNAs as NfeR1 targets, further uncovering redundant regulatory functions of the three ultraconserved aSD motifs, which likely act as interaction seeds for targeting.

2.6. Tagging of NfeR1 and MAPS setup

To further understand NfeR1 function we performed MAPS with a transcript tagged at the 5'-end with the MS2 aptamer (Fig. 31A). Intriguingly, Northern blot probing of bacteria expressing MS2NfeR1, revealed accumulation of both wild-type and tagged RNA species at similar levels (Fig. 31B). In this construct, NfeR1 was preceded by a *Xba*I recognition site, which is followed by the TC nucleotides (first nucleotides of *nfeR1*), thus generating a GATC site that might alter gene expression since it is sensitive to methylation (New England Biolabs) (Fig. 31B)

or, alternatively, promote MS2NfeR1 processing. Thus, we replaced one or three nucleotides within the *Xba*I site (i.e., MS2NfeR1_1 and MS2NfeR1_3, respectively), and probed RNA from bacteria expressing these variants. Only the triple mutation fully abolished the accumulation of the wild-type NfeR1 transcript and therefore, this variant (from here on MS2NfeR1) was used in the MAPS assay.

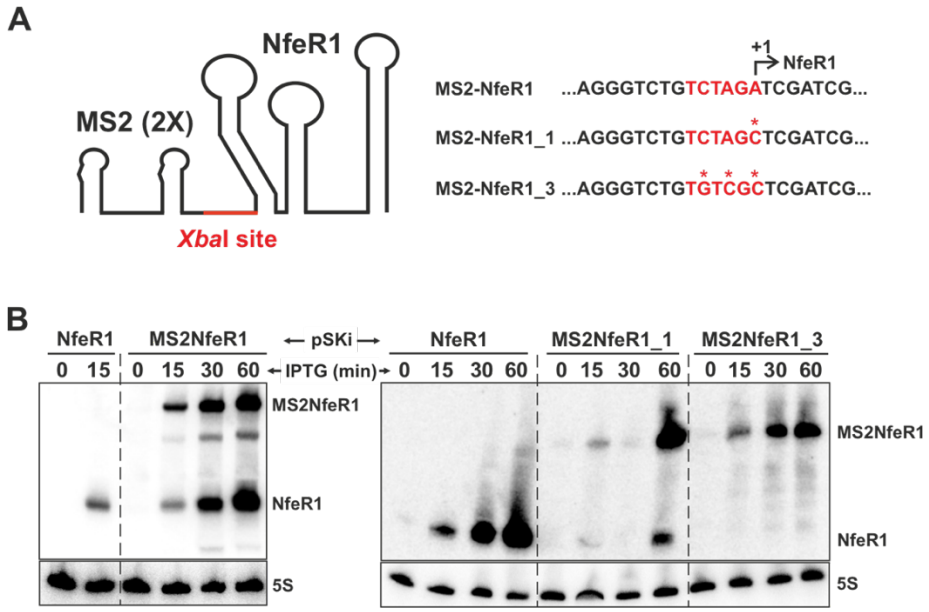


Figure 31. Construction of pSKiMS2NfeR1 overexpressing the MS2-tagged NfeR1. (A) MS2 aptamer fusion to the 5'-end of NfeR1 using the *Xba*I site (marked in red). The nucleotide sequence is showed to the right. Asterisks indicate nucleotides which were replaced to abolish generation of wild-type NfeR1 accumulation. **(B)** Northern blot analysis of MS2NfeR1 expression. Total RNA was obtained from Sm2020 cultures in MM upon IPTG-addition for 0, 15, 30 or 60 min. Probing of RNA from bacteria transformed with pSKiNfeR1 or pSKiMS2NfeR1 (initial construct) vectors is shown in the left panel. Right panel, probing of RNA from bacteria conjugated with pSKiNfeR1, pSKiMS2NfeR1_1 or pSKiMS2NfeR1_3. The 5S rRNA was probed as an RNA loading control.

For MAPS, pSKiNfeR1 and pSKiMS2NfeR1 were mobilized to Sm2020, and expression of wild-type and tagged NfeR1 were induced with IPTG in bacteria grown exponentially under N stress (i.e., MM-Glutamate and 0.5 mM MM-NH_4). Cell lysates were then subjected to affinity chromatography. RNA from the input and output chromatography fractions was analyzed by Northern

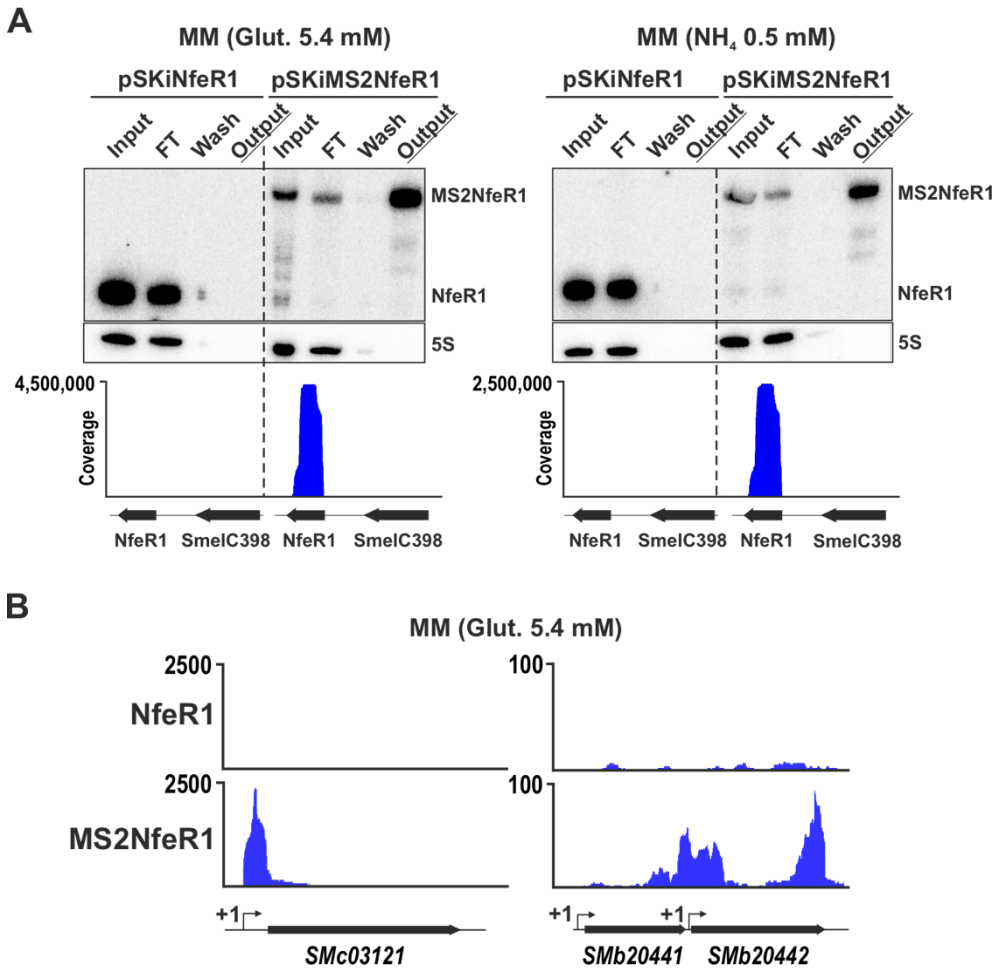


Figure 32. NfeR1 MAPS setup. (A) Monitoring of affinity chromatography in MM with glutamate and ammonia as N sources. Expression of wild-type and MS2NfeR1 was induced for 15 min with IPTG in Sm2020 transformed with pSKiNfeR1 or pSKiMS2NfeR1, respectively. RNA from input, flowthrough (FT), wash, and output chromatography fractions (as indicated on top of panels) was probed with PbNfeR1. 5S rRNA was probed as control. IGV plots of AbcR1/2 recovered in the elution (output) fractions are shown below. **(B)** Known NfeR1 target mRNAs co-purified efficiently with MS2NfeR1 in MM with glutamate. IGV plots show reads coverage and recovery profiles of *SMc03121* and *SMb20442* mRNAs upon affinity chromatography with wild-type and tagged NfeR1 as baits. The TSS of each mRNA is indicated with an arrow (+1).

blot probing, which confirmed that the tagged-sRNA was specifically retained by the MS2-MBP protein (Fig. 32A). Moreover, mapping of the sequencing reads from the eluted RNA demonstrated efficient recovery of MS2NfeR1 and co-purification of its target mRNAs, *SMc03121* and *SMb20442* (Fig. 32B). Patterns

of mapped reads on these mRNAs were similar in bacteria from MM and MM-NH_4 cultures (not shown).

2.7. Characterization of the NfeR1 targetome

To identify the NfeR1 mRNA partners, we normalized by coverage and compared read counts from NfeR1- and MS2NfeR1-binding transcripts mapping to different mRNA regions as described for AbcR1/2 in Chapter 1, i.e., full-length mRNA, virtual 5'-UTR, CDS and virtual 3'-UTR. In this case, we imposed a minimum of 100 mapped reads in either of these regions and a 2-fold difference between tagged-sRNA and control samples (log_2 FC >1) to score an mRNA as NfeR1 target. mRNAs likely captured by unspecific binding to the MS2 aptamer, as predicted by IntaRNA, were not considered further. This analysis unveiled an exceptionally large set of NfeR1-interacting mRNAs, i.e., 424 and 224 upon bacterial growth in MM and MM-NH_4, respectively, with only 14 target candidates (13 encoding proteins with unpredicted function) specifically recovered in stress of ammonia (Fig. 33; URL).

Putative NfeR1 target mRNAs are predominantly encoded in the *S. meliloti* chromosome (70%) (Fig. 33B). Of the total target candidates, 67% encode proteins with predicted functions and of those, 80% are likely involved in transport, metabolism or transcriptional regulation. Functional profiles of interacting mRNAs were equivalent in MM and MM-NH_4 (Fig. 33C).

We also examined reads coverage biases toward either region of the captured mRNAs to infer the most probable NfeR1 interaction sites at the targets. mRNAs with even reads distribution or similar enrichment in more than one region were regarded as unclassified (NC). This analysis revealed that NfeR1 preferentially binds to the 5'-UTR of its mRNA targets but targeting within the CDS is also appreciable (Fig. 34A; left bar graph). However, NfeR1 interactions at the 3'-UTR of mRNAs were rarely detected. Regardless of the mRNA interaction region

reported by MAPS, IntaRNA predicted NfeR1 base-pairing to ~80% of all target mRNA candidates, with ~60% of these antisense interactions likely favored thermodynamically, i.e., predicted hybridization energy (*E*)<-8 kcal/mol and involvement of nucleotide seeds of seven or more residues (Fig. 34A; right bar graph).

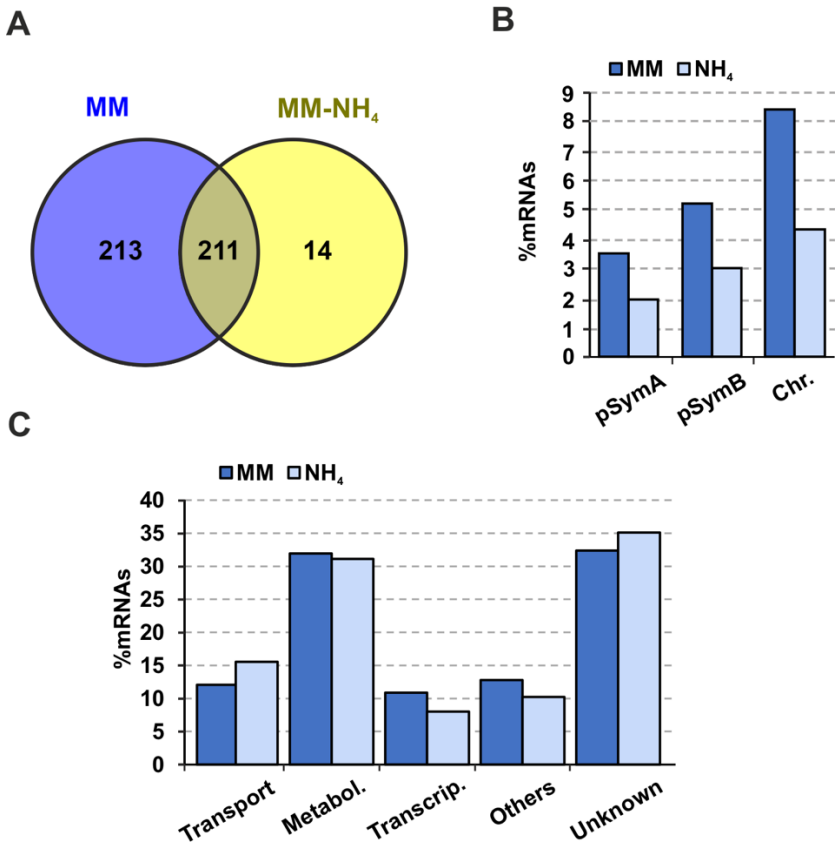


Figure 33. Overview of the NfeR1 mRNA interactomes determined by MAPS. (A) Venn diagram comparing the mRNA populations co-purified with NfeR1 in MM and MM-NH_4. **(B)** Distribution of captured mRNAs relative to the total number of protein-coding genes in each of the three *S. meliloti* replicons: the pSymA and pSymB megaplasmids, and the chromosome. **(C)** Functional categories of the co-purified mRNAs.

On the other hand, we also analyzed the population of sRNAs captured in the affinity chromatography, noticing a considerable number of putative NfeR1 targets within this group of transcripts (i.e., 72), which was similarly represented by sense RNAs, asRNAs and *trans*-sRNAs preferentially encoded in the

chromosome (i.e., 46). Further, IntaRNA predicted antisense interactions with NfeR1 via the aSD motifs for ~40% of these transcripts.

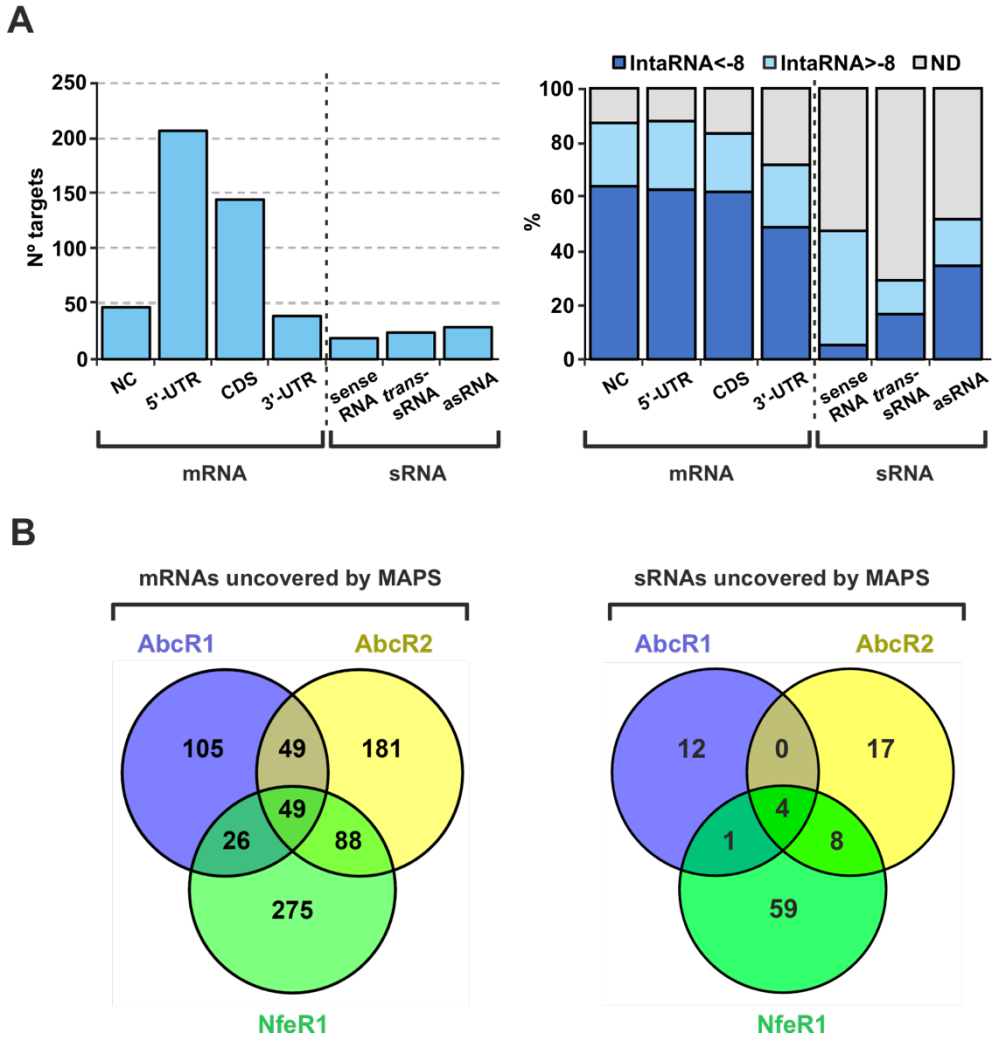


Figure 34. Recovery of mRNAs and sRNAs by NfeR1 MAPS. (A) Left, bar graph ploting the number of mRNAs and sRNAs scored as NfeR1 partners. Number of mRNAs enriched at the 5'-region, CDS and 3'-region with respect to the control experiments with the wild-type sRNA is specified. NC stands for mRNAs equally enriched at two or more regions. The different groups of sRNAs co-purified with MS2NfeR1 i.e., sense-RNA, *trans*-RNAs and asRNAs, are shown on the right. The *S. meliloti* genome was interrogated with IntaRNA for thermodynamically favored antisense interactions (minimum 7-nt seed) of NfeR1 in each enriched region of mRNAs and each full-length sRNAs. Data are ploted to the right. ND indicates no IntaRNA predictions **(B)** Venn diagram comparing mRNAs (left) and sRNAs (right) uncovered by MAPS for the AbcR!/2 and NfeR1.

AbcR1/2 and NfeR1 use similar aSD motifs for regulation, and therefore, the comparison of the three targetomes unveiled a large set of common mRNA targets, as expected (Fig. 34B). Remarkably, 36% and 63% of NfeR1 targets that encode metabolic and transport proteins, respectively, are shared with AbcR1/2-interacting mRNAs. Conversely, the sRNAs that likely interact with AbcR1/2 and NfeR1 are scarce. These comparisons revealed 275 and 59 NfeR1-specific potential target mRNAs and sRNAs, respectively (64% and 82%).

These data suggest that NfeR1 mostly regulates functions specified by the core genome through canonical translation inhibition of multiple mRNA targets. Furthermore, the large pool of sRNAs identified by MAPS anticipates additional levels of regulation of the NfeR1 network.

2.8. NfeR1 broadly regulates *S. meliloti* physiology

A relevant group of NfeR1-specific mRNA targets are related with DNA metabolism and cell division (Table 5). This set includes mRNAs encoding the cell cycle regulators DnaA and DivK, and MinC that represses cell division by inhibiting the Z-ring formation [196–199]. Another group of genes within the specific NfeR1 targetome are involved in motility and chemotaxis, i.e., those encoding the carboxylate chemoreceptor McpT, proteins for flagella biosynthesis, and the transcriptional regulators VisN and Rem [200–202]. However, motility assays did not evidence an altered competence for swarming and swimming of the *nfeR1* deletion mutant (Fig. 35).

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. Therefore, we also examined the absorption and colonization phenotypes of the Sm*ΔnfeR1* mutant on alfalfa roots. Plate counting of bacteria released from roots 2, 24, 48 and 72 h upon plants inoculation revealed that the wild-type and mutant strains were equally proficient in attaching to and prolifera-

Table 5. Some mRNAs specifically recovered as NfeR1-specific partners.

DNA replication and cell cycle	
ftsZ2	Cell division protein
dnaA	Chromosomal replication initiator
ftsY	Cell division protein
minC	Cell division inhibitor protein
smc	Chromosome partition protein
dnaG	DNA primase
recN	DNA repair protein
gyrB	DNA gyrase subunit B
SMc05018	DNA or RNA helicase
Motility	
rem	Exponential growth motility regulator
visN	Regulator of motility genes, LuxR family
flaF	Flagellin synthesis regulator
fliM	Flagellar motor switch transmembrane protein
flgL	Flagellar hook-associated protein
fliK	Hook length control protein
Microaerobic denitrification	
nosL	Copper chaperone
napF	Component of periplasmic nitrate reductase
napE	Component of periplasmic nitrate reductase
nosZ	Nitrous oxide reductase
Nitrogen assimilation (NSR)	
ntrB	Nitrogen assimilation regulatory protein
glnB	Nitrogen regulatory protein PII
glnE	Glutamate-ammonia-ligase adenylyltransferase
gdhA	Glutamate dehydrogenase
Nitrogen fixation	
fixI1	ATPase
fixL	Oxygen-regulated histidine kinase
fixP1	FixP1 di-heme c-type cytochrome

ting on the root surface (Fig. 35). This finding suggests that AbcR1 and NfeR1 have a different impact in *S. meliloti* physiology, despite of their large set of common target mRNAs involved in nutrient uptake and catabolism.

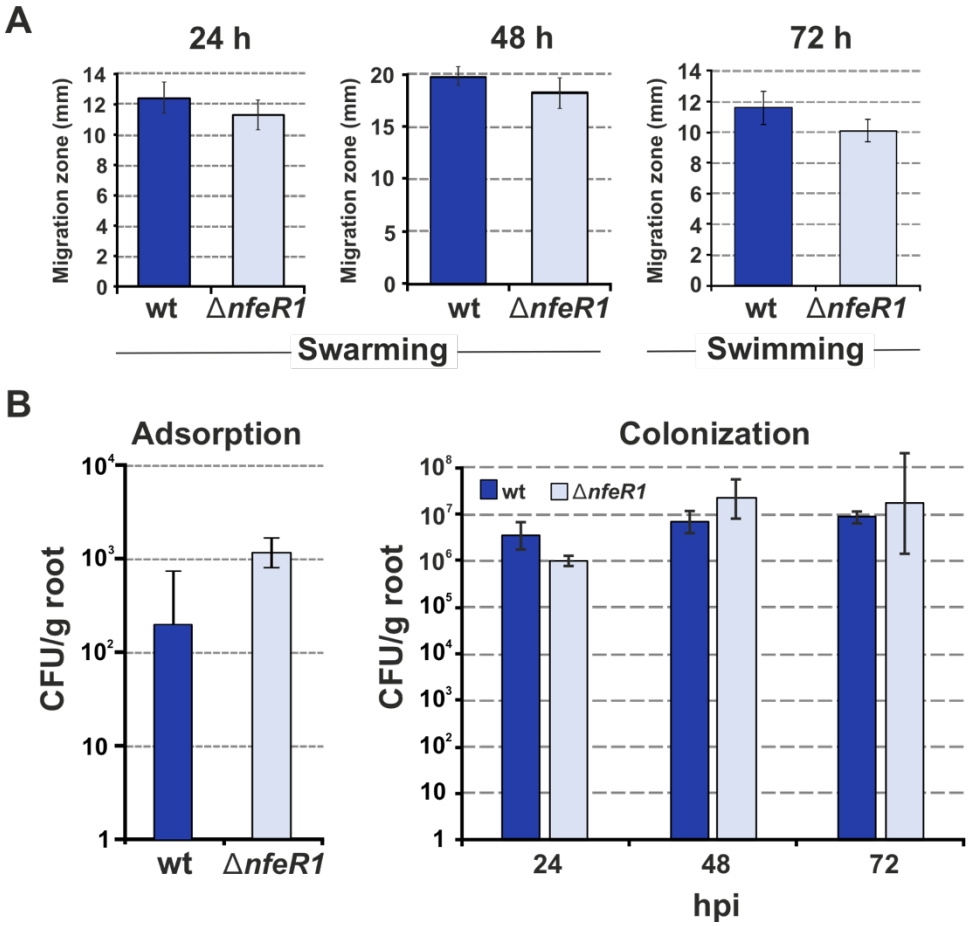


Figure 35. Motility and colonization phenotype of the SmΔ*nfeR1* mutant. (A) Swarming and swimming motility. Sizes of wild-type and SmΔ*nfeR1* colonies in swarming and swimming plates. Plotted values are means and SD of migration zone measurements on 20 (swarming) and 10 (swimming) colonies per strain. **(B)** Adsorption to and colonization of alfalfa roots by the wild-type strain and the SmΔ*nfeR1* mutant. Values reported are means and SD of the bacteria (CFU) released per g of root. Counting was performed on series of 30 plants inoculated with each strain 2 (adsorption), 24, 48 and 72 h post inoculation (hpi) of plants. The experiment was repeated twice with similar results.

2.9. Fine-tunning of N assimilation by NfeR1

The precise boundaries of three of the sRNAs captured as NfeR1-specific targets (SmelC549, Smel063, and SmelBR038) were determined previously by deep sequencing of the subpopulation of *S. meliloti* sRNA species [45] (Fig. 36). dRNA-Seq independently confirmed the TSS of the three transcripts [46]. SmelC549 and Smel063 are 250-nt and 72-nt long *trans*-encoded sRNAs, whereas SmelBR038 is a 79-nt sense sRNA likely derived from transcription within the *SMb21100* gene, which encodes a probable polysaccharide deacetylase. IntaRNA predicted favorable antisense interactions (E<-8 kcal/mol) of all three sRNAs with NfeR1, which explains their recovery in the MAPS assay. We used the SmelC549, Smel063 and SmelBR038 full-length sequences to interrogate the *S. meliloti* genome with IntaRNA, which rendered large lists of target mRNAs. Common predicted targets included the transcriptional regulators of nodulation (*nod*) genes *nodD1*, *nodD2* and *nodD3* (E<-9). Remarkably, we also found *ntrB* and *ntrC* among the common targets of the three sRNAs. IntaRNA further predicted that SmelC549, Smel063 and SmelBR038 target the CDS of these two mRNAs by favorable antisense interactions (E<-10 kcal/mol) involving seeds of at least 7-nt. These findings indicate that NfeR1 might silence sRNAs that are predicted to negatively regulate the NSR.

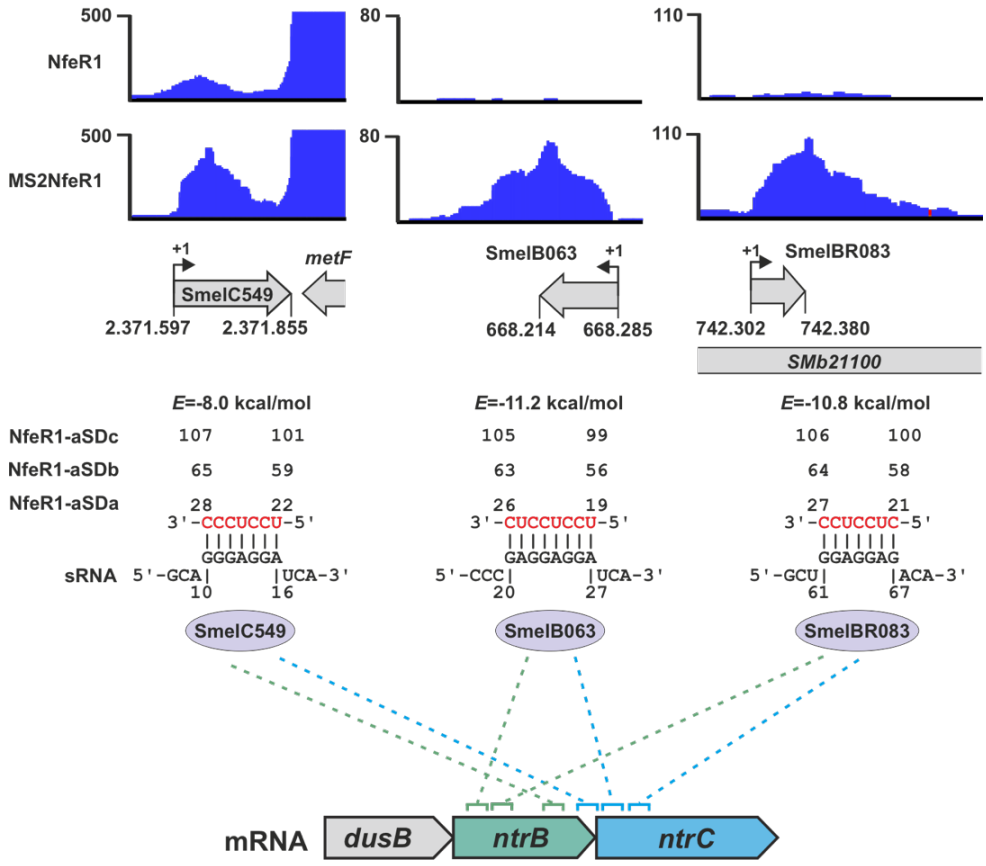


Figure 36. NfeR1 binds to three sRNAs that might silence *ntrBC*. IGV plots showing reads coverage and recovery profiles of SmelC549, SmelB063 and SmelBR038 sRNAs upon affinity chromatography with wild-type and tagged NfeR1 as baits. The TSS of each mRNA is indicated (+1). Numbers in the diagrams stand for coordinates within the Sm1021 genome. IntaRNA predicted base-pairing interactions between NfeR1 aSDa-c (in red), and the three sRNAs are shown below. Numbers denote nucleotide positions relative to the TSS of sRNAs. The predicted minimum hybridization energy (*E*) between NfeR1-aSDa and the mRNAs are indicated. *E* values for interactions involving aSDb and aSDc are similar. The predicted interaction sites (*E*<-8, minimum 7-nt seed) of each NfeR1-target sRNA at *ntrB* and *ntrC* are marked in green or blue, respectively.

MAPS also revealed putative interactions of NfeR1 with mRNAs involved in different pathways of the N metabolism, i.e., microaerobic denitrification, N fixation and N assimilation (Table 5). Of those, the NSR genes *ntrB, glnB, glnE* and *gdhA* do not belong to the AbcR1/2 regulons and, therefore are likely specific targets of NfeR1 regulation. Together with NtrB, GlnB and GlnE sense the intracellular N status whereas GdhA is an assimilatory glutamate dehydrogenase

encoded by a single copy gene within the *S. meliloti* pSymA symbiotic megaplasmid. MAPS profiles suggest that NfeR1 might interact at the 5'-region and deep into the CDS of *gdhA* (Fig. 37A). IntaRNA indeed predicted a plausible NfeR1-*gdhA* base-pairing at the mRNA CDS (*E*<-11.4 kcal/mol). Interestingly, RT-qPCR revealed upregulation of *gdhA* in the mutant Sm*ΔnfeR1* indicating that NfeR1 contributes to *gdhA* silencing (Fig. 37B).

On the other hand, we observed sequencing coverage biases toward the *glnE* CDS, whilst most of the recovered reads corresponding to *glnB* and *ntrB* mapped to the 5'-UTRs (Fig. 37C). Of note, IntaRNA predicted a thermodynamically favored antisense interaction between eight nucleotides within the translation initiation region of *ntrB* and either of the three NfeR1 aSD motifs (Fig. 37D). Data shown above suggest that NfeR1 is required for wild-type expression of *ntrB*. Therefore, we investigated the impact of NfeR1 on NtrB translation by the reporter assay. For that, the *ntrB::eGFP* reporter was mobilized to the Sm2020 strain harboring pSKiNfeR1 or pSKiNfeR1abc. Fluorescence of this translational fusion decreased upon IPTG-induction of NfeR1, whilst the simultaneous substitution of the three aSD motifs (NfeR1abc variant) abrogated *ntrB* regulation (Fig. 37E), thus validating *ntrB* as a target negatively regulated by NfeR1.

Taken together, these findings suggest that the plausible NfeR1-mediated depletion of the bicistronic *ntrBC* mRNA and the SmelC549, SmelB063 and SmelBR038 sRNAs, would counteract (auto)repression of the *dusBntrBC* operon, thus strengthening the *S. meliloti* NSR.

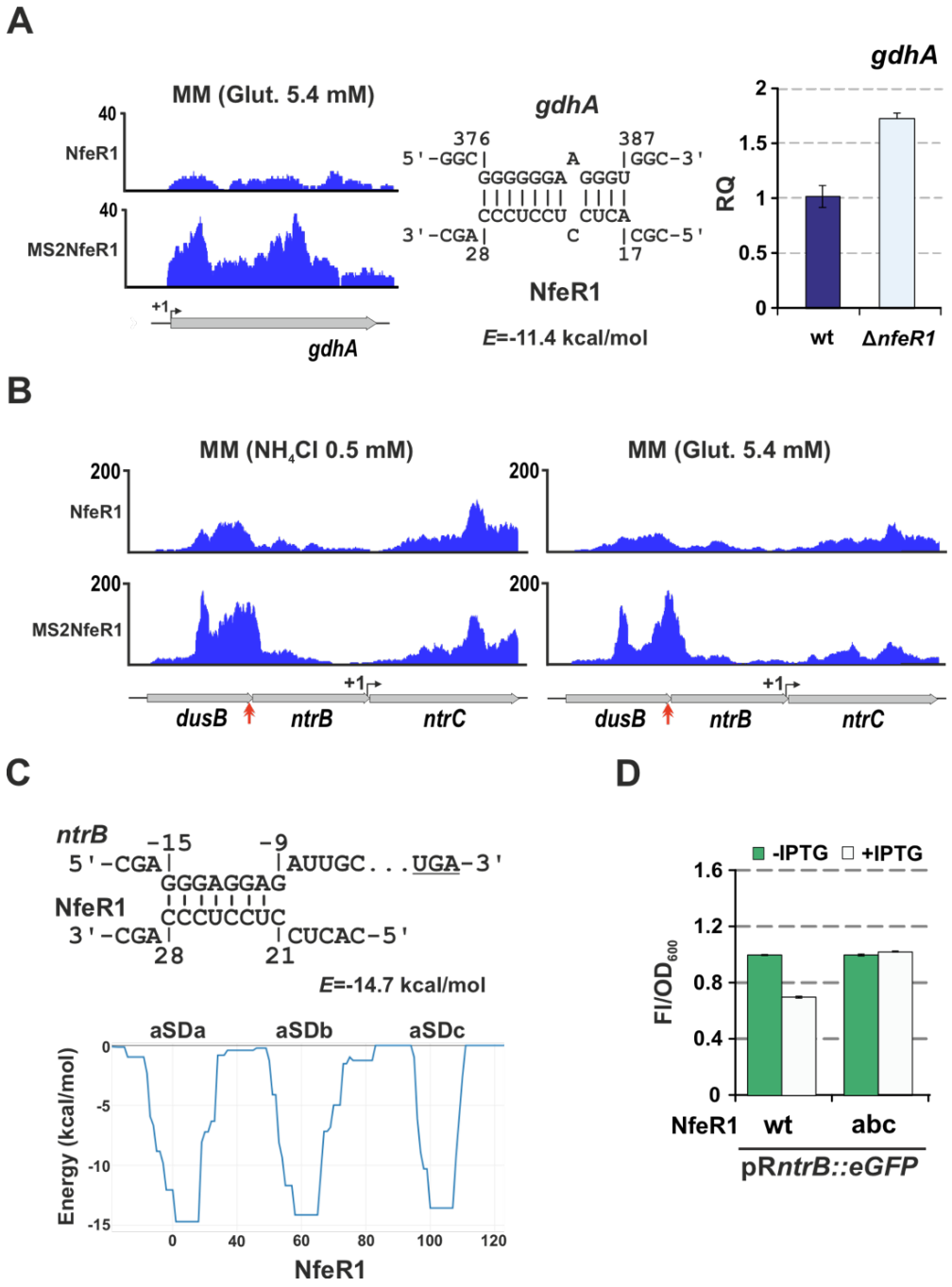


Figure 37. NfeR1 silencing of *gdhA* and *ntrB* mRNAs. Figure legend on the next page.

Figure 37. NfeR1 silencing of *gdhA* and *ntrB* mRNAs. (A) NfeR1 regulation of *gdhA*. Left, IGV plots showing recovery profiles of *gdhA* upon affinity chromatography with wild-type and tagged NfeR1 as baits. The TSS of *gdhA* is indicated (+1). IntaRNA predicted base-pairing interactions between NfeR1 aSDa-c and the *gdhA* CDS. Right, RT-qPCR analysis of *gdhA* abundance in wild-type Sm2B3001 and Sm*ΔnfeR1*. RNA was extracted from bacteria grown in TY to exponential phase, washed in PBS and cultured in MM supplemented with ammonia 0.5 mM for 4 h. Relative Quantification (RQ) values were normalized to *SMc01852* as a constitutive control. Values plotted in the bar graphs are means and SE of three replicates of two independent cultures. **(B)** IGV plots showing reads coverage of the *dusBntrBC* operon upon MS2-affinity chromatography. Red double arrowheads indicate the position of the predicted NfeR1 interaction site. **(C)** IntaRNA predicted base-pairing interactions between NfeR1 aSDa-c and the *ntrB* mRNA (denoted by red double arrowheads in B). The start codon of *ntrB* is underlined. IntaRNA predicted similar *E* for hybridization with the three aSD motifs (botton graph) **(D)** NfeR1 downregulates *ntrB*. Fluorescence of reporter strains co-transformed with the *ntrB::eGFP* translational fusions and plasmids overexpressing the wild-type NfeR1 or its mutant variant NfeR1abc upon IPTG-induction. Values plotted in the histogram correspond to the means and SD of 27 fluorescence measurements normalized to the OD_{600} of the cultures, i.e. three determinations of three independent cultures from three independent double transconjugants for each reporter strain. In the IntaRNA output diagrams, numbers denote nucleotide positions relative to the start codon of the *gdhA* and *ntrB* mRNAs and the NfeR1 TSS. The predicted minimum hybridization energy (*E*) between NfeR1-aSDa and the mRNAs are indicated. Energy values for interactions involving aSDb and aSDc are similar.

3. Discussion

NfeR1 is a stress-induced regulatory *trans*-sRNA that influences both the free-living and symbiotic *S. meliloti* lifestyles. Data presented in this Chapter show that NfeR1 transcription is controlled by a dual-mode promoter, activated by LsrB and repressed by NtrC, which specifies upregulation of the sRNA under N stress. Genome-wide deciphering of its RNA interactome anticipates a broad impact of NfeR1 in the control of *S. meliloti* physiology, metabolism, and N signalling at different levels during the symbiotic transition. Specifically, NfeR1 strengthens the *S. meliloti* NSR most likely by antisense post-transcritional silencing of the NtrBC two-component system repressors (Fig. 38).

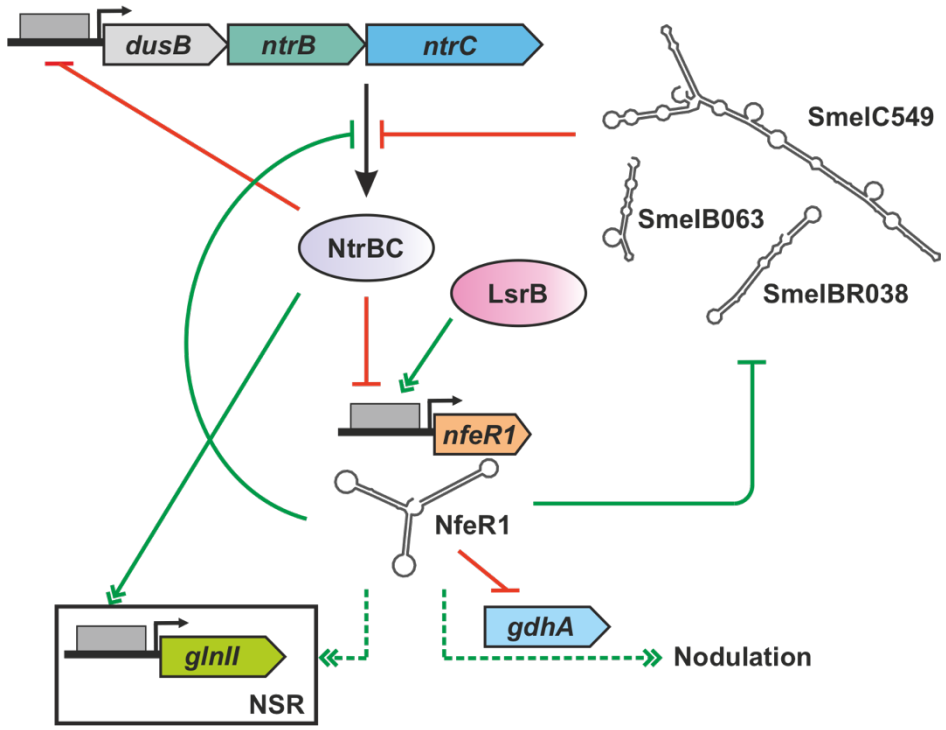


Figure 38. NfeR1 regulation of *S. meliloti* N metabolism: a proposed model. Solid and dashed lines indicate direct and indirect regulation, respectively. Truncated solid lines stand for negative regulation, whereas the double arrowheads indicate positive regulation. Negative regulation denoted by truncated solid lines in green likely results into a positive output from the regulatory loops. Details in the text.

3.1. NfeR1 is transcribed from a complex promoter

Transcription factors integrate specific external stimuli into bacterial promoters. Therefore, the function of a gene in environmental adaptation is linked to its transcriptional regulation. *In silico*, *in vitro* and *in vivo* approaches collectively demonstrated that the transcription factors LsrB and NtrC bind to and regulate the NfeR1 promoter. *S. meliloti* LsrB is required for alfalfa nodulation and N-fixation [151, 153], whereas rhizobial *ntrC* mutants are not affected in their ability to interact with their plant hosts [203]. However, NtrC mostly acts in its phosphorylated form, NtrC-P, as activator of σ^{54}-dependent promoters of N assimilation genes, and therefore, it is essential for bacterial free-living growth

under N limitation [168, 194]. Consistently with the functions assigned to these transcriptional regulators, NfeR1 has been previously shown to influence symbiosis [97], whilst data presented here uncovered an unprecedented role of this sRNA in the fine-tuning of the NSR in free-living *S. meliloti* bacteria. Remarkably, intracellular NfeR1 levels are the outcome of LsrB acting as indispensable seemingly constitutive activator, and NtrC as N-dependent repressor of NfeR1 transcription. Complex promoters such as that of NfeR1 occur frequently in bacteria, ensuring the integration of different environmental signals to achieve the accurate expression of genuine adaptive genes [204]. Indeed, other sRNAs have been shown to be regulated by more than one transcription factor. This is the case of *S. meliloti* MmgR, which is NtrC-activated and repressed by the global regulator of the C flux, AniA [95].

NtrC activity as transcriptional repressor is still poorly explored. Besides *trans*-repression of *gdhA* in *Pseudomonas putida* [193], genetic evidence suggests that NtrC negatively autoregulates *ntrBC* in several bacterial species, including *S. meliloti* [184, 192, 195, 205]. However, the mechanisms of NtrC-mediated repression of this bacterial operon have not been thoroughly investigated. Pulldown assays and expression profiling evidenced that NtrC-binding to the NfeR1 promoter and transcriptional repression preferentially occur under N surplus in complete or defined media, in which NtrC is mostly dephosphorylated. It is worthy to note that a mutant lacking the histidine-kinase NtrB did not show a growth defect in N-limiting media as obvious as that of the *ntrC* mutant, suggesting that the function of this long considered inactive form of NtrC is underestimated in *S. meliloti*, even in the frame of the NSR.

Most bacterial transcription factors that function as repressors bind to DNA blocking or occluding access to RNA Polymerase (RNAP), which does not seem to be the case of NtrC at the NfeR1 promoter. Alternatively, the regulatory outcome from complex promoters can be shaped by direct interactions between

repressor and activator, e.g., the LacI-type CytR repressor and the activator cyclic AMP receptor protein CRP at the *E. coli cdd* promoter [204]. The LsrB and NtrC-binding sites do not overlap but rather map 20-nt apart within the NfeR1 promoter. However, it cannot be ruled out that both proteins hinder efficient DNA-binding to each other. In this scenario, the effect of post-translational LsrB and/or NtrC modifications must be also considered. LsrB Cys residues that likely sense oxidative stress via the formation of intermolecular disulfide bonds to generate LsrB dimers, are required for expression of some target genes. In fact, LsrB Cys-to-Ser substitutions alter *S. meliloti* oxidative stress adaptation, nodulation, and N fixation [151]. On the other hand, our *in vitro* data suggest that NtrC binding to the NfeR1 promoter is independent of the phosphorylation status of the protein. Indeed, the conserved NtrC-binding motif of the repressed NfeR1 promoter (GC-N_{11}-GCA) is shared with the NtrC-activated *glnII* promoter in *B. japonicum, R. leguminosarum, R. etli* and *S. meliloti* [194, 206], suggesting that the σ^{54} site, rather than the protein affinity for the promoter, is a major determinant of NtrC-dependent transcriptional activation. Even though to a much lesser extent than in N excess media, our data also evidenced NtrC competence to repress the NfeR1 promoter under N starvation (e.g., in MM broth). Similarly, it has been reported that the *S. meliloti ntrBC* operon and *P. putida gdhA* are actively repressed by NtrC-P upon activation of the NSR [184, 193], further supporting that both forms of this transcription factor can bind to and repress simple promoters. In conditions of N stress (i.e., maximal levels of NfeR1), *ntrC* transcription is markedly upregulated, whereas accumulation of the *lsrB* mRNA remains invariable with respect to N excess (not shown). Therefore, the post-translational modifications rather than the cellular levels of both proteins are likely the major determinants of the transcriptional output from the complex NfeR1 promoter. Environmental stress, including N starvation, promotes endogenous ROS accumulation and LsrB dimerization. Therefore, LsrB dimers might efficiently outcompete NtrC-P at the

NfeR1 promoter during the NSR. Conversely, under N surplus, NtrC repressor activity would overcome activation by the likely prevalent monomeric LsrB, which nonetheless is evident in this condition at single-mode promoters such as that of AbcR1. This would also explain the recovery profiles of LsrB and NtrC in the pulldown assays. Within the nodules, *ntrBC* transcription and the NSR are silenced but LsrB dimerizes because of plant-derived ROS, which likely promotes the high LsrB-dependent accumulation of NfeR1 within the endosymbiotic compartments.

3.2. Redundant function in regulation of the NfeR1 aSD motifs

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 [97]. The double-plasmid reporter assay confirmed downregulation of *SMc03121* and *SMb20442* transport mRNAs by NfeR1. Remarkably, NfeR1 mutagenesis unveiled that the three single-stranded aSD sequences, strategically located within the predicted stem-loops of the transcript, are 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 [92, 160, 207–209]. In the latter case, differences in sequence 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. Of note, these analyses rendered the triple mutant NfeR1abc as a valuable negative control for the experimental validation of direct NfeR1 targets. Thus, 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, making NfeR1-mediated translational inhibition more feasible and independent of Hfq.

3.3. Overall insights into NfeR1 activity mechanism and function derived from MAPS

To further explore its function in the NSR, we determined the NfeR1 RNA interactome by MAPS upon *S. meliloti* growth under N stress. This approach delivered a large list of candidate target mRNAs (7% of protein-coding genes in the *S. meliloti* genome) mostly encoding proteins related to nutrient uptake and metabolism, many of which identified also as targets of AbcR1/2 regulation. This was an expected finding since all three sRNAs use aSD motifs of similar nucleotide sequence for targeting. Unlike the mRNAs of the AbcR1/2 regulons, NfeR1 targets are more represented in the *S. meliloti* chromosome, thus hinting at a major impact in regulation of core adaptive functions. MAPS profiles of the NfeR1 target mRNAs suggest a major canonical regulatory mechanism relying on base-pairing at the RBS leading to translation inhibition and accelerated mRNA decay. However, the set of targets recovered on their CDS were not minority, far outnumbering the AbcR1/2 interacting mRNAs with this read distribution pattern. Interactions within the CDS of the mRNAs have been reported to differently influence stability of targets regulated by diverse sRNAs in *Agrobacterium* and *E. coli* [92, 208, 210–212], and merits further investigation as alternative NfeR1 regulatory mechanism.

3.4. MAPS anticipates a broad impact of NfeR1 in *S. meliloti* physiology

Despite of the evident commonalities with the AbcR1/2 regulon and activity mechanism, MAPS uncovered relevant functional specificities for NfeR1. In this regard, it was noticeable the capture of a set of mRNAs encoding proteins

involved in chemotaxis, regulation and assembly of the flagellar apparatus, response to the osmotic stress, and cell cycle control. Flagella and chemotaxis mRNAs specifically targeted by NfeR1 contribute to swarming and/or swimming motility, which help rhizobia colonize rhizosphere, root and even nodules intracellularly (e.g., *R. leguminosarum*) [175, 199, 200, 213, 214]. *S. meliloti* bacteria lacking NfeR1 were not apparently defective either in motility or alfalfa root colonization. This latter phenotype further points to a NfeR1 function in regulation of *S. meliloti* metabolism differing from that proposed for AbcR1/2. However, gene knock-out usually fails to unveil sRNA function when this is assessed by such a complex end-point phenotypes that, in many bacteria, rely on diverse and redundant pathways [73, 215]. Therefore, the consequences of NfeR1 regulation of these targets in free-living and symbiotic bacteria must be investigated further.

We previously reported that NfeR1 contributes to salt/osmotic stress adaptation in *S. meliloti* [97]. Interestingly, some of the mRNAs catalogued here as NfeR1 targets are known to be misregulated in hypersaline environments, i.e., the *agl* operon for trehalose uptake, which is accumulated in osmostressed rhizobia and *greA,* encoding a transcription cleavage factor required for salt tolerance and symbiotic efficiency [216–218]. Besides metabolic reprogramming, adjustment of cell cycle progression is crucial for bacterial survival and adaptation under harsh environmental conditions [78]. Therefore, NfeR1 regulation of cell cycle related genes might further help *S. meliloti* survive abiotic stress in soil. Several lines of evidence indicate that modulation of the *S. meliloti* cell cycle is also critically important for bacteroid differentiation during symbiosis [199]. One of the putative NfeR1 targets is the mRNA encoding DivJ, a cell cycle regulator that controls CtrA activity and is required to establish an effective symbiosis [196]. Besides, misregulation of the NfeR1 target candidates *ftsZ*, *dnaA and minC* alters cell morphology [198, 219, 220]. It is known that some NCR peptides that

promote rhizobial endoreduplication and terminal bacteroid differentiation within nodules also influence *ftsZ* and *dnaA* expression [77, 221]. Profiling of RNA obtained by laser microdissection of all developmental zones of indeterminate alfalfa nodules revealed a particular NfeR1 abundance in the so-called interzone II-III, where rhizobia arrest division and start differentiation [54]. Therefore, these findings predict a novel role of NfeR1 in bacteroid differentiation.

3.5. NfeR1 influences symbiotic N signaling and the NSR

The N status of both bacteria and plant is a major metabolic signal in legume nodulation and symbiotic N fixation [168]. Interestingly, MAPS also placed NfeR1 as regulator of several pathways of N metabolism, i.e., N uptake, denitrification, N fixation, and aerobic N assimilation, which further explains the broad impact of NfeR1 in symbiosis. All transport and metabolic genes for ammonia assimilation are switched off in mature N-fixing bacteroids to guarantee the transfer of the fixed N to the plant [168, 222]. Complete reduction of nitrate or nitrite to N_2 by microerobic denitrification within nodules may counteract N fixation. Further, one of the gaseous intermediates of this process, nitric oxide, is toxic and harmful for nitrogenase [223–225]. It is therefore conceivable that NfeR1 contribution to the fine-tuning of mRNAs for the uptake of combined N (e.g., amino acids) and denitrification (i.e., *nap*, *nos*) might enhance the *S. meliloti* overall symbiotic efficiency.

The NtrBC two-component system is at the core of regulation of N assimilation in free-living rhizobia, which is fully active at the onset of nodulation, i.e., nodule formation is totally inhibited by N excess in soil [168]. Among many other genes involved in assimilation of combined N, *S. meliloti* NtrC has been shown to activate N fixation genes (*nif* and *fix*) in cultured rhizobia under N-limiting conditions [203, 226]. Post-transcriptional silencing of certain *nif/fix* genes likely regulated by NfeR1 could optimize their mRNA levels within

nodules and prevent the energy costs derived from translation when their coding proteins are fully dispensable in soil.

In rhizobia, ammonia assimilation proceeds mainly through the glutamine synthetase (GS)-glutamate synthase (GOGAT) pathway controlled by NtrBC. The glutamate dehydrogenase (GDH) encoded by *S. meliloti gdhA* in pSymA specifies a second but minor pathway for ammonia assimilation that may operate in both free-living and symbiotic conditions [222, 227–229]. GDH is not encoded by all rhizobial genomes and, therefore it has been likely acquired by *S. meliloti* with the accessory genome [30, 230]. Of note, a *S. meliloti* mutant with increased GDH activity has been shown to be defective in nodulation [229]. Similarly, in *R. etli*, which does not encode GDH, ectopic overexpression of *E. coli gdhA* strongly downregulates *nod* gene expression in a NtrC-dependent manner (i.e., a *ntrC* knock-out does not affect nodulation regardless of GDH levels), and compromises symbiotic N fixation by bacteroids. The high intracellular amino-N pool because of misregulated/constitutive GDH activity prior to or during nodule formation likely underlies these symbiotic defects [227, 228]. Our data showed that NfeR1 silences *S. meliloti gdhA* under N stress most likely by favored base-pairing within its CDS. *S. meliloti* NtrBC sense the N status and further adjusts the expression of *nod* genes accordingly. In fact, it has been reported that *S. meliloti* NtrC activates *nod* genes transcription via the flavonoid-independent NodD3 regulator at low ammonia concentrations [231–234]. Interestingly, MAPS also identified a particularly large set of sRNAs as probable NfeR1 partners. Of those, SmelC549, SmelB063 and SmelB038 are predicted to base-pair with the *nodD1*, *nodD2* and *nodD3*, and *ntrC* mRNAs. It is therefore plausible that NfeR1 optimizes nodulation by silencing of *gdhA*, and these *nodD* and *ntrC* ribo-repressors. Nonetheless, the major role of the NtrBC system is regulation of the NSR. Consistently with its accumulation under N stress, we found that NfeR1 promotes upregulation of *ntrBC* and *glnII*, thus indicating a positive contribution to the

NSR. Two-component systems are a dominant form of bacterial signal transduction commonly subject to feedback regulation to adjust their outputs [235]. This is also the case of *S. meliloti (dusB)ntrBC*, whose transcription has been shown to be repressed by NtrC. Our data predict that the NfeR1 sRNA partners SmelC549, SmelB063 and SmelB038 might be also unexpected new elements of NtrBC feedback regulation. Remarkably, affinity chromatography captured also *ntrB* as plausible NfeR1 target mRNA. NfeR1 overexpression promotes downregulation of NtrB as revealed by a genetic assay in which transcription of both sRNA and target mRNA is unlinked from their endogenous regulation. As part of a polycistronic mRNA, NfeR1-mediated silencing of *ntrB* might also promote decay of *ntrC* and/or alter the phosphorylation status of the operon autorepressor NtrC. In the context of the N-dependent endogenous regulation of NtrBC, NfeR1 would alleviate the autorepression of the operon to speed up and strengthen the NSR.

In summary, we have described unprecedented RNA elements that together with the NtrBC system are arranged into two complex negative feedback loops sharing NfeR1, which respond to N availability to pervasively shape the *S. meliloti* symbiotic behavior.

4. Experimental setup.

Bacterial strains and growth conditions. Bacterial strains and plasmids specifically used in this Chapter, along with their relevant characteristics, are listed in Table 6. To test the effect of shifts in N metabolism on NfeR1 accumulation, the L-glutamate (6.5 mM) of the standard MM was replaced by NH_4Cl (10 or 0.5 mM) or KNO_3 (10 or 0.5 mM).

Table 6. Bacterial strains and plasmids used in Chapter 2.

Strain/Plasmid	Relevant characteristics	Reference/Source
STRAIN		

S. meliloti		
Sm*Δlsr*B	Sm2011 Δ*lsrB* derivative; Sm^r	Chapter 1
Sm*ΔntrC*	Sm2B3001 Δ*ntrC derivative*; Sm^r	This work
Sm*ΔntrB*	Sm2B3001 Δ*ntrB derivative*; Sm^r	This work
Sm*ΔnfeR1*	Sm2B3001 Δ*nfeR1 derivative*; Sm^r	[97]
PLASMIDS		
pK18*Δlsr*B	Suicide plasmid for *lsrB* deletion; Km^r	This work
pK18*ΔntrC*	Suicide plasmid for *ntrC* deletion; Km^r	This work
pK18*ΔntrB*	Suicide plasmid for *ntrB* deletion; Km^r	This work
pSRK-NfeR1	pSRK derivative constitutively expressing NfeR1	[119]
pSRKMS2NfeR1	pSRK derivative constitutively expressing MS2NfeR1	[119]
pSKiNfeR1	pSRKKm carrying the NfeR1 coding sequence fused to *sinR*-P_{sinI}	This work
pSKiMS2NfeR1	pSKMS2 derivative expressing MS2NfeR1	This work
pSKiNfeR1a/b/c	pSRKKm derivatives expressing NfeR1 mutants in loops a, b or c	This work
pSKiNfeR1ab	pSRKKm derivatives expressing NfeR1 mutants in loops a and b	This work
pSKiNfeR1abc	pSRKKm derivatives expressing NfeR1 mutants in loops a, b and c	This work
pBBNfeR1-40*::eGFP*	pBBR1MCS-2 derivative expressing a transcriptional fusion of a truncated *nfeR1* promoter (54-bp) to *eGFP*; Km^r	[97]
pBBNfeR1-100*::eGFP*	pBBR1MCS-2 derivative expressing a transcriptional fusion of a truncated *nfeR1* promoter (114 bp) to *eGFP*; Km^r	[97]
pBBNfeR1-213*::eGFP*	pBBR1MCS-2 derivative expressing a transcriptional fusion of a full-length *nfeR1* promoter (227-bp) to *eGFP p*; Km^r	This work
pBBNfeR1-100*::eGFP*	pBBR1MCS-2 derivative expressing a transcriptional fusion of a *nfeR1* promoter (114-bp) including mutations to *eGFP*; Km^r	This work
pBBP$_{glnII}$*::eGFP*	pBBR1MCS-2 derivative expressing a transcriptional fusion of a full-length *glnII* promoter (400-bp) to *egfp*; Km^r	This work
pABCaNfeR1-40	pABCa::*GFP* derivate expressing a transcriptional fusion of a truncated *nfeR1* promoter (54-bp) to *egfp*; Gm^r	This work
pABCaNfeR1-100	pABCa::*GFP* derivate expressing a transcriptional fusion of a truncated *nfeR1* promoter (114 bp) to *egfp*; Gm^r	This work
pABCaNfeR1-213	pABCa::*GFP* derivate expressing a transcriptional fusion of a full-length *nfeR1* promoter (227-bp) to *egfp*; Gm^r	This work

pR*SMc03121::eGFP*	*SMc03121::eGFP* translational fusion (-156/+36 relative to *SMc03121* AUG) ; Ap^r, Tc^r	[97]
pR*SMb20442::eGFP*	*SMb20442::eGFP* translational fusion (-18/+45 relative to *SMb20442* AUG) ; Ap^r, Tc^r	[97]
pR*ntrB::eGFP*	*ntrB::eGFP* translational fusion (-143/+48 relative to *ntrB* AUG) ; Ap^r, Tc^r	This work
p16*lsrB*	pET16b derivate carrying N-terminally 10xHis-tagged LsrB	This work
p29*ntrC*	pET29a derivate carrying NtrC	This work

Oligonucleotides. Sequences of the oligonucleotides used as probes for Northern hybridization, or as amplification primers for cloning and RT-qPCR are listed in Table 7.

Table 7. Oligonucleotides specifically used in Chapter 2. Restriction sites are underlined**.**

Oligonucleotides	5'-sequence-3'	Use
PbNfeR1	TGCTTGATCTGATTGGCAACCGGGA	Northern blot probing
Pb5S	TACTCTCCCGCGTCTTAAGACGAA	
BamHIntrBFw	ATTAGGATTCCGGTCCATGGCGGACAAG	Construction and PCR verification of Sm*ΔntrC* mutant
ATGXbaIRv	CGAATCTAGACGCCCGTCATCCATTGGT	
XbaITGAFw	GTTGTCTAGAAGCGCTTGACTGAGATGC	
ntrYHindIIIRv	GCCGAAGCTTATTGCGCTCCAGGTCGTT	
ntrCoutFw	CGGCCGATGAAGAAACTC	
ntrCoutRv	TGTCATCAGCTCGACGAAG	
BamHIntrBupFw	CGTGGGATCCTGAAGAATTCCGGCATCG	Construction and PCR verification of Sm*ΔntrB* mutant
ntrBupXbaIRv	CTACTCTAGACTTCTCGGTCATGCCGCA	
XbaIntrBdownFw	CTTGTCTAGACCGGCATCCAAGGGGCTG	
ntrBdownHindIIIRv	CTTGAAGCTTGCGCATCCATCGGCATGT	
ntrBoutFw	TCGTGACCGAGATGGTG	
ntrBoutRv	TGCGCCCGTCATCCATT	

NfeR1OexfusTSSI	GAGCCTGACAGCATCGCTACATCG ATCGGGCAGCGCAC	Construction of pSKiNfeR1
Mut1Fw	GGGTCTGTCTAGCTCGATCGG	Construction of pSKiMS2NfeR1_1 and pSKiMS2NfeR1_3 combined with sinR_NdeIF (R) and SecSRK (F)
Mut1Rv	CCGATCGAGCTAGACAGACCC	
Mut3Fw	GGGTCTGTGTCGCTCGATCGG	
Mut3Rv	CCGATCGAGCGACACAGACCC	
Smr14C2a_F	GCAGCGCAAGAAGAAGAAAAGCG CGCCCGATACGGATTGG	Construction of pSKiNfeR1a/b/c combined with sinR_NdeIF (R) and SecSRK (F)
Smr14C2a_R	TTTTCTTCTTCTTGCGCTGCCCGAT CGATGGA	
Smr14C2b_F	TGGCGACAAGAAGAAGAAAGGTT GCCAATCAGATCAAGC	
Smr14C2b_R	TTTCTTCTTCTTGTCGCCAATCCGT ATCGGG	
Smr14C2c_F	CGCCCGCTGGAGGAGGAGGGCCA GCGGGCGTTGTTTATTTC	
Smr14C2c_R	CCCTCCTCCTCCAGCGGGCGTTGC TTGATC	
P14C2Fw	ACTAGTATTCTGTGATCATTCGGC GC	$P_{nfeR1\text{-}100}$ amplification
P14C2Rv	TCTAGAGCTGCCCGATCGATGATT GG	
P14C2_54	CTAGTGCCCCTGGTAAAATCCGGG GGTTCGGCCTATATTCCAATCATC GATCGGGCAGCT	Generation of $P_{abcR1\text{-}40}$ by annealing
P14C2_54i	CTAGAGCTGCCCGATCGATGATTG GAATATAGGCCGAACCCCCGGATT TTACCAGGGGCA	
P14C2EcoRIFw	CGTAGAATTCCGGTTGCCAATCGC CT	$P_{nfeR1\text{-}213}$ amplification
P14C2XbaIRv	TGAGTCTAGAGCTGCCCGATCGAT GA	
EcoRIPc14mutFw	AATTCATTCTGTGATCATTCGGCG CCTGAGCCAacgATCACTacgATAGG TGCCATTCGCGGCAGCCCCTGGTA AAATCCGGGGGTTCGGCCTATATT CCAATCATCGATCGGGCAGCC	Generation of $P_{nfeR1\text{-}100}$* by annealing
XhoIPc14mutRv	TCGAGGCTGCCCGATCGATGATTG GAATATAGGCCGAACCCCCGGATT TTACCAGGGGCTGCCGCGAATGGC ACCTATcgtAGTGATcgtTGGCTCAG GCGCCGAATGATCACAGAATG	
XbaImTSSglnIIR	GAAGTCTAGATGAGCAAATCCTGC CGG	P_{glnII} amplification
HindIIIPglnII-400F	GTTCAAGCTTATTGATGCAACGGC CGC	

avrIISRFw	CAATCCTAGGCACCGCGGGGAAGTACGCCA	Construction of pABCa::*GFP* transcriptional fusions
avrIIGFPRv	CGGCCCTAGGTTAGCAGCCGGATCCTTTGTATAG	
BtnPC14Fw	BiotinCGGTTGCCAATCGCCTTTATGACGCC	DNA-chromatography pull down assay and construction of $P_{nfeR1\Delta}$ probe
PC14FusFw	TTGCCCATTATTTCGGCGCCCCTGGTAAAATCCGGG	
PC14FusRv	GCCGAAATAATGGGCAAGATCGTTATACAAAATGCG	
P14C2Rv	TCTAGAGCTGCCCGATCGATGATTGG	
ntrB_Fw	ACCAGGATCCTATCTCGATCGTTTCGC	Amplification of the *ntrB* 5'-region fused to *eGFP*
ntrB_Rv	TAATGCTAGCGGAAAGATCGTTGGCACC	
LsrB_Fw_NdeI	ATATCATATGGGGGATTCTATGTCGCT	Amplification of the LsrB CDS
LsrB_Rv_BamHI	ATGGATCCTTATCAGAAGTTCCAGTTTCTCG	
NtrC_Fw_NdeI	ATATCATATGACGGGCGCAACGATCCT	Amplification of the NtrC CDS
NtrC_Rv_HindIII	ATAAGCTTATCAAGCGCTACGCGAGCTGC	
gdhA125F	ATCCCCGTTACGCCGAAAAC	qRT-PCR of *gdhA ntrBC glnII* and *lsrB*
gdhA239R	AACCCCCGGTTGATCTGGAC	
ntrB957F	CCCGTTCATCACCACCAAGA	
ntrB1074R	GAAGGTCGTGCGGCTATGCT	
ntrC172F	CCGGATGAAAACGCCTTCG	
ntrC290R	CCCTTCTCCGAAGCCTTGATG	
glnII500F	AGGCATCAACGCCGAAGTG	
glnII614R	GTTAGGCGCAGCAGAAGGTAG	
lsrB517F	GCGCCGTCCTACATCAACAA	
lsrB631R	GCCAGTTCACGTCGAGCAGA	
Smc01852F	TCACCAACACTGCCGACTGC	
Smc01852R	TCGTGTGCAGGATGCTGATG	

Generation of DNA probes for the pull down assay. The BtnPC14Fw/P14C2Rv primer pair was used for amplification of Btn-P_{nfeR1} from genomic DNA. The promoter fragment lacking the conserved motif was

generated by a two-step PCR on genomic DNA, the first amplification round with the primer pairs BtnPC14Fw/PC14FusRv and PC14FusFw/P14C2Rv to yield overlapping fragments flanking this motif, and the second with BtnPC14Fw/P14C2Rv to generate the full-lenght Btn-$P_{nfeR1\Delta}$ DNA probe. Both Btn-P_{nfeR1} and Btn-$P_{nfeR1\Delta}$ DNA fragments were concentrated and purified by phenol-chloroform extraction followed by ethanol precipitation. The DNA-chromatography pull down assays were performed following a described protocol in Material and Methods.

Construction of *S. meliloti* mutants. Knock-out mutants were generated using the suicide plasmid pK18*mobsacB* as described in Material and Methods. SmΔ*ntrC* and SmΔ*ntrB* were generated in Sm2B3001 by a markerless in-frame deletion of the *ntrC* and *ntrB* CDS using pK18Δ*ntrC* and pK18Δ*ntrB,* respectively. To construct pK18Δ*ntrC,* 792-nt and 749-nt DNA fragments flanking the *ntrC* ORF were amplified from genomic DNA with the BamHIntrBFw/ATGXbaIRv and XbaITGAFw/ntrYHindIIIRv primer pairs. PCR fragments were digested with *Bam*HI/*Xba*I and *Xba*I/*Hind*III, respectively, and ligated to the pK18*mobsacB Bam*HI and *Hind*III restriction sites, leading to insertion of the tandem fragments via their common *Xba*I site. Similarly, pK18Δ*ntrB* were generated by amplification of 816-nt and 801-nt DNA fragments from genomic DNA with the BamHIntrBupFw/ntrBupXbaIRv and XbaIntrBdownFw/ntrBdownHindIIIRv primer pairs and subsequent digestion *Bam*HI/*Xba*I and *Xba*I/*Hin*dIII and ligation between the pK18*mobsacB Bam*HI and *Hin*dIII restriction sites.

EMSA with LsrB and NtrC. The NtrC CDS was PCR amplified from genomic DNA using the primers NtrC_Fw_NdeI/NtrC_Rv_HindIII and cloned into the vector pET-29a (Novagen) between the *Nde*I/*Hin*dIII restriction sites, yielding p29NtrC for native NtrC overexpression and purification as described in Material and Methods. The P14C2Fw/P14C2Rv and

P14C2EcoRIFw/P14C2XbaIRv primer pairs were used to amplify $P_{nfeR1\text{-}100}$/$P_{nfeR1\text{-}100*}$ and $P_{nfeR1\text{-}213}$, respectively, using pBB-*eGFP* carrying promoter fragments of different lenght as DNA templates. PCR products were further purified with the GFX™ PCR DNA and Gel Band Purification Kit (GE Healthcare). A deleted version of the *nfeR1* promoter ($P_{nfeR1\text{-}40}$) was generated by annealing of oligonucleotides P14C2_54 and P14C2_54i. This product was purified from agarose gel. Binding reactions were performed with 1 nM radiolabeled probes in the absence or presence of purified LsrB or NtrC, which were then subjected to electrophoresis and analyzed with the Personal FX equipment and Quantity One software (Bio-Rad). For the footprint on the bottom strand, P14C2EcoRIFw/P14C2XbaIRv primer pair were used, but in this case, P14C2XbaIRv was end-labeled. Binding reactions were performed with 20 nM radiolabeled probes in the absence or presence of purified NtrC before DNAse I digestion. These protocols are described in detail in Material and Methods.

Construction of plasmids for induced NfeR1 expression and tagging. For the IPTG-induced expression of wild-type and MS2 aptamer-tagged NfeR1, we constructed plasmids pSKiNfeR1 and pSKiMS2NfeR1. NfeR1 was amplified from pSRKMS2NfeR1 (constitutively expressing MS2NfeR1) using the PCR1/PCR2 primers (Table 2). The PCR product was digested with *Xba*I and *Xho*I and inserted into pSKMS2 to generate pSKiMS2NfeR1. Alternatively, NfeR1 was amplified from pSRK-NfeR1 (constitutively expressing the wild-type transcript) using the NfeR1OexfusTSSI/secSRK primer pair [97]. The forward primer contains a complementary sequence to TSS3_28bp_b_sinIR, which is used together with the sinR_NdeIF primer for amplification of the *sinR*-P_{sinI}module. These first PCR products were used as template for a second PCR using the primer pairs sinR_NdeIF/HindIIIvec (Table 2). The resulting fragments were restricted with *Nde*I and *Xba*I and inserted into pSRKKm to generate pSKiNfeR1.

Replacements of specific nucleotides within aSDa-c were performed using a two-step PCR strategy based on overlapping fragments using pSKiNfeR1 or derivative plasmid constructs as the template. The first round of PCR amplifications was performed with sinR_NdeIF or secSRK (Table 2) (both hybridizing to all plasmid templates) and their respective primer pairs carrying the desired mutations (Smr14C2a_F, Smr14C2a_R, Smr14C2b_F, Smr14C2b_R, Smr14C2c_F; Table7). Each pair of complementary PCR products was used as the template in the second PCR with sinR_NdeIF/secSRK. The resulting products were digested with *Nde*I/*Xba*I and ligated to pSRKKm to yield plasmids pSKiNfeR1a, pSKiNfeR1b, pSKiNfeR1c, pSKiNfeR1ab and pSKiNfeR1abc that were mobilized to *S. meliloti* strain Sm2020 by biparental matings. Similarly, pSKiMS2NfeR1_1 and pSKiMS2NfeR1_3 were generated using Mut1Fw, Mut1Rv, Mut3Fw and Mut3Rv primer pairs carrying the desired mutations.

MAPS assays. Sm2020 cells carrying pSKiMS2NfeR1 or pSKiNfeR1 (control of column-binding specificity) were grown in MM and MM-NH_4 (0.5 mM ammonia) media to exponential phase. Bacterial lysates were subjected to affinity purification following the protocol described in Material and Methods.

Fluorescence reporter assays. The transcriptional fusions reporting promoter activity were generated in the promoterless vector pBB-*eGFP*. The full-length NfeR1 ($P_{nfeR1\text{-}213}$) and *glnII* (P_{glnII}) promoters were amplified with the primer pairs P14C2EcoRIFw/P14C2XbaIRv and HindIIIPglnII-400F /XbaImTSSglnIIR, respectively, from genomic DNA. The PCR products were digested with *Eco*RI/*Xba*I and cloned into pBB-*eGFP* to generate $pBBP_{nfeR1\text{-}213}$::*eGFP* and $pBBP_{glnII}$::*eGFP*. P_{nfeR1} devoid of the LsrB and NtrC binding sites ($P_{nfeR1\text{-}40}$) was generated by annealing the oligonucleotides P14C2_54/ P14C2_54i and cloning the resulting product into pGEM-T. $P_{nfeR1\text{-}100}$ was amplified using P14C2Fw/P14C2Rv primers and also cloned into pGEM-T. $P_{nfeR1\text{-}40}$ and $P_{nfeR1\text{-}100}$ were retrieved from pGEM-T by *Spe*I-*Xba*I restriction, and finally inserted in

pBB*eGFP* to yield pBBP$_{nfeR1\text{-}40}$::*eGFP* and pBBP$_{nfeR1\text{-}100}$::*eGFP*. The 100-nt long promoter $P_{nfeR1\text{-}100*}$ containing point mutations at the LsrB binding site was generated by annealing the oligonucleotides EcoRIPc14mutFw/XhoIPc14mutRv and further insertion of the resulting product between the *Eco*RI and *Xho*I restriction sites in pBB-*eGFP,* yielding pBBP$_{nfeR1\text{-}100*}$::*eGFP*.

The different transcriptional fusions of the NfeR1 promoter were also inserted in in the single-copy plasmid pABCa by amplification with avrIISRFw and avrIIGFPRv primers using pBB$_{PnfeR1\text{-}213}$::*eGFP*, pBBP$_{nfeR1\text{-}100}$::*eGFP* and pBBP$_{nfeR1\text{-}40}$::*eGFP* as templates. Then, PCR products were digested with *Avr*II and cloned into pABCa.

The translational reporter fusion of *ntrB* to *eGFP* was generated in plasmid pR-*eGFP*. For this, a genomic region of *ntrB* from its TSS to the 16th codon was amplified with the ntrBF/ntrBR primer pair. The resulting PCR product was digested with *Bam*HI/*Nhe*I and cloned into pR-*eGFP* to yield pR*ntrB*::*eGFP*. The reporter plasmid was transferred by biparental conjugation to Sm2020 harboring plasmids expressing either wild-type NfeR1 or the NfeR1abc variant. Transconjugants were evaluated using a two-plasmid genetic reporter assay *in vivo* described in Material and Methods.

Data availability. Raw RNAseq data can be accessed at the URL

Chapter 3 Proteins assisting regulation by AbcR1/2 and NfeR1

1. Background

RBPs are involved in the processing, stability, and activity of sRNAs. Chaperones assist and bring accuracy to sRNA-mRNA base-pairing and RNases have nucleolytic activity that is responsible for sRNA turnover and decay of the target mRNAs in riboregulation [101]. The long considered major, if not unique, RNA chaperone assisting *trans*-sRNA activity in bacteria, Hfq, is not ubiquitous, i.e., almost half of species lack a recognizable homolog. Moreover, in many bacteria expressing a canonical Hfq, including *S. meliloti*, this protein has a much more limited role in riboregulation than in enterobacteria [50, 236, 237]. More recently, gradient profiling by sequencing (Grad-seq) of *Salmonella* ribonucleoprotein complexes identified ProQ as a novel binding partner of a large set of highly structured sRNAs, which envisages a global Hfq-like role of this protein in *trans*-encoded mRNA regulation and virulence[116, 238, 239]. Indeed, both Hfq and ProQ share several RNA partners, indicating overlapping or competing functions of these two RNA chaperones [240]. However, ProQ is not as widespread as Hfq in bacteria, which suggests that other yet undiscovered RBPs with constrained phylogenetic distribution or unique to particular species may fulfil a chaperone function in riboregulation [241]. Affinity chromatography using aptamer-tagged RNAs as baits has been one of the experimental approaches of choice to capture the proteomes associated to several *trans*-sRNAs in enterobacteria and in the Hfq-less ε-proteobacteria *Helicobacter pylori* [242–245].

A functionally relevant Hfq-interacting protein is the major single-strand endoribonuclease RNase E, which is a catalytic component of the enterobacterial RNA degradosome that also includes the exoribonuclease polynucleotide phosphorylase (PNPase), the RNA helicase RhlB and the glycolytic enzyme enolase [16]. Through interaction with RNase E, Hfq recruits the degradosome

complex to the sRNA-mRNA interplay, thus promoting irreversible target mRNA degradation subsequent to primary sRNA-mediated translational repression [98]. RNases remain poorly characterized in most bacterial species. In *S. meliloti,* only YbeY and RNase III have been thoroughly studied. YbeY exhibits an unprecedented catalytic versatility, being able to cleave both ssRNA and dsRNA. Further, first experimental evidence suggests that *S. meliloti* YbeY promotes degradation of mRNAs from transporter genes upon their antisense interaction with the AbcR2 sRNA (e.g., *prbA*) [111]. On the other hand, RNase III, encoded by the *S. meliloti rnc* gene, is a widely distributed endoribonuclease specifically involved in cleavage of RNA duplexes. Remarkably, both endoribonucleases have a great impact in free-living and symbiotic *S. meliloti* physiology [108].

Instead of acting by base-pairing interactions with mRNAs, several sRNAs bind to and antagonize the activity of certain proteins. This minor class of RNA regulators typically mimic DNA or RNA motifs that are specifically recognized by certain proteins. The widely conserved 6S RNA and the CsrB family of sRNAs are well-characterized examples of riboregulators that act by target mimicry, outcompeting the σ^{70} RNAP holoenzyme and the C storage regulator CsrA at their cognate targets in gene promoters and mRNAs, respectively [61, 246]. Base-pairing and protein titration have long been considered mutually exclusive mechanisms of sRNA activity. A remarkable exception is McaS, an *E. coli trans*-RNA that relies on a dual mechanism involving both Hfq-dependent antisense interaction with a target mRNA and CsrA titration for controlling biofilm formation [247]. More recently, a new role has been uncovered for CsrA as RNA matchmaker in *Bacillus subtillis*, which further highlights the functional plasticity and diversity of the RBPs for regulation of transcription, translation, and RNA turnover [100, 248].

Affinity chromatography to pull-down novel proteins associated to MS2-tagged versions of three stress-induced *S. meliloti trans*-sRNAs, AbcR2, NfeR1

and EcpR1 unveiled several proteins functionally related to the flow of genetic information, probably linked to the activity of these sRNAs as translation inhibitors [119]. Strikingly, these experiments failed to identify RNA chaperones alternative to Hfq in the proteomes associated to the Hfq-independent NfeR1 and EcpR1 sRNAs. Interestingly, this assay identified the enzyme S-adenosylmethionine (SAM) synthetase (MetK) as a common binding partner of AbcR2, NfeR1 and EcpR1. In this Chapter, we have confirmed MetK binding to RNA and started to explore its possible role in riboregulation. On the other hand, we mined the RNase III and YbeY-dependent transcriptomes to search for misregulated AbcR1/2 and NfeR1 target mRNAs, as a first global approach to the function of both endoribonucleases in RNA silencing.

2. Results

2.1. MetK binds diverse RNA species

To search for novel RBPs, we performed affinity chromatography with wild-type and MS2-tagged AbcR2, NfeR1 and EcpR1 sRNAs expressed constitutively from a modified *lac* promoter in plasmid pSRK-C [91] (Fig. 39A). Additionally, a plasmid expressing the aptamer (MS2Term) was also used as common negative control to assess unspecific protein binding in affinity chromatography. Wild-type bacteria transformed with the different plasmids were grown under stress conditions that induce the endogenous expression of the three sRNAs, i.e., salt stress for NfeR1 and EcpR1 and stationary phase growth for AbcR2. Hfq was specifically recovered with AbcR2, which validates the procedure (Fig. 39C). Interestingly, MetK (a conserved SAM synthetase) was the only protein recovered with all the MS2-tagged sRNAs but not with any of the controls (wild-type RNA species and MS2Term) (Fig. 39B-C).

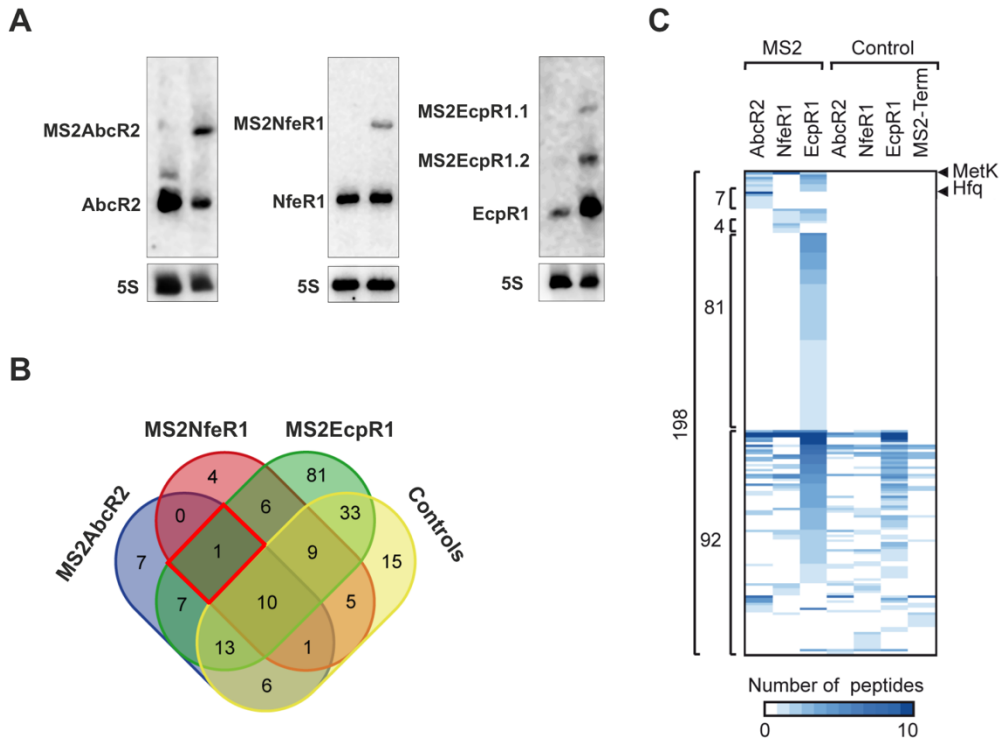


Figure 39. Identification of protein partners of the AbcR2, NfeR1, and EcpR1 *trans*-sRNAs. **(A)** Northern blot detection of aptamer-tagged sRNAs. Total RNA from strains transformed with either pSRK-MS2Term, pSRKMS2AbcR2/NfeR1/EcpR1, was probed with specific oligonucleotides targeting AbcR2, NfeR1 and EcpR1, respectively. The detected RNA species are indicated to the left of each panel. MS2EcpR1.1 and MS2EcpR1.2 correspond to full-length and processed forms of the sRNA, respectively. 5S rRNA was probed as RNA loading control. **(B)** Venn diagram showing the overlap of all identified proteins associated to the tagged-sRNAs and all control samples. **(C)** Heatmap representation of protein abundance as revealed by MS. Columns refer to the different MS2-sRNAs used in these experiments and the respective controls, rows represent individual proteins. For visualization purposes, the white-blue color bar represents the number of peptides identified for the corresponding proteins.

To further confirm MetK-sRNA interactions we performed binding assays with the purified protein and radiolabelled *in vitro* transcribed full-length AbcR2, NfeR1 and EcpR1 RNA species. Mass spectrometry discarded contamination of our MetK preparation with Hfq, as a possibility reported previously for other Ni-affinity purifications of His-tagged proteins from *E. coli* [249]. Binding reactions were analyzed by dot-blot on nitrocellulose and PVDF membranes that collect RNA-protein complexes and free RNA, respectively (Fig. 40; left panel).

Increased protein concentrations in the reaction mixtures resulted in reliable detection of increased amounts of bound sRNA for all three transcripts. K_D values derived from data of three independent experiments were 45.6, 21.2 and 6.7 nM, for AbcR2, NfeR1 and EcpR1, respectively, indicating different affinities of each sRNA for association with MetK (Fig. 40; right panel). Addition of a 100-fold molar excess of cold sRNA outcompeted binding of the radiolabelled sRNAs to MetK, whereas each sRNA was exclusively recovered in the PVDF membrane upon incubation with 2 μM BSA, which has no recognized RNA-binding ability. These two negative binding controls validate all three sRNA-MetK interactions.

To assess specificity of MetK interaction with RNA, in a new series of experiments, we similarly probed MetK binding to a set of different Hfq-independent RNA species expressed by *S. meliloti* (Fig. 40): the asRNA SmelC812 (165 nt) [45], $tRNA^{Met}$ (76 nt), the group II ribozyme RmInt1 (748 nt) [250], the mRNA annotated as *SMb20420* (627 nt) and the 5S rRNA (120 nt). The assays revealed binding of MetK to all these *in vitro* synthetized transcripts except to the 5S rRNA. In this case, K_D values calculated from 3-4 independent assays were 12.4, 1.3, 28.4 and 10.31 nM for SmelC812, $tRNA^{Met}$, RmInt1 and *SMb20420*, respectively, predicting a remarkable strong binding affinity of MetK for tRNA substrates. Together, our data thus uncover an unexpected promiscuity of the metabolic enzyme MetK for RNA binding.

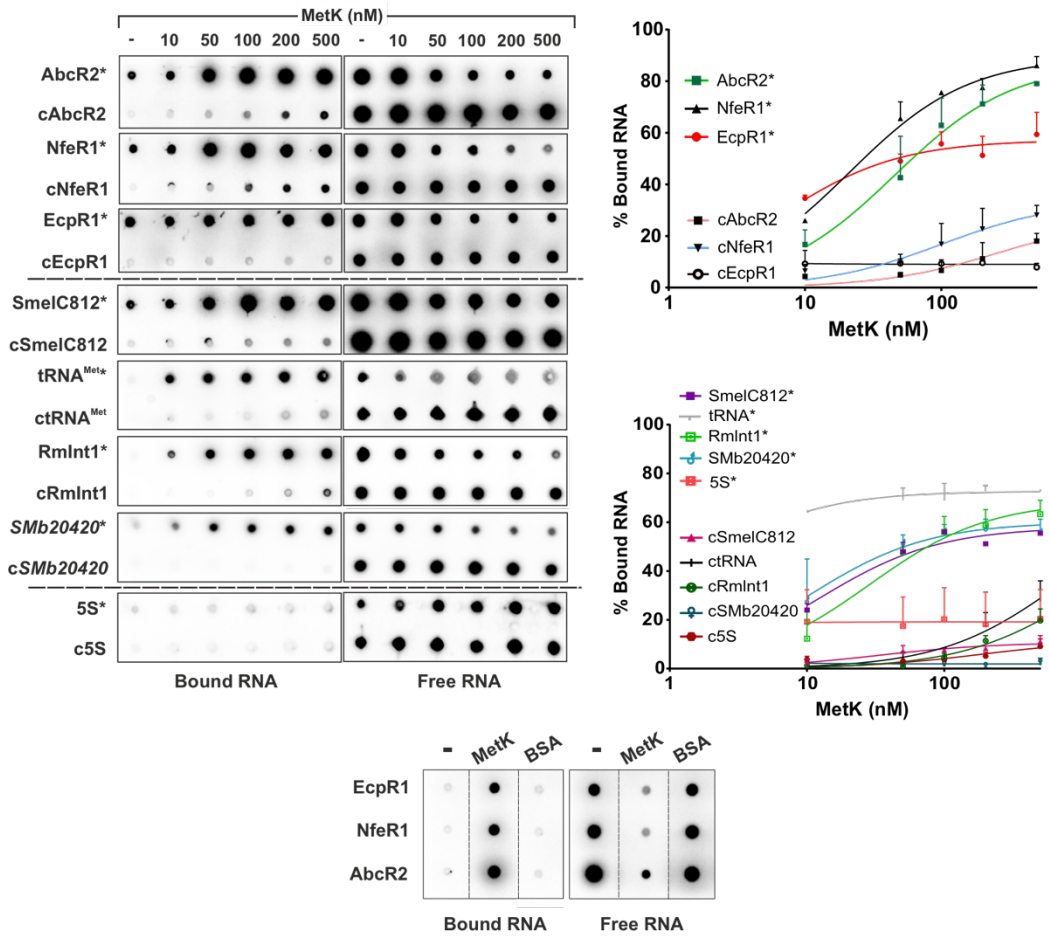


Figure 40. ***In vitro*** **analysis of MetK-RNA interactions.** Dot-blot assays to test binding of MetK to AbcR2, NfeR1, EcpR1, SmelC812, $tRNA^{Met}$, RmInt1, *SMb20420* and 5S RNA species. Radiolabelled transcripts (*) alone (1 nM) or with a molar excess of the corresponding cold RNA (cRNA; 100 nM) were incubated with increased concentrations of purified MetK as indicated on top of the panel. RNA-protein complexes and free RNA were collected by blotting of the reaction mixtures on nitrocellulose (left) and PVDF (right) membranes, respectively. Signal intensities were plotted in the graphs shown to the right. Binding affinities (K_D values) quoted in the text were calculated from data of three (in some cases four) independent experiments. Control filter binding assays were performed with 2 μM BSA or MetK (lower panel).

2.2. MetK-sRNA interaction does not influence canonical riboregulatory traits

In bacteria, sRNA-protein interactions typically result in either interference with protein activity or sRNA protection from degradation. Since MetK is not a

recognizable RNA chaperone, we first assessed whether binding to NfeR1, AbcR2 and EcpR1 influenced SAM homeostasis. We reasoned that SAM accumulation might parallel increased transcription of *metK*, which occurs in a single copy and is essential in *S. meliloti*. We therefore constructed a *metK* conditional deletion strain (SmΔ*metK*) that was complemented with plasmid pSK*metK*^{FLAG} expressing a FLAG-tagged MetK protein upon IPTG induction, as verified by Western-blot (Fig. 41A). Leaky transcription of *metK*^{FLAG} from the inducible *lacZ* promoter was enough to render cells viable (data not shown). We next used this strain as recipient of a translational fusion of the *S. meliloti metZ* gene to the *eGFP* reporter, that includes a predicted SAMII riboswitch (pR*metZ*::eGFP) (Fig. 41B). MetZ is involved in methionine biosynthesis, which is inhibited by the MetK-dependent intracellular SAM accumulation presumably sensed by the riboswitch fused to *eGFP*. As predicted, fluorescence of SmΔ*metK* bacteria co-transformed with plasmids pSK*metK*^{FLAG} and pR*metZ*::eGFP decreased as MetK accumulated with increased concentrations of the inducer (Fig. 41B). These data validate the *metZ* SAMII riboswitch as reporter of intracellular SAM levels in *S. meliloti*. Therefore, we next mobilized pR*metZ*::eGFP to NfeR1/AbcR2 and EcpR1 knock-out mutants (Sm2020 and Sm4011*ecpR1*) and co-transformed the resulting strains with plasmids expressing each sRNA upon IPTG induction, i.e., Sm2020 with pSKiNfeR1 or pSKiAbcR2, and Sm4011*ecpR1* with pSKiEcpR1. Double transconjugants were grown in TY broth to exponential phase and then induced for sRNA expression with 0.5 mM IPTG for a further 24 h. Induced expression of NfeR1 and EcpR1 in these conditions has been shown to result in productive target regulation and gain-of-function phenotypes [73, 97]. In our assays, fluorescence of the reporter strains was not altered upon IPTG addition to cultures (Fig. 41C), suggesting that overexpression of the sRNAs did not influence SAM levels. On the other hand, Western-blot probing of lysates from a *S. meliloti* derivative strain expressing

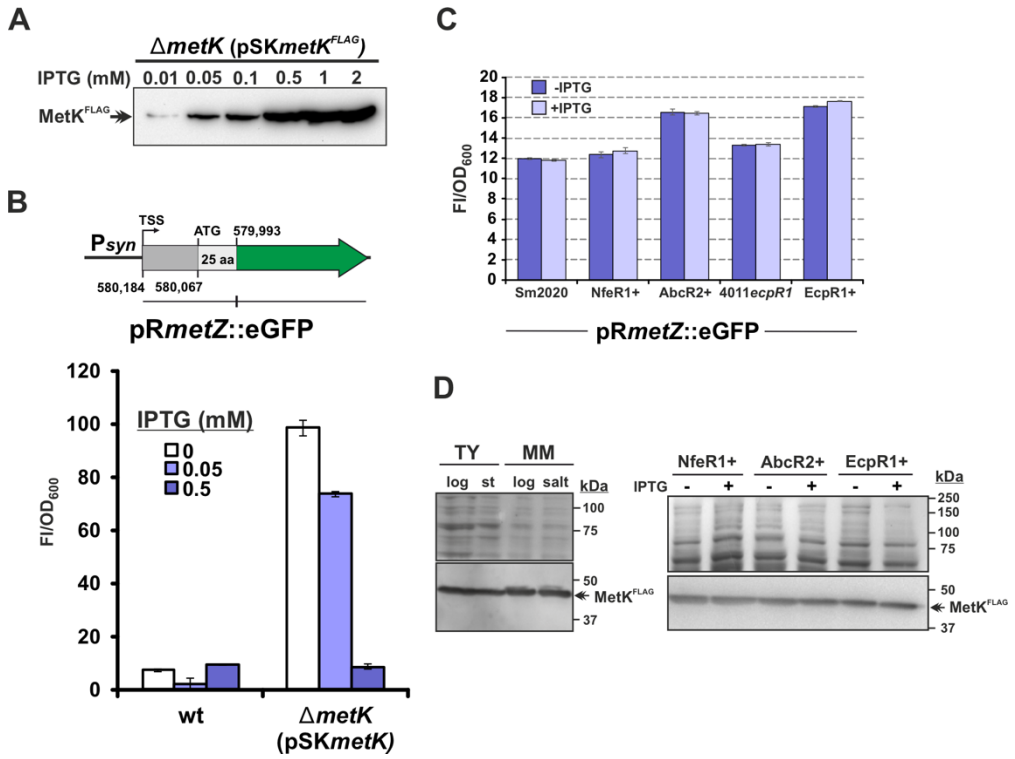


Figure 41. NfeR1, AbcR2 or EcpR1 (over)expression does not influence MetK activity. (A) Western-blot detection of MetK^{FLAG} accumulation in strain SmΔ*metK* complemented with pSK*metK*^{FLAG}. Bacteria were cultured in complete TY medium and IPTG was added at the indicated concentrations for induction of *metK*^{FLAG}. **(B)** SAMII riboswitch-based sensing of SAM levels. The bar graph shows fluorescence of the pR*metZ*::eGFP fusion in the wild-type strain and its SmΔ*metK* mutant derivative complemented with pSK*metK*, which expresses wild-type *metK* upon IPTG induction. The concentrations of IPTG in TY cultures is indicated. Values reported are means of three independent determinations and represent % of the fluorescence of uninduced cultures (100%). The map of the *metZ*::eGFP fusion is shown above the graph. The SAMII riboswitch is depicted as a grey box. Numbers indicate coordinates in the Sm1021 genome. P*syn*, constitutive Syn promoter. **(C)** Fluorescence of pR*metZ*::eGFP upon IPTG-induced expression of NfeR1, AbcR2 and EcpR1. Sm2020 is the NfeR1/AbcR2 double knock-out mutant whereas Sm4011 is the EcpR1 knock-out mutant. Expression of the sRNAs was induced from plasmids pSKiNfeR1, pSKiAbcR2 or pSKiEcpR1 with addition of 0.5 mM IPTG to TY cultures (OD_{600} 0.2) and further growth in this condition during 24 h. Values are means of three independent experiments**. (D)** MetK accumulation upon endogenous and induced (over)expression of the sRNAs. Western-blot detection of MetK^{FLAG} in Sm*metK*^{FLAG} (left panel) and Sm2019*metK*^{FLAG} transformed with pSKiNfeR1, pSKiAbcR2 or pSKiEcpR1 (right panel). Culture conditions to promote endogenous sRNA expression in Sm*metK*^{FLAG} are indicated on top (log, exponential phase cultures; st, stationary phase cultures; salt, osmotic upshift). In the Sm2019*metK*^{FLAG} derivatives expression of the sRNAs was induced with addition of 0.5 mM IPTG to TY cultures (OD_{600} 0.2) and further growth in this condition during 4 h. A portion of both Coomassie stained SDS-PAGE gels is shown as protein loading control on top.

MetK^{FLAG} from the chromosome (Sm*metK*^{FLAG}) cultured in conditions that upregulate the three sRNAs (i.e. stationary phase and salt shock), did not reveal obvious alterations of MetK accumulation (Fig. 41D; left panel), consistently with the housekeeping function of the enzyme. In line with this observation, similar experiments with and Sm2019*metK*^{FLAG} harbouring either pSKiNfeR1, pSKiAbcR2 or pSKiEcpR1, revealed that accumulation of MetK was not altered either by sRNA induced expression (Fig. 41D; right panel).

Finally, we tested whether MetK depletion influenced sRNA stability using NfeR1 as a proof of principle. This sRNA was reliably detected by Northern blot probing of MetK^{FLAG} CoIP-RNA from Sm*metK*^{FLAG} (Fig. 42A). However, its accumulation pattern in rifampicin-treated SmΔ*metK* bacteria complemented with pSK*metK*^{FLAG} was not altered regardless IPTG-induced MetK levels (Fig.

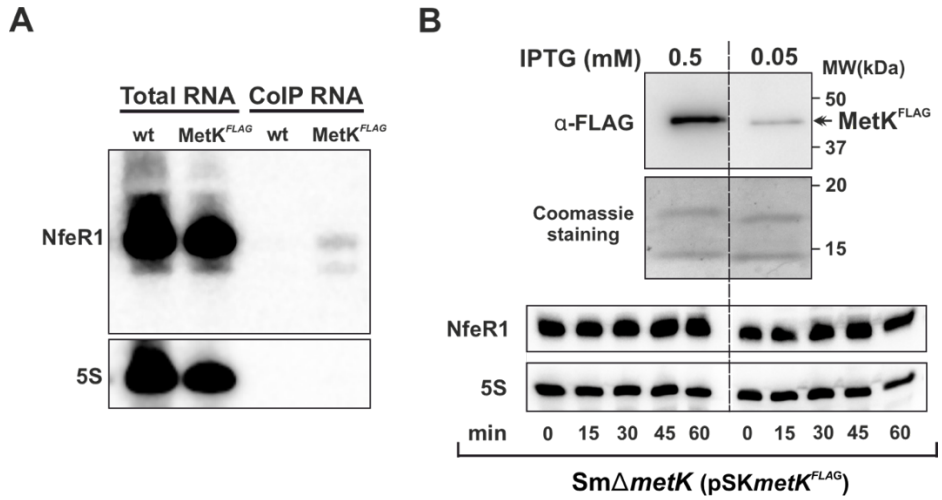


Figure 42. MetK does not impact NfeR1 stability. (A) Recovery of NfeR1 in MetK^{FLAG} CoIP-RNA. Northern blot probing of total and CoIP-RNA from wild-type and Sm*metK*^{FLAG} strains. RNA samples were obtained from MM cultures upon induction of NfeR1 expression by salt shock (400 mM NaCl for 1 h). 5S rRNA was probed as RNA loading control**. (B)** MetK depletion does not influence NfeR1 stability. Upper panel, Western-blot detection of MetK^{FLAG} in Sm*Δmet*K (pSK*metK*^{FLAG}) bacteria grown in MM supplemented with 0.5 or 0.05 mM IPTG for induction of *metK*^{FLAG} expression. A portion of the Coomassie-stained SDS-PAGE gel is shown below as protein loading control. Lower panel, Northern blot probing of total RNA derived from Sm*ΔmetK* (pSK*metK*^{FLAG}) grown in MM medium with 0.5 or 0.05 mM IPTG. NfeR1 expression was induced by salt shock and transcription was arrested by addition of rifampicin. Samples were withdrawn at the time points indicated below the panel. 5S rRNA was probed as RNA loading control.

42B). Together these results hint at a non-canonical novel function of MetK as sRNA-binding protein in bacteria.

2.3. YbeY and RNase III partially assist regulation by AbcR1/2 and NfeR1

mRNA decay initiated by RNases is usually concomitant to base-pairing with regulatory sRNAs. We used strand specific RNAseq to profile the transcriptomes of both *S. meliloti* YbeY and RNase III knock-out mutants [111, 251]. For that, we extracted total RNA from wild-type and mutant strains grown in TY under aerobic and micro-aerobic conditions. The latter induces a partial set of symbiosis-related genes that increase their expression in the microoxic environment of the nodule. RNA-Seq revealed a strong impact of these RNases in the *S. meliloti* transcriptome as expected from the pleiotropic phenotype of the respective mutants. Protein-coding genes misregulated in the *ybeY* mutant with respect to the wild-type strain ($log_2FC > 1$ or < -1 in the combined aerobic and microaerobic transcriptome) represent 27% of ORFs annotated in the *S. meliloti* genome, whereas in the case of RNase III were 42%. Indeed, 67% and 68% of mRNAs influenced by YbeY and RNase III, respectively, are expressed in micro-aerobic conditions, consistently with the reported impact of these RNases in symbiosis. Of note, 694 and 470 sRNAs were also scored as differentially accumulated between wild-type and *rnc* or *ybeY* mutant strains, respectively. Among them, AbcR1 is differentially regulated by both RNases whereas AbcR2 seems to be only YbeY-dependent. Differential accumulation of these sRNAs between strains only occur under microoxia, suggesting that both enzymes contribute to AbcR1/2 silencing within nodules.

We next compared the lists of AbcR1, AbcR2 and NfeR1 MAPS-identified targets with the set of mRNAs influenced by YbeY and RNase III as preliminary evidence of a general role of these RNases in riboregulation (URL) (Fig. 43).

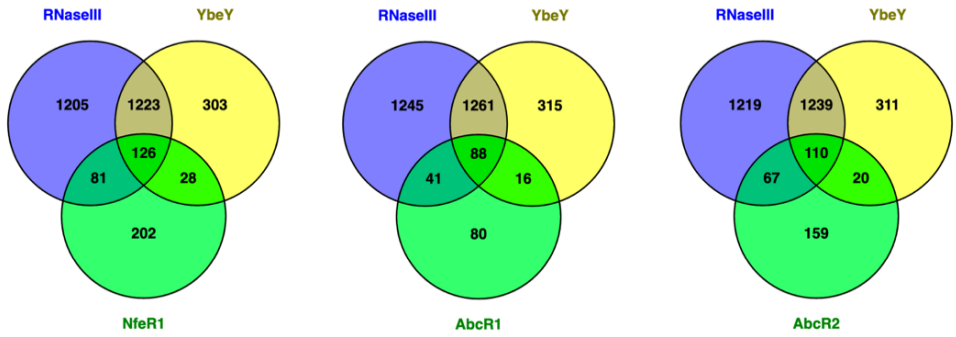


Figure 43. Impact of YbeY and RNase III in regulation by AbcR1/2 and NfeR1. Venn diagram showing the overlap of all MAPS-identified targets of AbcR1/2 and NfeR1 and mRNAs misregulated in YbeY and RNase III knock-out mutant strains cultured in aerobic and micro-aerobic conditions.

These comparisons revealed that more than 50% of mRNAs listed in each targetome were misregulated in the RNase mutants. NfeR1 target mRNAs involved in cell division and motility *rem*, *mcpT* and *minC*, in regulation of surface polysaccharides production and metabolism, (*exoX* and *exoW*), and the validated transporter *SMb20442,* seem to be exclusively affected by RNase III. Conversely, mRNAs involved in the response to osmotic stress *greA* and *aglE* are specifically misregulated in the YbeY mutant. On the other hand, *SMc03121,* together with mRNAs for the transport and regulation of mannitol and fructose metabolism, i.e., *smo* and *fcr*, are likely co-regulated by RNase III and YbeY. Linked to the NSR, we found that only the putative NfeR1 target *glnB* seems to be co-regulated by both RNases.

Regarding the AbcR1/2 target mRNA candidates, a subset of those encoding proteins for the transport and metabolism of diverse carbohydrates, sugar alcohols and amino acids, including the validated targets *aapJQ, livK*, and *prbA* are YbeY and RNase III-dependent, at least in one culture condition. This analysis thus anticipates a great impact of these endoribonulceases in the RNA silencing of metabolic genes, further indicating that AbcR1/2 and NfeR1 downregulation of some of these metabolic mRNAs could be redundantly assisted by both enzymes.

3. Discussion

Besides the well-known RNA chaperone Hfq, the repertoire of RBPs and their role in riboregulation remains unexplored in *S. meliloti* [50]. The reported profiling of glycerol gradient-sorted ribonucleoprotein complexes in *Salmonella* identified ProQ as a novel bacterial RNA chaperone with expected widespread Hfq-like functions [116, 238, 240]. ProQ orthologues are identifiable by a domain of the FinO protein that mediates antisense RNA regulation of F plasmid transfer [241]. However, in the large α–subgroup of proteobacteria, occurrence of ProQ/FinO-domain proteins is likely restricted to the species *R. leguminosarum* and *Caulobacter cresecentus* [241]. Similarly, rhizobial genomes lack genes putatively encoding members of the well-characterized CsrA translational inhibitors [36]. The larger sets of sRNAs uncovered in α-proteobacteria as compared to enterobacteria, and the known functional protein redundancy (e.g., RNases) anticipate an even larger number of RBPs encoded by the multipartite rhizobial genomes. Therefore, specific screens are required to identify RBPs with putative roles in riboregulation and overall RNA metabolism within this group of bacteria. In this scenario, our laboratory described the *S. meliloti* sRNA-associated proteomes as revealed by affinity chromatography-based capture [119]. Remarkably, this screening and its further validation by independent methodologies identified the major methyl donor synthetase MetK as unexpected non-canonical bacterial sRNA-binding protein. Notably, our results also provide evidence of a striking ability of MetK for binding functionally diverse RNA species.

3.1. *S. meliloti* MetK is a novel RBP

Three characterized *S. meliloti trans*-sRNAs, which are stress-induced and widely conserved in α-proteobacterial species, AbcR2, NfeR1 and EcpR1, were

selected [62, 63] to search new RBP in *S. meliloti*. Intriguingly, no promising candidates to fulfil an Hfq-like chaperone role (i.e., with recognizable RNA-binding domains; RBDs) were envisaged among the protein partners of the Hfq-independent NfeR1 and EcpR1 sRNAs. However, all the three proteomes shared components related to transcription, translation and RNA turnover. Besides the proteins related to the flow of genetic information, our curated lists of putative RBPs contain a striking number of metabolic enzymes. Nonetheless, it has been already reported that binding of the small alarmone synthetase RelQ from the Gram-positive pathogen *Enterococcus faecalis* to ssRNA allosterically antagonizes the enzyme activity [252, 253], thus anticipating the ability of metabolic enzymes to bind RNA in bacteria. In addition, recent screenings for RBPs in eukaryotic organisms have revealed that a large fraction of metabolic enzymes indeed bind to polyadenylated RNAs via non-classical RBDs [254–256]. Moreover, comprehensive studies of two conserved enzymes across all animal kingdoms, glyceraldehyde-3-phosphate and aconitase, demonstrated that these proteins have dual functions where RNA binding is incompatible with the enzymatic activity [257–259]. Affinity chromatography, CoIP with the FLAG-tagged protein and *in vitro* assays unambiguously confirmed that one of such metabolic enzymes, MetK, is a common interacting partner of AbcR2, NfeR1 and EcpR1 sRNAs.

Massive binding to sRNAs may hint at a chaperone-like role of MetK in promoting RNA stability. However, probing of NfeR1 sRNA upon rifampicin treatment of MetK-depleted cells argued against this possibility. As documented for other prokaryotic and eukaryotic RNA-protein interactions, we therefore considered protein titration as a plausible consequence of MetK binding to sRNAs. To test this hypothesis, we set up a SAMII riboswitch-based assay that reliably sensed increased SAM accumulation upon MetK synthesis, thus becoming a novel genetic tool to investigate biology of macromolecule

methylation in rhizobia. However, our assays did not reveal either significative alterations of intracellular SAM levels or MetK accumulation upon endogenous or induced (over)expression of either of the three sRNAs. These findings predict that the function of MetK in riboregulation, if any, is not canonical.

Of note, our *in vitro* assays reliably revealed that MetK is also able to bind other non-coding and protein-coding RNA species with diverse cellular functions, thus anticipating an unexpected promiscuity of a protein lacking recognizable RBDs for RNA binding in bacteria. All the tested transcripts are highly divergent in length (76 to 748 nt) and primary nucleotide sequence, which suggest that interaction with MetK most likely rely on the recognition of structural modules with widespread occurrence among bacterial transcripts, as reported for ProQ [260]. In this regard, MetK evidenced a particular strong affinity for binding to tRNA. This broad RNA-binding specificity is also a feature of well-characterized RNA chaperones such as Hfq. The plethora of RNA ligands and molecular interactions explain the diversity of newly discovered Hfq functions in protein synthesis beyond its recognized role in sRNA-mediated regulation [261]. To date, MetK has been solely viewed as a key enzyme involved in macromolecule metabolism and epigenetic control of gene expression. Thus, our data add this protein to the emerging group of moonlighting enzymes with likely novel functions in riboregulation and other RNA-dependent cellular processes that merit further investigation in the near future.

3.2. Broad impact of YbeY and RNase III in regulation by AbcR1/2 and NfeR1

Ribonucleases are key elements of post-transcriptional regulatory networks that are poorly characterized in rhizobia. Both *S. meliloti* YbeY and RNase III are endoribonucleases highly efficient in degrading RNA duplexes, yielding similar cleavage patterns of generic structured dsRNA substrates *in vitro* [108, 111].

Accordingly, transcriptomics uncovered large sets of mRNAs whose steady-state levels are influenced by both enzymes, but also hinted at specific activity features. This functional specificity likely relies on a broader YbeY activity that includes degradation of ssRNA, thus predicting a yet uncharacterized RNase E-like role of this novel enzyme in RNA processing. Interestingly, besides mRNAs, we found that a great number of sRNAs of all types, including AbcR1 and AbcR2, were misregulated by a lack of YbeY and RNase III, thus emphasizing the importance of these RNases in the dynamics of the *S. meliloti* non-coding transcriptome. We therefore reasoned that both YbeY and RNase III might be major players in RNA silencing promoted by antisense interactions between AbcR1/2 or NfeR1, and their respective arrays of target mRNAs. The extensive comparisons among the YbeY- and RNase III-dependent transcriptomes, and the AbcR1/2 and NfeR1 mRNA interactomes uncovered by MAPS supported this hypothesis. These analyses further delineated metabolic traits post-transcriptionally regulated by these sRNAs that are either common or specific targets of YbeY and RNase III activity. Nonetheless, direct RNase-mRNA interactions and concomitant mRNA degradation must be demonstrated to exclude indirect RNase effects on the turnover of these AbcR1/2 and NfeR1 targets. As double-strand RNases, YbeY and RNase III might initiate cleavage of sRNA-mRNA duplexes, whereas activity of YbeY as single-strand endoribonuclease might be involved in the further decay of the target mRNA. Our dataset provides a foundation to further investigate YbeY and RNase III function in riboregulation and, most specifically, the activity mechanisms of AbcR1/2 and NfeR1 as *trans*-acting sRNAs.

4. Experimental setup

Bacterial strains and growth conditions. Bacterial strains and plasmids specifically used in this Chapter, along with their relevant characteristics, are listed in Table 8. To assess the MetK-dependent stability of NfeR1 upon salt

stress, exponentially growing bacteria in MM were cultured for a further 1 h after addition of 400 mM NaCl. For induction of FLAG-tagged protein and sRNA expression, IPTG was added to a final concentration of 0.5 mM to exponential phase cultures, unless otherwise indicated.

Table 8. Bacterial strains and plasmids used in Chapter 3.

Strain/Plasmid	Relevant characteristics	Reference/Source
STRAIN		
S. meliloti		
Sm4011Δ*ecpR1*	Sm4011 Δ*ecpR1* derivative; *expR*/*sinI* double mutant, Nx^r, Sm^r	[73]
Sm*metK*^{FLAG}	Sm2011 derivative chromosomally ncoding MetK^{FLAG}	This work
Sm2019*metK*^{FLAG}	Sm2019 derivative chromosomally ncoding MetK^{FLAG}	This work
Sm2011Δ*metK*	Sm2011 Δ*metK* derivative. Only viable if complemented with *metK*-expressing plasmids	This work
PLASMIDS		
pK18*metK*^{FLAG}	Suicide plasmid for FLAG-tagging of chromosomal *metK* gene	This work
pK18Δ*metK*	Suicide plasmid for *metK* deletion; Er^r, Km^r	This work
pSK*metK*^{FLAG}	pSK_FLAG expressing *metK*^{FLAG}; Km^r	This work
pSK*metK*	pSRKKm expressing wild-type *metK*; Km^r	This work
pSKGm*metK*^{FLAG}	pSRKGm expressing *metK*^{FLAG}; Gm^r	This work
pSRK-R2	pSRK-C derivative constitutively expressing AbcR2	[91]
pSRK-NfeR1	pSRK-C derivative constitutively expressing NfeR1	[97]
pSRK-EcpR1	pSRK-C derivative constitutively expressing EcpR1	[119]
pSKiAbcR2	pSRKKm carrying the AbcR2 coding sequence fused to *sinR*-P_{sinI}	Chapter 1
pSKiNfeR1	pSRKKm carrying the AbcR1 coding sequence fused to *sinR*-P_{sinI}	Chapter 2
pSKiEcpR1	pSRKKm carrying the EcpR1 coding sequence fused to *sinR*-P_{sinI}	[73]
pR*metZ::eGFP*	*metZ::eGFP* translational fusion (-/+ relative to *metZ* AUG); Ap^r, Tc^r	This work
p16*metK*	pET16b derivate carrying N-terminally 10xHis-tagged MetK	This work

Oligonucleotides. Sequences of the oligonucleotides used as probes for Northern hybridization, or as amplification primers for cloning and RT-qPCR are listed in Table 9.

Table 9. Oligonucleotides specifically used in Chapter 3. Restriction sites are underlined.

Oligonucleotides	5'-sequence-3'	Use
PbAbcR2	GAGGAGAAAGCCGCTAGATGCACCA	Northern blot probing
PbNfeR1	TGCTTGATCTGATTGGCAACCGGGA	
PbEcpR1	AGCACGAGGAATCCGCAAGTCCGAG	
Pb5S	TACTCTCCCGCGTCTTAAGACGAA	
NdeImetkF2	GCTGCATATGCGCGCAAACTACCTGTTCACAAGTG	*metK* amplification and FLAG tagging
metK3'XhoI	ATCTCGAGTTAGGCCGCGCGGCTCGCCG	
EcoNdemetkF	GCTGGAATTCCATATGCGCGCAAACTACCTG	
BamTGAflagR	ATTAGGATCCTCACTTGTCATCGTCATCCTTG	
3metKOBamHI_F	ATTAGGATCCGCCGTCCCGGCGGGGCAGCTT	
4metKOPstII_R	CGGCTCTAGATGGCTTCGTAGCGCGTGCTC	
XbaImetkR	TAATTCTAGAGGCCGCGCGGCTCGCCGCTCTT	
pT7C15	GTTTAATACGACTCACTATAGAGCTGGTGCATCTAGCGG	*In vitro* transcription
termC15	AAAAAAAGAGGGCCACAAGAAACTTGGG	
pT7C14	GTTTAATACGACTCACTATAGATCGATCGGGCAGCGCAC	
termC14	AACAACGCCCGCTGGGGGAGGA	
pT7C291	GTTTAATACGACTCACTATAGTTACCCCATGATGCTCAG	
termC291	GTCGACTGGTCCGGTTGCAGCACGAGGAAT	
pT7C812	GTTTAATACGACTCACTATAGATCCTTGTCGTCCCTGCC	
termC812	ATGCCTTTGGCCGGGCGG	
pT7tRNA	GTTTAATACGACTCACTATAGGGCGGAGTAGCTCAGTAG	

termtRNA	TGGTAGCGGAGGAGGGATT	
pT75S	GTTTAATACGACTCACTATAGCGACCTGGTGGTTCTG	
term5S	AGACCTGGCAGCGACCTA	
pT7SMb20420	GTTTAATACGACTCACTATAGAGAATCAGGAGAACCGA	
termSMb20420	CTAACTCTGAAGGCTCGCAA	
2metKKOBamHI_R	AGATGGATCCAGCGGTCCAAACGTTTAGTGG	*metK* deletion
3metKOBamHI_F	ATTAGGATCCGCCGTCCCGGCGGGGCAGCTT	
4metKOPstII_R	CGGCTCTAGATGGCTTCGTAGCGCGTGCTC	
TSS_metZ_Fw	CTTAGGATCCATCCCGTGGTGATTTGGC	Amplification of *metZ*-riboswitch
metZ_75_Rv	CATTGCTAGCCGAGGTCTCGCCATACTG	

Construction of *S. meliloti* mutants. To endow *S. meliloti* MetK protein with a FLAG-tag at their C-termini, the MetK CDS was amplified by PCR using primers EcoNdemetkF/XbaImetkR that introduce *Nde*I and *Xba*I sites. The PCR fragment was inserted into the corresponding restriction sites of pSK_FLAG to yield pSK-*metK*^{FLAG}. pSK_FLAG carries the 3×FLAG sequence followed by a TGA stop codon and an IPTG-inducible P_{lac} closely upstream the *Nde*I restriction site (Table 2).

Strain Sm*metK*^{FLAG} and Sm2019*metK*^{FLAG} was generated by replacement of the wild-type allele via pK18*metK*^{FLAG}-mediated double recombination. This plasmid was generated by PCR amplification of the MetK^{FLAG} CDS from pSK_*metK*^{FLAG} and a 1,300-bp fragment of its downstream region from genomic DNA with primer pairs EcoNdemetkF/BamTGAflagR and 3metKOBamHI_F/4metKOPstII_R, respectively. The first fragment was restricted with *Eco*RI/*Bam*HI and the second with *Bam*HI/*Xba*I. All restriction sites were incorporated into the PCR primers except *Xba*I, which is encoded in the genome. In the same ligation reaction both digestion products were inserted between the *Eco*RI and *Xba*I sites of the suicide vector pK18*mobsacB* to yield

pK18*metK*^{FLAG}. Proper expression of MetK^{FLAG} in strain Sm*metK*^{FLAG} was confirmed by Western blot. For the CoIP assay, wild-type and Sm*metK*^{FLAG} strains were grown in MM media to exponential phase and then, bacterial lysates were subjected to affinity purification. These protocols are described in detail in Material and Methods.

To achieve MetK depletion in strain Sm2011, its CDS was deleted by pK18Δ*metK*-mediated double recombination in the presence of plasmid pSKGm*metK*^{FLAG}. Plasmid pK18Δ*metK* was generated by PCR amplification of genomic DNA with the pair of primers 1metKOEcoRI_F/2metKKOBamHI_R and 3metKOBamHI_F/4metKOPstII_R, ligation of the resulting fragments at the cohesive *Bam*HI overhangs and insertion of the tandem between the *Eco*RI and *Pst*II sites of pK18*mobsacB*. pSKGm*metK*^{FLAG} was constructed by extraction of a fragment containing the *metK*^{FLAG} sequence by *Nde*I/*Xba*I digestion from pSK*metK*^{FLAG} and cloned into pSRKGm.

Plasmids pSKiEcpR1, pSKiNfeR1 and pSKiAbcR2 were mobilized as required to strains Sm4011*ecpR1*, Sm2020 and Sm2019*metK*^{FLAG} for IPTG-induced transcription of the wild-type EcpR1, NfeR1 and AbcR2 sRNAs, respectively.

Fluorescence reporter assays. As reporter of intracellular SAM accumulation we constructed plasmid *pRmetZ::eGFP*, which expresses a translational fusion of the SAM-II riboswitch, preceding the *metZ* gene, to *eGFP*. For that, a *metZ*-derived fragment extending from its TSS to the 25th codon was amplified with primers TSS_metZ_Fw and metZ_75_Rv, and inserted between the *Bam*HI and *Nhe*I sites of *pR_eGFP*. The reporter plasmid was transferred by biparental conjugation to Sm4011*ecpR1* and Sm2020 harboring plasmids expressing either wild-type EcpR1, NfeR1 or AbcR2, as well as to SmΔ*metK*.

Filter binding assays. The MetK CDS was PCR amplified from genomic DNA using the primers metK3'XhoI/NdeImetKF2 and cloned into the vector

pET-16b (Novagen) between the *Nde*I/*Bam*HI restriction sites, yielding p16MetK encoding a His-tagged MetK. Recombinant MetK was produced and purified as described in Material and Methods.

For *in vitro* synthesis and labelling of RNAs, the AbcR2, NfeR1, EcpR1, SmelC812 [38], tRNAMet, *SMb20420* and 5S rRNA transcripts were amplified from genomic DNA and the group II ribozyme RmInt1 from plasmid pKGEMA4 [42] with the primer pairs listed in Table 9 All forward primers incorporate the T7 promoter sequence for transcription with T7 RNA polymerase from PCR templates (250 ng). The protocol for the purification of the synthetized transcripts is described in Material and Methods.

Binding reactions were performed with 1 nM radiolabeled probes and the indicated concentrations of purified recombinant MetK (0-500 nM). Binding specificity was assessed by competition experiments in the presence of a molar excess (100 nM) of the corresponding unlabelled sRNA. BSA (2 μM) was also used instead of MetK as negative protein control for binding to the *trans*-sRNAs. Reactions were loaded into a dot-blot device (Bio-Dot; Bio-Rad) and analyzed with the Personal FX equipment and Quantity One software (Bio-Rad) as described in Material and Methods.

Data availability. Comparisons between endoribonucleases RNaseIII and YbeY-regulated mRNAs and AbcR1/2 and NfeR1 interactomes can be accessed at the URL

General Discussion

The mutualistic N-fixing root nodule rhizobia-legume symbioses are built on a strict metabolic cooperation between the partners. Rhizobia must cope with drastic shifts in nutrient availability during colonization of soil, rhizosphere, root, and nodule cells [24, 175]. Thus, the symbiotic success greatly depends on the ability of rhizobia to accurately adapt their metabolism to the soil and plant environments. The metabolic versatility of rhizobia is supported by large and multipartite genomes, with a generous genetic endowment arranged in complex networks devoted to nutrient uptake and catabolism [85, 262]. To date, regulation of these metabolic pathways has been almost exclusively attributed to proteins involved in differential transcription or specific post-translational modifications, but the underlying post-transcriptional mechanisms are largely unknown [32–35]. However, since the massive discovery of sRNAs in prokaryotes, RNA-mediated post-transcriptional rewiring of metabolism has been regarded as a major and ubiquitous adaptive trait, contributing greatly to bacterial fitness in fluctuating environments [263]. Knowledge about this level of genetic regulation in rhizobia is still rather scarce and mostly derives from work on the α-proteobacterium *S. meliloti* [45, 46, 264], but even in this model plant symbiont only a handful of sRNAs and proteins that assist riboregulation have been characterized.

In this Thesis, we have deciphered the regulation, architecture, and mechanistic principles of the regulatory networks controlled by the *S. meliloti* base-pairing *trans*-sRNAs AbcR1, AbcR2 and NfeR1. Our data show that these three sRNAs, their transcriptional regulators and their target mRNAs are arranged into a dense overlapping post-transcriptional regulon that pervasively contributes to *S. melitoti* metabolic reprogramming throughout the symbiotic transition (Fig. 44).

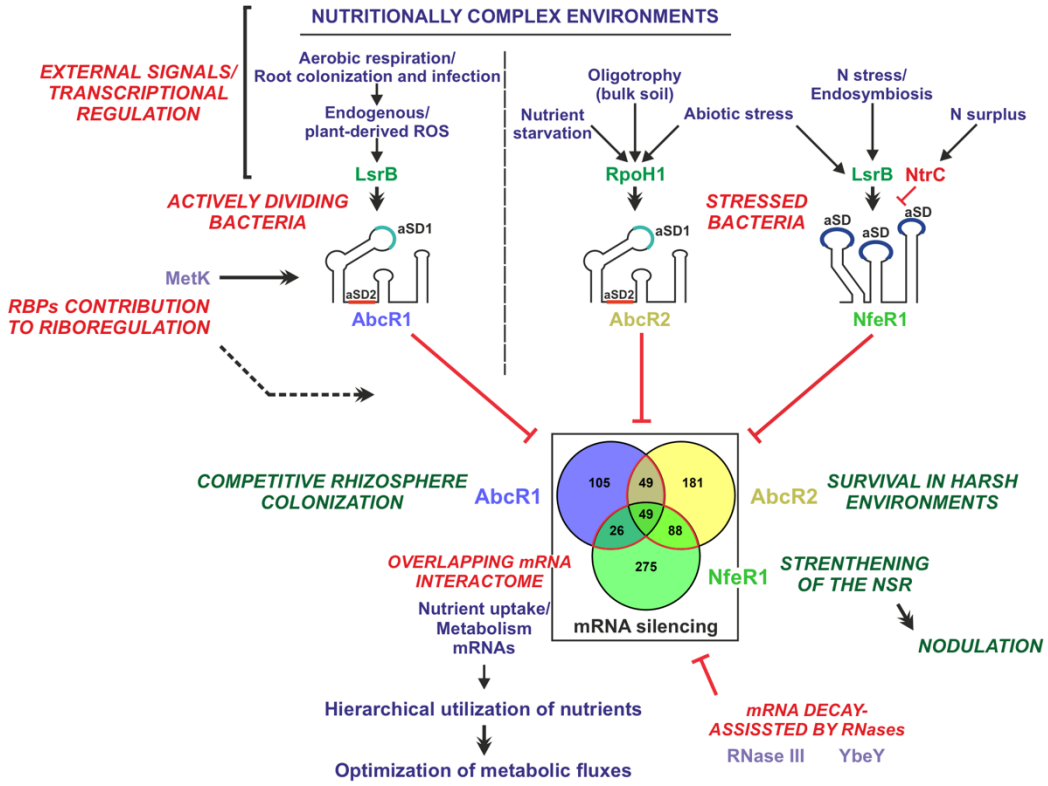


Fig. 44. The AbcR1/2 and NfeR1 dense overlapping post-transcriptional metabolic regulon. Graphical summary of data presented in this work. See text for details.

1. Transcriptional regulation of AbcR1/2 and NfeR1

Previous expression profiling and data presented here place AbcR1/2 activity at nutritionally complex environments during rhizobial free-living growth in soil and rhizosphere, whilst NfeR1 responds more specifically to the N signalling that operates during nodulation and symbiotic nitrogen-fixation. Here, we identified the LysR-type symbiotic regulator LsrB, the alternative σ factor RpoH1, and the major NSR regulator NtrC as the primary determinants of the differential accumulation of these sRNAs. LsrB, which transduces the cell redox state, is absolutely required for AbcR1 and NfeR1 expression. LsrB is a seemingly constitutive transcription factor that undergoes redox-dependent post-translational modifications that alter its specificity for promoter binding [151],

which might be a first determinant of the differential AbcR1 and NfeR1 expression in free-living and symbiotic bacteria. Remarkably, NfeR1 is expressed from a dual-mode promoter, whose transcriptional output is further tuned by NtrC-mediated repression boosting sRNA accumulation under N stress conditions prevailing in the legume rhizosphere at the onset of nodulation [168, 234]. Both AbcR1 and NfeR1 are likely transcribed by the major σ RNAP subunit, RpoD (σ^{70}), whereas the stress-induced transcription of AbcR2 depends on the alternative σ factor RpoH1. Strikingly, RpoH1 operates in both free-living and symbiotic rhizobia [149], which suggests that AbcR2 silencing within nodules involves additional levels of regulation.

2. AbcR1/2 and NfeR1 interactomes and activity mechanisms

The identification of mRNA and protein partners is key to understand the function and activity mechanisms of bacterial sRNAs. In particular, the identification of the target mRNAs, usually multiple for a single *trans*-sRNA, remains challenging [265, 266]. Computational tools typically predict large sets of target mRNA candidates, but the limited complementarity between the partners often leads to exceedingly high false-positive prediction rates [71]. Affinity chromatography of aptamer-tagged *trans*-sRNAs allow tackling the comprehensive genome-wide profiling of their interactomes (mRNAs and proteins) in growth conditions that stimulate endogenous upregulation of each sRNA [159]. In this work, we implemented this technology to further delineate the function and mode of action of the AbcR1/2 and NfeR1 *trans*-sRNAs. RNA-Seq profiling of the RNA species captured by affinity chromatography of all three MS2-tagged transcripts (MAPS) unveiled exceptionally large and overlapping mRNA interactomes, which jointly represent more than 7% of the *S. meliloti* protein coding genes. Most of these mRNAs encode transport and metabolic proteins for the uptake and assimilation of amino acids and carbohydrates,

anticipating a prominent role of these three sRNAs in the pervasive regulation of *S. meliloti* metabolism.

The large sets of mRNAs identified as common targets of AbcR1/2 and NfeR1 support computational predictions suggesting regulation via aSD motifs of very similar, but not identical, nucleotide sequence that most likely remain single-stranded in the three molecules. As described for other bacterial *trans*-sRNAs that regulate multiple targets [88, 267], these regulatory motifs are not unique but rather occur in two and three sites within AbcR1/2 and NfeR1, respectively. Further genetic dissection of some of these sRNA-mRNA base-pairing interactions uncovered independent regulatory functions for the similar AbcR1/2 aSD seeds, whilst demonstrating an unprecedented redundant contribution of the three identical interaction sites within NfeR1 for target regulation. Most importantly, the same genetic assays evidenced that the productive AbcR1/2, and most certainly NfeR1, interactions with their mRNA partners are modifiable for regulation of non-cognate targets, which opens unexplored possibilities for the engineering of this large metabolic RNA network.

Proteins acting as RNA chaperones (e.g., Hfq or ProQ) typically assist short base-pairing between the *trans*-sRNAs and their targets [240], as that described for AbcR1/2 and NfeR1. Unlike in enterobacteria, the well-characterized RNA-chaperone Hfq seems to have a rather limited role in riboregulation in rhizobia [50]. Indeed, NfeR1 belongs to the group of *S. meliloti* Hfq-independent sRNAs, which is prevalent in this bacterium [50, 97]. Neither Hfq nor other protein with recognizable RBDs were identifiable in the NfeR1 associated proteome captured by affinity chromatography of the MS2-tagged RNA. We therefore speculate that the redundancy of the identical aSD motifs for targeting, might render NfeR1 activity independent of RNA matchmakers. Instead of RNA chaperones, we strikingly identified the SAM synthetase, MetK, as common binding partner of diverse RNA species, including AbcR2 and NfeR1. Binding of likely

moonlighting metabolic enzymes to eukaryotic and prokaryotic non-coding RNAs has been previously reported, but the biological significance of these interactions remains largely elusive [254–256, 258]. Our findings hint at a non-canonical function of MetK in regulation by *trans*-sRNAs, i.e., binding to AbcR2/NfeR1 neither influences sRNA stability nor alters intracellular SAM levels. As synthetase of the major methyl donor for methyl transferases [268, 269], MetK likely plays a role in the epigenetic control of gene expression. In such scenario, the methylation patterns of AbcR2 (probably also of AbcR1) and NfeR1 and their target mRNAs might be novel determinants for the fine-tuning of riboregulation that must be investigated.

MAPS captures sRNA-mRNA base-pairing, but it is not inherently designed to inform about the impact of these interactions on target mRNA stability. However, the markedly uneven distribution of sequencing reads over large sets of mRNAs co-purified with tagged AbcR1/2 and NfeR1 suggests a major canonical RNA silencing mechanism involving accelerated message decay upon primary translation inhibition by base-pairing at the mRNA RBS. MAPS profiles also predict alternative activity mechanisms relying on interactions at the CDS or 3' regions of certain target mRNAs that merit further investigation. The set of ribonucleases and degradosome-like assemblies containing Hfq are the effectors of sRNA-mediated mRNA silencing [162, 264]. Profiling of the AbcR2 and NfeR1 associated proteomes did not provide solid clues about the composition of these silencing higher order protein complexes in *S. meliloti* [119]. However, the comparison among the AbcR1/2 and NfeR1 targetomes and the arrays of mRNAs influenced by RNase III and YbeY, which are the only characterized ribonucleases in *S. meliloti* [108, 111], predicts an extensive contribution of the two enzymes to target mRNA regulation. In fact, a considerable set of mRNAs are co-regulated by both endoribonucleases, suggesting that mRNA silencing could be redundantly catalyzed by RNases with similar substrate preference such

as RNase III and YbeY for dsRNA. Nonetheless, these comparisons also envisaged likely specific contributions of RNase III and YbeY to mRNA regulation, which might involve activity of YbeY on ssRNA species released upon initial cleavage of the RNA duplexes. Our MAPS setup is thus suitable to investigate the turnover dynamics of the AbcR1/2 and NfeR1 mRNA interactomes upon expression of the sRNA baits in the relevant ribonuclease knock-out mutants.

3. A target-centric view of AbcR1/2 and NfeR1 function in *S. meliloti*

Despite of their overlapping targeting potential, MAPS uncovered functional specificities for AbcR1/2 and NfeR1 largely linked to the transcriptionally regulated expression of each sRNA in different environments. AbcR1 most likely operates in actively dividing rhizobia colonizing niches with highly diverse and available enough nutrients, e.g., complete growth media, rhizosphere, or infection threads within the nodule primordia [24, 173]. Massive but controlled AbcR1-mediated downregulation of transport and metabolic mRNAs in response to shifts in nutrient availability would enable *S. meliloti* the hierarchical use of the available substrates to optimize metabolic fluxes. Thus, AbcR1 would confer bacteria with an advantage to colonize extremely selective environments such as the rhizosphere or the root rhizoplane under mild conditions. In contrast, AbcR2 expression requires an external stress imposed by either nutrient depletion (e.g., stationary phase growth) or abiotic variables (e.g., salinity) [91]. Therefore, we predict an AbcR1-like function for AbcR2 that would help *S. meliloti* survive the oligotrophy of soil or colonize the legume rhizosphere under the abiotic stress conditions featuring most agroecosystems.

Unlike AbcR1/2, NfeR1 is not detected in nutrient rich media but rather responds specifically to the N signalling and other conditions of the endosymbiotic compartments (e.g., osmotic stress) in an NtrC-dependent manner.

Under N stress imposed to free-living *S. meliloti* cultures, NfeR1 was shown to specifically target arrays of mRNAs for N assimilation (e.g., *gdhA*), cell cycle progression or the regulation of the NSR itself by NtrBC (e.g., *ntrB*). Silencing of the uptake and assimilation of combined N at the rhizosphere or within nodules is required for nodulation and transfer of the fixed nitrogen to the plant [168, 175, 234], whereas NfeR1-mediated fine-tuning of the cell cycle might impact bacteroid differentiation [196, 198, 199]. Regulation of morphological differentiation by N-responsive *trans*-sRNAs is unprecedented in rhizobia but has been already described during heterocyst formation in N-fixing filamentous cyanobacteria [170, 270–272]. MAPS-based deciphering of the NfeR1 mRNA interactome from infective *S. meliloti* bacteria prior to or during bacteroid differentiation will help faithfully dissect the NfeR1 symbiotic function. Remarkably, these experiments would also serve to explore a possible NfeR1 cross-regulation of alfalfa genes operating in nodule organogenesis, as already shown for *B. japonicum* tRNA-derived small RNA fragments (tRFs) during soybean nodulation [273].

Fine-tuning of the bicistronic mRNA *ntrBC* either by direct *ntrB* targeting or via silencing of putative ribo-repressors places NfeR1 at the core of regulation of the *S. meliloti* NSR. These novel RNA elements described here would guarantee the robust feed-back regulation of the NtrBC two-component system, thereby strengthening the NSR and helping *S. meliloti* to competitively survive the N stress of the rhizosphere environment.

Large RNA networks such as that governed by AbcR1/2 and NfeR1 provide additional levels of regulation relying on mRNA competition for the sRNA [274]. First, computational predictions suggest that the binding affinity between the sRNA and the target mRNAs is a determinant of the hierarchy in the network, i.e., more extensive base-pairing to the sRNA would provide priority for regulation [275]. Second, competing endogenous RNAs (ceRNAs) acting as sRNA

antagonists can further shape regulatory outputs and mediate cross-regulation of the target mRNAs within the network [276]. NfeR1 could be thus regarded as a ceRNA for its *trans*-sRNAs partners that are predicted to silence *ntrBC*. Cross talk between ABC transporter mRNAs via a target mRNA-derived ceRNA has already been demonstrated in the GcvB regulon [88]. Therefore, this is a plausible mechanism for the post-transcriptional control of AbcR1/2 and NfeR1 levels, which likely occurs upon specific metabolic shifts or in endosymbiotic bacteria. Our MAPS dataset can be further inspected to search for such RNA sponges [277].

To conclude, data presented in this Thesis depict a singularly large RNA network that fine-tunes the *S. meliloti* adaptive symbiotic metabolism and is predicted to operate in diverse rhizobial species. Because riboregulation relies on the functional plasticity of the RNA molecules and on modifiable base-pairing interactions, this network could be rewired at different levels, thereby opening yet unexplored avenues for the engineering of highly competitive biofertilizers and symbiotic N-fixation in the sustainable agricultural practices.

Conclusiones

1. El regulador simbiótico de tipo LysR, LsrB, es indispensable para la expresión del RNA AbcR1 (ABC Regulator) en la división activa de *S. meliloti*, mientras que la transcripción inducida por estrés de AbcR2 depende del factor σ alternativo RpoH1.
2. NfeR1 (Nodule Formation Efficiency Regulator) es un RNAs inducido por estrés de nitrógeno que se transcribe a partir de un promotor dual activado por LsrB y reprimido por el regulador maestro de la respuesta al estrés de nitrógeno, NtrC
3. . Las excepcionalmente grandes y superpuestas listas de mRNAs diana de AbcR1, AbcR2 y NfeR1 codifican mayoritariamente proteínas dedicadas a la captación y asimilación de fuentes de carbono y nitrógeno muy diversas, lo que sugiere un papel importante de estos tres sRNAs en la riboregulación generalizada del metabolismo de *S. meliloti* durante la transición simbiótica.
4. AbcR1/2 y NfeR1 regulan principalmente sus mRNA diana mediante un mecanismo de silenciamiento post-transcripcional canónico que implica el apareamiento de bases de sus motivos aSD monocatenarios en la región de inicio de la traducción de las dianas, lo que conduce al bloqueo de la traducción y la posterior degradación del mRNA.
5. AbcR1/2 llevan dos motivos aSD distintos y NfeR1 tres idénticos que actúan de forma independiente y redundante en el apareamiento de bases y regulación del mRNA, respectivamente. Las interacciones sRNA-mRNA son modificables, de modo que AbcR1/2 (y muy probablemente NfeR1) se pueden redireccionar para la regulación de mRNAs adicionales.
6. Las endoribonucleasas de *S. meliloti* RNasa III e YbeY influyen en los niveles de acumulación de grandes listas de mRNAs catalogados como

dianas de AbcR1/2 y NfeR1, anticipando funciones redundantes y específicas de ambas enzimas en el silenciamiento del mRNA.

7. MetK, la sintetasa del principal donante de grupos metilo, S-adenosilmetionina, es una nueva proteína de unión a RNA con una función no canónica en la regulación mediada por los *trans*-ARNs AbcR2 y NfeR1.
8. El silenciamiento masivo pero controlado de los mRNAs metabólicos y de transporte por AbcR1 probablemente permite la utilización jerárquica de los sustratos disponibles para optimizar los flujos metabólicos de la bacteria, lo que confiere a *S. meliloti* una ventaja en la colonización competitiva de nichos nutricionalmente ricos pero extremadamente selectivos, como lo es la rizosfera de las leguminosas, en condiciones ambientales suaves. Los perfiles de expresión y el targetoma de AbcR2 sugieren un impacto similar de éste *trans*-sRNA en la colonización de la raíz bajo estreses abióticos.
9. Bajo estrés de nitrógeno, NfeR1 se une específicamente a mRNAs que codifican proteínas involucradas en la regulación de la absorción y asimilación de nitrógeno combinado, en el control de la progresión del ciclo celular, en la motilidad y en la desnitrificación microaeróbica. El ajuste fino de estos mRNAs mediado por NfeR1 posiblemente afecta a la nodulación, a la diferenciación de bacteroides y/o a la eficiencia de fijación de N, lo que explica los fenotipos simbióticos vinculados a la pérdida de función de NfeR1.
10. NfeR1 es un elemento común de dos bucles de retroalimentación negativa complejos que fortalecen la respuesta al estrés de nitrógeno de *S. meliloti* al aliviar la (auto)represión del sistema de dos componentes NtrBC mediante la regulación de NtrC y probablemente de tres *trans*-RNAs aún

no caracterizados capturados como específicos en el interactoma de NfeR1.

Conclusions

1. The LysR-type symbiotic regulator LsrB is indispensable for the expression of the ABC Regulator AbcR1 sRNA in actively dividing *S. meliloti* bacteria, whereas the stress-induced transcription of AbcR2 depends on the alternative σ factor RpoH1.

2. The Nodule Formation Efficiency Regulator NfeR1 is a nitrogen stress induced sRNA that is transcribed from a dual-mode promoter activated by LsrB and repressed by the master regulator of the nitrogen stress response, NtrC.

3. The AbcR1, AbcR2 and NfeR1 targetomes consist of exceptionally large and overlapping arrays of mRNAs encoding proteins devoted to the uptake and assimilation of widely diverse carbon and nitrogen substrates, suggesting a major role of these three sRNAs in the pervasive riboregulation of *S. meliloti* metabolism during the symbiotic transition.

4. AbcR1/2 and NfeR1 mostly regulate their target mRNAs by a canonical post-transcriptional silencing mechanism involving base-pairing of single-stranded aSD motifs at the translation initiation region, leading to blocking of translation and concomitant mRNA decay.

5. AbcR1/2 carry two distinct, and NfeR1 three identical, aSD motifs that act independently and redundantly in mRNA targeting and regulation, respectively. sRNA-target mRNA base-pairing interactions are modifiable, so that AbcR1/2 (and most probably NfeR1) can be retargeted to regulate non-cognate mRNAs.

6. The *S. meliloti* endoribonucleases RNase III and YbeY influence the steady-state levels of large sets of mRNAs catalogued as AbcR1/2 and

NfeR1 targets, anticipating redundant and specific roles of both enzymes in mRNA silencing.

7. MetK, the synthetase of the major methyl donor S-adenosylmethionine, is a novel RNA-binding protein with a non-canonical function in regulation by the AbcR2 and NfeR1 *trans*-sRNAs.

8. Massive but controlled silencing of transport and metabolic mRNAs by AbcR1 likely enables the hierarchical utilization of available substrates to optimize the metabolic fluxes, thereby conferring *S. meliloti* an advantage for the competitive colonization of nutrient rich but extremely selective niches such as the legume rhizosphere in mild environmental conditions. Expression and targetome profiles suggest a similar impact of AbcR2 in the colonization of the plant environments under abiotic stress.

9. Under nitrogen stress, NfeR1 targets specifically mRNAs encoding proteins involved in the regulation of the uptake and assimilation of combined nitrogen, the control of cell cycle progression, motility, and microaerobic denitrification. NfeR1-mediated fine-tuning of these mRNAs presumably impacts nodulation, bacteroid differentiation and/or N-fixation efficiency, thus explaining the symbiotic phenotypes linked to the NfeR1 loss-of-function.

10. NfeR1 is a common element of two complex negative feedback loops that strengthens the *S. meliloti* nitrogen stress response by alleviating (auto)repression of the NtrBC two-component system by NtrC and probably by three yet uncharacterized *trans*-sRNAs captured as NfeR1 specific partners.

Bibliography

1. Storz G (2002) An expanding universe of noncoding RNAs. Science (1979) 296:1260–1263. https://doi.org/296/5571/1260 [pii] 10.1126/science.1072249
2. Mattick JS, Makunin I v. (2006) Non-coding RNA. Human Molecular Genetics 15:R17–R29. https://doi.org/10.1093/HMG/DDL046
3. Husser C, Dentz N, Ryckelynck M (2021) Structure-Switching RNAs: From Gene Expression Regulation to Small Molecule Detection. Small Structures 2:2000132. https://doi.org/10.1002/SSTR.202000132
4. Cech TR (2012) The RNA Worlds in Context. Cold Spring Harbor Perspectives in Biology 4:a006742. https://doi.org/10.1101/CSHPERSPECT.A006742
5. Sorek R, Cossart P (2010) Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. Nat Rev Genet 11:9–16. https://doi.org/nrg2695 [pii] 10.1038/nrg2695
6. Gottesman S, Storz G (2011) Bacterial small RNA regulators: versatile roles and rapidly evolving variations. Cold Spring Harb Perspect Biol 3:. https://doi.org/10.1101/cshperspect.a003798
7. Serganov A, Nudler E (2013) A Decade of Riboswitches. Cell 152:17–24. https://doi.org/10.1016/j.cell.2012.12.024
8. Mellin JR, Cossart P (2015) Unexpected versatility in bacterial riboswitches. Trends in Genetics 31:150–156. https://doi.org/10.1016/J.TIG.2015.01.005
9. Kolb FA, Engdahl HM, Slagter-Jäger JG, et al (2000) Progression of a loop–loop complex to a four-way junction is crucial for the activity of a regulatory antisense RNA. The EMBO Journal 19:5905–5915. https://doi.org/10.1093/EMBOJ/19.21.5905

10. Gottesman S, Storz G (2011) Bacterial small RNA regulators: versatile roles and rapidly evolving variations. Cold Spring Harb Perspect Biol 3:. https://doi.org/10.1101/cshperspect.a003798
11. Lasa I, Toledo-Arana A, Gingeras TR (2012) An effort to make sense of antisense transcription in bacteria. RNA Biology 9:1039–1044. https://doi.org/10.4161/rna.21167
12. Ruiz de los Mozos I, Vergara-Irigaray M, Segura V, et al (2013) Base Pairing Interaction between 5′- and 3′-UTRs Controls icaR mRNA Translation in Staphylococcus aureus. PLOS Genetics 9:e1004001. https://doi.org/10.1371/JOURNAL.PGEN.1004001
13. Toledo-Arana A, Lasa I (2020) Advances in bacterial transcriptome understanding: From overlapping transcription to the excludon concept. Molecular Microbiology 113:593–602. https://doi.org/10.1111/MMI.14456
14. Menendez-Gil P, Caballero CJ, Catalan-Moreno A, et al (2020) Differential evolution in 3′UTRs leads to specific gene expression in Staphylococcus. Nucleic Acids Research 48:2544–2563. https://doi.org/10.1093/nar/gkaa047
15. Lejars M, Hajnsdorf E (2020) The world of asRNAs in Gram-negative and Gram-positive bacteria. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms 1863:194489. https://doi.org/10.1016/J.BBAGRM.2020.194489
16. Sobrero P, Valverde C (2012) The bacterial protein Hfq: much more than a mere RNA-binding factor. Crit Rev Microbiol 38:276–299. https://doi.org/10.3109/1040841X.2012.664540
17. Saramago M, Bárria C, dos Santos RF, et al (2014) The role of RNases in the regulation of small RNAs. Curr Opin Microbiol 18:105–115. https://doi.org/http://dx.doi.org/10.1016/j.mib.2014.02.009

18. Opdyke JA, Kang JG, Storz G (2004) GadY, a small-RNA regulator of acid response genes in Escherichia coli. Journal of Bacteriology 186:6698–6705. https://doi.org/10.1128/JB.186.20.6698-6705.2004/ASSET/8447D515-5527-415F-A70E-92BCC76FCBA5/ASSETS/GRAPHIC/ZJB0200441140006.JPEG
19. Fröhlich KS, Vogel J (2009) Activation of gene expression by small RNA. Curr Opin Microbiol 12:674–682. https://doi.org/10.1016/j.mib.2009.09.009
20. Sedlyarova N, Shamovsky I, Bharati BK, et al (2016) sRNA-Mediated Control of Transcription Termination in E. coli. Cell 167:111-121.e13. https://doi.org/10.1016/J.CELL.2016.09.004
21. Wagner EGH, Romby P (2015) Small RNAs in bacteria and archaea: who they are, what they do, and how they do it. In: Theodore Friedmann JCD, Stephen FG (eds) Adv Genet. Academic Press, pp 133–208
22. Gage D (2004) Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. MicrobiolMolBiolRev 68:280–300
23. Gibson KE, Kobayashi H, Walker GC (2008) Molecular determinants of a symbiotic chronic infection. Annu Rev Genet 42:413–441. https://doi.org/10.1146/annurev.genet.42.110807.091427
24. Poole P, Ramachandran V, Terpolilli J (2018) Rhizobia: from saprophytes to endosymbionts. Nature Reviews Microbiology 16:291. https://doi.org/10.1038/nrmicro.2017.171
25. Zahran HH (1999) Rhizobium-legume symbiosis and nitrogen fixation under severe conditions and in an arid climate. MicrobiolMolBiolRev 63:968–989

26. Philippot L, Raaijmakers JM, Lemanceau P, van der Putten WH (2013) Going back to the roots: the microbial ecology of the rhizosphere. Nature Reviews Microbiology 11:789–799. https://doi.org/10.1038/nrmicro3109
27. Masson-Boivin C, Giraud E, Perret X, Batut J (2009) Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes? 17:458–466. https://doi.org/10.1016/j.tim.2009.07.004
28. Soto MJ, Domínguez-Ferreras A, Pérez-Mendoza D, et al (2009) Mutualism versus pathogenesis: the give-and-take in plant-bacteria interactions. Cell Microbiol 11:381–388. https://doi.org/10.1111/j.1462-5822.2008.01282.x
29. van de Velde W, Zehirov G, Szatmari A, et al (2010) Plant Peptides Govern Terminal Differentiation of Bacteria in Symbiosis. Science (1979) 327:1122–1126. https://doi.org/10.1126/science.1184057
30. Galibert F, Finan TM, Long SR, et al (2001) The composite genome of the legume symbiont Sinorhizobium meliloti. Science (1979) 293:668–672. https://doi.org/10.1126/SCIENCE.1060966
31. Brilli M, Fondi M, Fani R, et al (2010) The diversity and evolution of cell cycle regulation in alpha-proteobacteria: a comparative genomic analysis. BMC Syst Biol 4:52. https://doi.org/10.1186/1752-0509-4-52
32. Yurgel SN, Rice J, Mulder M, Kahn ML (2010) GlnB/GlnK PII proteins and regulation of the Sinorhizobium meliloti Rm1021 nitrogen stress response and symbiotic function. J Bacteriol 192:2473–2481. https://doi.org/10.1128/jb.01657-09
33. Krol E, Blom J, Winnebald J, et al (2011) RhizoRegNet—A database of rhizobial transcription factors and regulatory networks. Journal of Biotechnology 155:127–134. https://doi.org/10.1016/j.jbiotec.2010.11.004

34. Galardini M, Brilli M, Spini G, et al (2015) Evolution of intra-specific regulatory networks in a multipartite bacterial genome. PLoS Computational Biology 11:e1004478–e1004478. https://doi.org/10.1371/journal.pcbi.1004478
35. Lang C, Barnett MJ, Fisher RF, et al (2018) Most Sinorhizobium meliloti extracytoplasmic function sigma factors control accessory functions. mSphere 3:e00454-18. https://doi.org/10.1128/mspheredirect.00454-18
36. Jiménez-Zurdo JIJI, Valverde C, Becker A (2013) Insights into the noncoding RNome of nitrogen-fixing endosymbiotic α-proteobacteria. Mol Plant-Microbe Interact 26:160–167. https://doi.org/10.1094/mpmi-07-12-0186-cr
37. Becker A, Overlöper A, Schlüter J-P, et al (2014) Riboregulation in plant-associated α-proteobacteria. RNA Biol 11:550–562. https://doi.org/10.4161/rna.29625
38. Jiménez-Zurdo JIJI, Robledo M (2015) Unraveling the universe of small RNA regulators in the legume symbiont Sinorhizobium meliloti. Symbiosis 67:43–54. https://doi.org/10.1007/s13199-015-0345-z
39. Capela D, Barloy-Hubler F, Gouzy J, et al (2001) Analysis of the chromosome sequence of the legume symbiont Sinorhizobium meliloti strain 1021. Proc Natl Acad Sci U S A 98:9877–9882. https://doi.org/10.1073/pnas.161294398
40. del Val C, Rivas E, Torres-Quesada O, et al (2007) Identification of differentially expressed small non-coding RNAs in the legume endosymbiont Sinorhizobium meliloti by comparative genomics. Mol Microbiol 66:1080–1091. https://doi.org/MMI5978 [pii] 10.1111/j.1365-2958.2007.05978.x
41. Ulve VM, Sevin EW, Cheron A, Barloy-Hubler F (2007) Identification of chromosomal alpha-proteobacterial small RNAs by comparative genome

analysis and detection in Sinorhizobium meliloti strain 1021. BMC Genomics 8:467. https://doi.org/1471-2164-8-467 [pii] 10.1186/1471-2164-8-467

42. Valverde C, Livny J, Schluter JP, et al (2008) Prediction of Sinorhizobium meliloti sRNA genes and experimental detection in strain 2011. BMC Genomics 9:416. https://doi.org/10.1186/1471-2164-9-416
43. Sharma CM, Vogel J (2014) Differential RNA-seq: the approach behind and the biological insight gained. Curr Opin Microbiol 19:97–105. https://doi.org/http://dx.doi.org/10.1016/j.mib.2014.06.010
44. Ettwiller L, Buswell J, Yigit E, Schildkraut I (2016) A novel enrichment strategy reveals unprecedented number of novel transcription start sites at single base resolution in a model prokaryote and the gut microbiome. BMC Genomics 17:199. https://doi.org/10.1186/s12864-016-2539-z
45. Schlüter JP, Reinkensmeier J, Daschkey S, et al (2010) A genome-wide survey of sRNAs in the symbiotic nitrogen-fixing alpha-proteobacterium Sinorhizobium meliloti. BMC Genomics 11:245. https://doi.org/1471-2164-11-245 [pii] 10.1186/1471-2164-11-245
46. Schlüter JP, Reinkensmeier J, Barnett MJ, et al (2013) Global mapping of transcription start sites and promoter motifs in the symbiotic alpha-proteobacterium Sinorhizobium meliloti 1021. BMC Genomics 14:156. https://doi.org/10.1186/1471-2164-14-156
47. Sallet E, Roux B, Sauviac L, et al (2013) Next-generation annotation of prokaryotic genomes with EuGene-P: application to Sinorhizobium meliloti 2011. DNA Res 20:339–354. https://doi.org/10.1093/dnares/dst014
48. Sittka A, Lucchini S, Papenfort K, et al (2008) Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-

transcriptional regulator, Hfq. PLoS Genet 4:e1000163–e1000163. https://doi.org/10.1371/journal.pgen.1000163

49. Sittka A, Sharma CM, Rolle K, 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 6:266–275. https://doi.org/10.4161/rna.6.3.8332
50. Torres-Quesada O, Reinkensmeier J, Schluter JP, et al (2014) Genome-wide profiling of Hfq-binding RNAs uncovers extensive post-transcriptional rewiring of major stress response and symbiotic regulons in Sinorhizobium meliloti. RNA Biology 11:563. https://doi.org/10.4161/RNA.28239
51. Vercruysse M, Fauvart M, Cloots L, et al (2010) Genome-wide detection of predicted non-coding RNAs in Rhizobium etli expressed during free-living and host-associated growth using a high-resolution tiling array. BMC Genomics 11:53. https://doi.org/10.1186/1471-2164-11-53
52. Madhugiri R, Pessi G, Voss B, et al (2012) Small RNAs of the Bradyrhizobium/Rhodopseudomonas lineage and their analysis. RNA Biol 9:47–58. https://doi.org/10.4161/rna.9.1.18008
53. Čuklina J, Hahn J, Imakaev M, et al (2016) Genome-wide transcription start site mapping of Bradyrhizobium japonicum grown free-living or in symbiosis - a rich resource to identify new transcripts, proteins and to study gene regulation. BMC Genomics 17:302. https://doi.org/10.1186/s12864-016-2602-9
54. Roux B, Rodde N, Jardinaud M-F, et al (2014) An integrated analysis of plant and bacterial gene expression in symbiotic root nodules using laser-capture microdissection coupled to RNA sequencing. Plant J 77:817–837. https://doi.org/10.1111/tpj.12442

55. Kalvari I, Argasinska J, Quinones-Olvera N, et al (2017) Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families. Nucleic Acids Res 46:D335–D342. https://doi.org/10.1093/nar/gkx1038
56. Guerrier-Takada C, Gardiner K, Marsh T, et al (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. Cell 35:849–857. https://doi.org/https://doi.org/10.1016/0092-8674(83)90117-4
57. Keiler KC (2008) Biology of trans-Translation. Annual Review of Microbiology 62:133–151. https://doi.org/10.1146/annurev.micro.62.081307.162948
58. Ebeling S, Kündig C, Hennecke H (1991) Discovery of a rhizobial RNA that is essential for symbiotic root nodule development. J Bacteriol 173:6373–6382
59. Ulve VM, Cheron A, Trautwetter A, et al (2007) Characterization and expression patterns of Sinorhizobium meliloti tmRNA (ssrA). FEMS Microbiol Lett 269:117–123. https://doi.org/10.1111/j.1574-6968.2006.00616.x
60. Keenan RJ, Freymann DM, Stroud RM, Walter P (2001) The signal recognition particle. Annual Review of Biochemistry 70:755–775. https://doi.org/10.1146/annurev.biochem.70.1.755
61. Wassarman KM (2018) 6S RNA, a Global Regulator of Transcription. Microbiology Spectrum 6:10.1128/microbiolspec.RWR–0019–2018. https://doi.org/10.1128/microbiolspec.RWR-0019-2018
62. Reinkensmeier J, Schlüter J-P, Giegerich R, Becker A (2011) Conservation and Occurrence of Trans-Encoded sRNAs in the Rhizobiales. Genes (Basel) 2:925–956
63. del Val C, Romero-Zaliz R, Torres-Quesada O, et al (2012) A survey of sRNA families in alpha-proteobacteria. RNA Biol 9:119–129. https://doi.org/10.4161/rna.18643

64. Reinkensmeier J, Giegerich R (2015) Thermodynamic matchers for the construction of the cuckoo RNA family. RNA Biol 12:197–207. https://doi.org/10.1080/15476286.2015.1017206
65. Lagares Jr A, Roux I, Valverde C (2016) Phylogenetic distribution and evolutionary pattern of an α-proteobacterial small RNA gene that controls polyhydroxybutyrate accumulation in Sinorhizobium meliloti. Molecular Phylogenetics and Evolution 99:182–193. https://doi.org/http://dx.doi.org/10.1016/j.ympev.2016.03.026
66. Robledo M, Jiménez-Zurdo JIJI, Becker A (2015) Antisense transcription of symbiotic genes in Sinorhizobium meliloti. Symbiosis 67:55–67. https://doi.org/10.1007/s13199-015-0358-7
67. J G, WR H (2018) Widespread Antisense Transcription in Prokaryotes. In: Regulating with RNA in Bacteria and Archaea. American Society of Microbiology, pp 191–210
68. MacLellan SR, Smallbone LA, Sibley CD, Finan TM (2005) The expression of a novel antisense gene mediates incompatibility within the large repABC family of α-proteobacterial plasmids. Mol Microbiol 55:611–623. https://doi.org/10.1111/j.1365-2958.2004.04412.x
69. Robledo M, García-Tomsig NINI, Jiménez-Zurdo JIJI (2018) Primary Characterization of Small RNAs in Symbiotic Nitrogen-Fixing Bacteria. In: Medina C, López-Baena FJ (eds) Host-Pathogen Interactions: Methods and Protocols. Springer New York, New York, NY, pp 277–295
70. Wright PR, Richter AS, Papenfort K, et al (2013) Comparative genomics boosts target prediction for bacterial small RNAs. Proc Natl Acad Sci USA 110:E3487–E3496. https://doi.org/10.1073/pnas.1303248110
71. Wright PR, Georg J, Mann M, et al (2014) CopraRNA and IntaRNA: predicting small RNA targets, networks and interaction domains. Nucleic Acids Res 42:W119–W123. https://doi.org/10.1093/nar/gku359

72. McIntosh M, Meyer S, Becker A (2009) Novel Sinorhizobium meliloti quorum sensing positive and negative regulatory feedback mechanisms respond to phosphate availability. Mol Microbiol 74:1238–1256. https://doi.org/10.1111/j.1365-2958.2009.06930.x
73. Robledo M, Frage B, Wright PR, Becker A (2015) A stress-induced small RNA modulates alpha-rhizobial cell cycle progression. PLoS Genet 11:e1005153–e1005153. https://doi.org/10.1371/journal.pgen.1005153
74. Sharma CM, Vogel J (2009) Experimental approaches for the discovery and characterization of regulatory small RNA. Curr Opin Microbiol 12:536–546. https://doi.org/S1369-5274(09)00111-8 [pii] 10.1016/j.mib.2009.07.006
75. Lalaouna D, Massé E (2015) Identification of sRNA interacting with a transcript of interest using MS2-affinity purification coupled with RNA sequencing (MAPS) technology. Genomics Data 5:136–138. https://doi.org/10.1016/j.gdata.2015.05.033
76. Jonas K (2014) To divide or not to divide: control of the bacterial cell cycle by environmental cues. Curr Opin Microbiol 18:54–60. https://doi.org/http://dx.doi.org/10.1016/j.mib.2014.02.006
77. Penterman J, Abo RP, de Nisco NJ, et al (2014) Host plant peptides elicit a transcriptional response to control the Sinorhizobium meliloti cell cycle during symbiosis. Proc Natl Acad Sci U S A 111:3561–3566. https://doi.org/10.1073/pnas.1400450111
78. Robledo M, Schlüter J-P, Loehr LO, et al (2018) An sRNA and cold shock protein homolog-based feedforward loop post-transcriptionally controls cell cycle master regulator CtrA. Frontiers in Microbiology 9:763. https://doi.org/10.3389/fmicb.2018.00763
79. Hoang HH, Becker A, Gonzalez JE (2004) The LuxR homolog ExpR, in combination with the Sin quorum sensing system, plays a central role in

Sinorhizobium meliloti gene expression. J Bacteriol 186:5460–5472. https://doi.org/10.1128/JB.186.16.5460-5472.2004

80. Hoang HH, Gurich N, Gonzalez JE (2008) Regulation of motility by the ExpR/Sin quorum-sensing system in Sinorhizobium meliloti. J Bacteriol 190:861–871. https://doi.org/10.1128/JB.01310-07
81. Baumgardt K, Charoenpanich P, McIntosh M, et al (2014) RNase E Affects the Expression of the Acyl-Homoserine Lactone Synthase Gene sinI in Sinorhizobium meliloti. J Bacteriol 196:1435–1447. https://doi.org/10.1128/JB.01471-13
82. Baumgardt K, Šmídová K, Rahn H, et al (2016) The stress-related, rhizobial small RNA RcsR1 destabilizes the autoinducer synthase encoding mRNA sinI in Sinorhizobium meliloti. RNA Biology 13:486–499. https://doi.org/10.1080/15476286.2015.1110673
83. Melior H, Li S, Madhugiri R, et al (2019) Transcription attenuation-derived small RNA rnTrpL regulates tryptophan biosynthesis gene expression in trans. Nucleic Acids Research 47:6396–6410. https://doi.org/10.1093/nar/gkz274
84. Higgins CF (2001) ABC transporters: physiology, structure and mechanism – an overview. Research in Microbiology 152:205–210. https://doi.org/10.1016/S0923-2508(01)01193-7
85. Mauchline TH, Fowler JE, East AK, et al (2006) Mapping the Sinorhizobium meliloti 1021 solute-binding protein-dependent transportome. Proc Natl Acad Sci USA 103:17933–17938
86. Nogales J, Muñoz S, Muñoz M, et al (2009) Genetic characterization of oligopeptide uptake systems in Sinorhizobium meliloti. FEMS Microbiology Letters 293:177–187. https://doi.org/10.1111/J.1574-6968.2009.01527.X

87. Djordjevic MA (2004) Sinorhizobium meliloti metabolism in the root nodule: A proteomic perspective. PROTEOMICS 4:1859–1872. https://doi.org/10.1002/PMIC.200300802

88. Miyakoshi M, Chao Y, Vogel J (2015) Cross talk between ABC transporter mRNAs via a target mRNA-derived sponge of the GcvB small RNA. EMBO J 34:1478–1492. https://doi.org/10.15252/embj.201490546

89. Wilms I, Voss B, Hess WR, et al (2011) Small RNA-mediated control of the Agrobacterium tumefaciens GABA binding protein. Mol Microbiol 80 (2):492–506. https://doi.org/10.1111/j.1365-2958.2011.07589.x

90. Caswell CC, Gaines JM, Ciborowski P, et al (2012) Identification of two small regulatory RNAs linked to virulence in Brucella abortus 2308. Mol Microbiol 85:345–360. https://doi.org/10.1111/j.1365-2958.2012.08117.x

91. Torres-Quesada O, Millán V, Nisa-Martínez R, et al (2013) Independent activity of the homologous small regulatory RNAs AbcR1 and AbcR2 in the legume symbiont Sinorhizobium meliloti. PLoS ONE 8:1080–1091. https://doi.org/10.1371/journal.pone.0068147

92. Overlöper A, Kraus A, Gurski R, et al (2014) Two separate modules of the conserved regulatory RNA AbcR1 address multiple target mRNAs in and outside of the translation initiation region. RNA Biology 11:624–640. https://doi.org/10.4161/rna.29145

93. Sheehan LM, Caswell CC (2018) An account of evolutionary specialization: the AbcR small RNAs in the Rhizobiales. Mol Microbiol 107:24–33. https://doi.org/10.1111/mmi.13869

94. Lagares Jr. A, Ceizel Borella G, Linne U, et al (2017) Regulation of polyhydroxybutyrate accumulation in Sinorhizobium meliloti by the trans-encoded small RNA MmgR. J Bacteriol 199:e00776–16. https://doi.org/10.1128/JB.00776-16

95. Ceizel Borella G, Lagares Jr. A, Valverde C (2018) Expression of the small regulatory RNA gene mmgR is regulated negatively by AniA and positively by NtrC in Sinorhizobium meliloti 2011. Microbiology (N Y) 164:88–98. https://doi.org/doi:10.1099/mic.0.000586
96. Barnett MJ, Toman CJ, Fisher RF, Long SR (2004) A dual-genome Symbiosis Chip for coordinate study of signal exchange and development in a prokaryote-host interaction. Proc Natl Acad Sci U S A 101:16636–16641. https://doi.org/10.1073/pnas.0407269101
97. Robledo M, Peregrina A, Millán V, et al (2017) A conserved α-proteobacterial small RNA contributes to osmoadaptation and symbiotic efficiency of rhizobia on legume roots. Environmental Microbiology 19:2661–2680. https://doi.org/doi:10.1111/1462-2920.13757
98. de Lay N, Schu DJ, Gottesman S (2013) Bacterial small RNA-based negative regulation: Hfq and its accomplices. J Biol Chem 288:7996–8003. https://doi.org/10.1074/jbc.R112.441386
99. Jiménez-Zurdo JI, Robledo M (2017) RNA silencing in plant symbiotic bacteria: Insights from a protein-centric view. RNA Biology 14:1672–1677. https://doi.org/10.1080/15476286.2017.1356565
100. Holmqvist E, Vogel J (2018) RNA-binding proteins in bacteria. Nature Reviews Microbiology 601–615. https://doi.org/10.1038/s41579-018-0049-5
101. Quendera AP, Seixas AF, dos Santos RF, et al (2020) RNA-Binding Proteins Driving the Regulatory Activity of Small Non-coding RNAs in Bacteria. Frontiers in Molecular Biosciences 7:78. https://doi.org/10.3389/FMOLB.2020.00078/BIBTEX
102. Arraiano CM, Andrade JM, Domingues S, et al (2010) The critical role of RNA processing and degradation in the control of gene expression. FEMS

Microbiol Rev 34:883–923. https://doi.org/10.1111/j.1574-6976.2010.00242.x

103. Silva IJ, Saramago M, Dressaire C, et al (2011) Importance and key events of prokaryotic RNA decay: the ultimate fate of an RNA molecule. Wiley Interdisciplinary Reviews: RNA 2:818–836. https://doi.org/doi:10.1002/wrna.94
104. Lalaouna D, Simoneau-Roy M, Lafontaine D, Massé E (2013) Regulatory RNAs and target mRNA decay in prokaryotes. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms 1829:742–747. https://doi.org/http://dx.doi.org/10.1016/j.bbagrm.2013.02.013
105. Becker A, Barnett MJ, Capela D, et al (2009) A portal for rhizobial genomes: RhizoGATE integrates a Sinorhizobium meliloti genome annotation update with postgenome data. J Biotechnol 140:45–50. https://doi.org/10.1016/j.jbiotec.2008.11.006
106. Voss B, Holscher M, Baumgarth B, et al (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 390:331–336. https://doi.org/S0006-291X(09)01942-1 [pii] 10.1016/j.bbrc.2009.09.125
107. Zhang Y, Hong G (2009) Post-transcriptional regulation of NifA expression by Hfq and RNase E complex in Rhizobium leguminosarum bv. viciae. Acta Biochim Biophys Sin (Shanghai) 41:719–730
108. Saramago M, Robledo M, Matos RGRG, et al (2018) Sinorhizobium meliloti RNase III: Catalytic Features and Impact on Symbiosis. Frontiers in Genetics 9:. https://doi.org/10.3389/fgene.2018.00350
109. Gil R, Silva FJ, Peretó J, Moya A (2004) Determination of the core of a minimal bacterial gene set. Microbiol Mol Biol Rev 68:518–537. https://doi.org/10.1128/MMBR.68.3.518-537.2004

110. Jacob AI, Köhrer C, Davies BW, et al (2013) Conserved bacterial RNase YbeY plays key roles in 70S ribosome quality control and 16S rRNA maturation. Mol Cell 49:427–438. https://doi.org/10.1016/j.molcel.2012.11.025
111. Saramago M, Peregrina A, Robledo M, et al (2017) Sinorhizobium meliloti YbeY is an endoribonuclease with unprecedented catalytic features, acting as silencing enzyme in riboregulation. Nucleic Acids Res 45:1371–1391. https://doi.org/10.1093/nar/gkw1234
112. Babu VMP, Sankari S, Budnick JA, et al (2019) Sinorhizobium meliloti YbeY is a zinc-dependent single-strand specific endoribonuclease that plays an important role in 16S ribosomal RNA processing. Nucleic Acids Res. https://doi.org/10.1093/nar/gkz1095
113. Davies BW, Walker GC (2008) A highly conserved protein of unknown function is required by Sinorhizobium meliloti for symbiosis and environmental stress protection. J Bacteriol 190:1118–1123
114. Caballero CJ, Menendez-Gil P, Catalan-Moreno A, et al (2018) The regulon of the RNA chaperone CspA and its auto-regulation in Staphylococcus aureus. Nucleic Acids Res 46:1345–1361. https://doi.org/10.1093/nar/gkx1284
115. Wagner EG (2013) Cycling of RNAs on Hfq. RNA Biol 10:619–626. https://doi.org/10.4161/rna.24044
116. Smirnov A, Förstner KU, Holmqvist E, et al (2016) Grad-seq guides the discovery of ProQ as a major small RNA-binding protein. Proc Natl Acad Sci USA 113:11591–11596. https://doi.org/10.1073/pnas.1609981113
117. Vogel J, Luisi BF (2011) Hfq and its constellation of RNA. Nat Rev Microbiol 9:578–589. https://doi.org/10.1038/nrmicro2615
118. Gao M, Benge A, Mesa JM, et al (2018) Use of RNA Immunoprecipitation Method for Determining Sinorhizobium meliloti RNA-Hfq Protein

Associations In Vivo. Biological Procedures Online 20:8. https://doi.org/10.1186/s12575-018-0075-8

119. Robledo M, García-Tomsig NI, Matia-González AM, et al (2021) Synthetase of the methyl donor S-adenosylmethionine from nitrogen-fixing α-rhizobia can bind functionally diverse RNA species. RNA Biology 18:1111–1123. https://doi.org/10.1080/15476286.2020.1829365
120. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning. A laboratory Manual. Cold Spring Harbor Laboratory Press Cold Sprin:
121. Beringer JE (1974) R factor transfer in Rhizobium leguminosarum. J Gen Microbiol 84:188–198
122. Robertsen BK, Aman P, Darvill AG, et al (1981) Host-symbiont interactions : V. the structure of acidic extracellular polysaccharides secreted by Rhizobium leguminosarum and Rhizobium trifolii. Plant Physiol 67:389–400
123. Ditta G, Virts E, Palomares A, Kim CH (1987) The nifA gene of Rhizobium meliloti is oxygen regulated. J Bacteriol 169:3217–3223
124. Foussard M, Garnerone AM, Ni F, et al (1997) Negative autoregulation of the Rhizobium meliloti fixK gene is indirect and requires a newly identified regulator, FixT. Molecular Microbiology 25:27–37. https://doi.org/10.1046/j.1365-2958.1997.4501814.x
125. Meade HM, Long SR, Ruvkun GB, et al (1982) Physical and genetic characterization of symbiotic and auxotrophic mutants of Rhizobium meliloti induced by transposon Tn5 mutagenesis. J Bacteriol 149:114–122
126. Bahlawane C, McIntosh M, Krol E, Becker A (2008) Sinorhizobium meliloti regulator MucR couples exopolysaccharide synthesis and motility. Mol Plant Microbe Interact 21:1498–1509. https://doi.org/10.1094/MPMI-21-11-1498

127. Simon R, Priefer U, Puhler A (1983) A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in gram negative bacteria. Nat Biotech 1:784–791
128. Schafer A, Tauch A, Jager W, et al (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 145:69–73
129. Studier FW (1991) Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. Journal of Molecular Biology 219:37–44
130. Khan SR, Gaines J, Roop RM, Farrand SK (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 74:5053–5062. https://doi.org/10.1128/AEM.01098-08
131. Döhlemann J, Wagner M, Happel C, et al (2017) A Family of Single Copy repABC-Type Shuttle Vectors Stably Maintained in the Alpha-Proteobacterium Sinorhizobium meliloti. ACS Synthetic Biology 6:968–984. https://doi.org/10.1021/ACSSYNBIO.6B00320/SUPPL_FILE/SB6B00320_SI_001.PDF
132. Cabanes D, Boistard P, Batut J (2000) Identification of Sinorhizobium meliloti genes regulated during symbiosis. J Bacteriol 182:3632–3637
133. Krol E, Becker A (2004) Global transcriptional analysis of the phosphate starvation response in Sinorhizobium meliloti strains 1021 and 2011. Mol Genet Genomics 272:1–17. https://doi.org/10.1007/s00438-004-1030-8
134. Lalaouna D, Prévost K, Eyraud A, Massé E (2017) Identification of unknown RNA partners using MAPS. Methods 117:28–34. https://doi.org/10.1016/j.ymeth.2016.11.011

135. Carrier MC, Laliberté G, Massé E (2018) Identification of new bacterial small RNA targets using MS2 affinity purification coupled to RNA sequencing. In: Methods in Molecular Biology. Humana Press Inc., pp 77–88

136. García-Tomsig NI, Robledo M, diCenzo GC, et al (2022) Pervasive RNA Regulation of Metabolism Enhances the Root Colonization Ability of Nitrogen-Fixing Symbiotic α-Rhizobia. mBio 13:. https://doi.org/10.1128/MBIO.03576-21

137. LeCuyer KA, Behlen LS, Uhlenbeck OC (1995) Mutants of the Bacteriophage MS2 Coat Protein That Alter Its Cooperative Binding to RNA. Biochemistry 34:10600–10606. https://doi.org/10.1021/bi00033a035

138. Wei Z, Zhang W, Fang H, et al (2018) EsATAC: An easy-to-use systematic pipeline for ATAC-seq data analysis. Bioinformatics 34:2664–2665. https://doi.org/10.1093/bioinformatics/bty141

139. Liao Y, Smyth GK, Shi W (2019) The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. Nucleic Acids Research 47:e47–e47. https://doi.org/10.1093/nar/gkz114

140. Robinson MD, Oshlack A (2010) A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biology 11:1–9. https://doi.org/10.1186/gb-2010-11-3-r25

141. Robinson JT, Thorvaldsdóttir H, Winckler W, et al (2011) Integrative genomics viewer. Nature Biotechnology 29:24–26

142. Larkin MA, Blackshields G, Brown NP, et al (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23:2947–2948. https://doi.org/10.1093/bioinformatics/btm404

143. Bailey TL, Boden M, Buske FA, et al (2009) MEME Suite: Tools for motif discovery and searching. Nucleic Acids Research 37:doi: 10.1093/nar/gkp335. https://doi.org/10.1093/nar/gkp335
144. Mann M, Wright PR, Backofen R (2017) IntaRNA 2.0: Enhanced and customizable prediction of RNA-RNA interactions. Nucleic Acids Research 45:W435–W439. https://doi.org/10.1093/nar/gkx279
145. Sheehan LM, Caswell CC (2017) A 6-nucleotide regulatory motif within the AbcR small RNAs of Brucella abortus mediates host-pathogen interactions. mBio 8:e00473–17. https://doi.org/10.1128/mBio.00473-17
146. Sheehan LM, Budnick JA, Blanchard C, et al (2015) A LysR-family transcriptional regulator required for virulence in Brucella abortus is highly conserved among the α-proteobacteria. Mol Microbiol 98:318–328. https://doi.org/10.1111/mmi.13123
147. Eisfeld J, Kraus A, Ronge C, et al (2021) A LysR-type transcriptional regulator controls the expression of numerous small RNAs in Agrobacterium tumefaciens. Molecular Microbiology 116:126–139. https://doi.org/10.1111/mmi.14695
148. Budnick JA, Sheehan LM, Ginder MJ, et al (2020) A central role for the transcriptional regulator VtlR in small RNA-mediated gene regulation in Agrobacterium tumefaciens. Scientific Reports 10:14968. https://doi.org/10.1038/s41598-020-72117-0
149. Barnett MJ, Bittner AN, Toman CJ, et al (2012) Dual RpoH sigma factors and transcriptional plasticity in a symbiotic bacterium. J Bacteriol 194:4983–4994. https://doi.org/10.1128/JB.00449-12
150. diCenzo GC, Tesi M, Pfau T, et al (2020) Genome-scale metabolic reconstruction of the symbiosis between a leguminous plant and a nitrogen-fixing bacterium. Nature Communications 11:2574. https://doi.org/10.1038/s41467-020-16484-2

151. Tang G, Xing S, Wang S, et al (2017) Regulation of cysteine residues in LsrB proteins from Sinorhizobium meliloti under free-living and symbiotic oxidative stress. Environmental Microbiology 19:5130–5145. https://doi.org/10.1111/1462-2920.13992
152. Sen A, Imlay JA (2021) How microbes defend themselves from incoming hydrogen peroxide. Frontiers in Immunology 12:667343. https://doi.org/10.3389/fimmu.2021.667343
153. Tang G, Lu D, Wang D, Luo L (2013) Sinorhizobium meliloti lsrB is involved in alfalfa root nodule development and nitrogen-fixing bacteroid differentiation. Chinese Science Bulletin 58:4077–4083. https://doi.org/10.1007/s11434-013-5960-6
154. Lu D, Tang G, Wang D, Luo L (2013) The Sinorhizobium meliloti LysR family transcriptional factor LsrB is involved in regulation of glutathione biosynthesis. Acta Biochimica et Biophysica Sinica 45:882–888. https://doi.org/10.1093/abbs/gmt083
155. Tang G, Wang Y, Luo L (2014) Transcriptional regulator LsrB of Sinorhizobium meliloti positively regulates the expression of genes involved in lipopolysaccharide biosynthesis. Appl Environ Microbiol 80:5265–5273. https://doi.org/10.1128/aem.01393-14
156. Mitsui H, Sato T, Sato Y, et al (2004) Sinorhizobium meliloti RpoH1 is required for effective nitrogen-fixing symbiosis with alfalfa. Molecular Genetics and Genomics 271:416–425. https://doi.org/10.1007/s00438-004-0992-x
157. Luo L, Yao SY, Becker A, et al (2005) Two new Sinorhizobium meliloti LysR-type transcriptional regulators required for nodulation. J Bacteriol 187:4562–4572
158. Mika F, Hengge R (2014) Small RNAs in the control of RpoS, CsgD, and biofilm architecture of Escherichia coli. RNA Biology 11:494–507

159. Carrier MC, Lalaouna D, Massé E (2016) A game of tag: MAPS catches up on RNA interactomes. RNA Biology 13:473–476. https://doi.org/10.1080/15476286.2016.1156830

160. Sharma CM, Papenfort K, Pernitzsch SR, et al (2011) Pervasive post-transcriptional control of genes involved in amino acid metabolism by the Hfq-dependent GcvB small RNA. Mol Microbiol 81:1144–1165. https://doi.org/10.1111/j.1365-2958.2011.07751.x

161. Melamed S, Peer A, Faigenbaum-Romm R, et al (2016) Global mapping of small RNA-target interactions in bacteria. Molecular Cell 63:884–897. https://doi.org/10.1016/j.molcel.2016.07.026

162. Torres-Quesada O, Oruezabal RII, Peregrina A, et al (2010) The Sinorhizobium meliloti RNA chaperone Hfq influences central carbon metabolism and the symbiotic interaction with alfalfa. BMC Microbiol 10:71. https://doi.org/1471-2180-10-71 [pii] 10.1186/1471-2180-10-71

163. Adams PP, Storz G (2020) Prevalence of small base-pairing RNAs derived from diverse genomic loci. Biochimica et Biophysica Acta - Gene Regulatory Mechanisms 1863:194524

164. Povolo S, Casella S (2000) A critical role for aniA in energy-carbon flux and symbiotic nitrogen fixation in Sinorhizobium meliloti. Archives of Microbiology 174:42–49. https://doi.org/10.1007/s002030000171

165. Eberlein C, Baumgarten T, Starke S, Heipieper HJ (2018) Immediate response mechanisms of Gram-negative solvent-tolerant bacteria to cope with environmental stress: cis-trans isomerization of unsaturated fatty acids and outer membrane vesicle secretion. Applied Microbiology and Biotechnology 102:2583–2593

166. Delhaye A, Collet JF, Laloux G (2019) A fly on the wall: how stress response systems can sense and respond to damage to peptidoglycan. Frontiers in Cellular and Infection Microbiology 9:380

167. Brenes-Álvarez M, Vioque A, Muro-Pastor AM (2020) The integrity of the cell wall and its remodeling during heterocyst differentiation are regulated by phylogenetically conserved small RNA Yfr1 in Nostoc sp. Strain PCC 7120. mBio 11:e02599-19. https://doi.org/10.1128/mBio.02599-19
168. Patriarca EJ, Tatè R, Iaccarino M (2002) Key Role of Bacterial NH(4)(+) Metabolism in Rhizobium-Plant Symbiosis. Microbiology and Molecular Biology Reviews 66:203–222. https://doi.org/10.1128/MMBR.66.2.203-222.2002
169. Ruiz B, le Scornet A, Sauviac L, et al (2019) The nitrate assimilatory pathway in Sinorhizobium meliloti: Contribution to NO production. Frontiers in Microbiology 10:1526. https://doi.org/10.3389/fmicb.2019.01526
170. Muro-Pastor AM, Hess WR (2020) Regulatory RNA at the crossroads of carbon and nitrogen metabolism in photosynthetic cyanobacteria. Biochimica et Biophysica Acta - Gene Regulatory Mechanisms 1863:194477
171. Prell J, Bourdes A, Kumar S, et al (2010) Role of symbiotic auxotrophy in the Rhizobium-legume symbioses. PLoS One 5:e13933. https://doi.org/10.1371/journal.pone.0013933
172. Udvardi M, Poole PS (2013) Transport and metabolism in legume-rhizobia symbioses. Annu Rev Plant Biol 64:781–805. https://doi.org/doi:10.1146/annurev-arplant-050312-120235
173. Tkacz A, Bestion E, Bo Z, et al (2020) Influence of plant fraction, soil, and plant species on microbiota: A multikingdom comparison. mBio 11:1–17. https://doi.org/10.1128/mBio.02785-19
174. Fagorzi C, Bacci G, Huang R, et al (2021) Nonadditive transcriptomic signatures of genotype-by-genotype interactions during the initiation of

plant-Rhizobium symbiosis. mSystems 6:e00974-20. https://doi.org/10.1128/msystems.00974-20

175. Wheatley RM, Ford BL, Li L, et al (2020) Lifestyle adaptations of Rhizobium from rhizosphere to symbiosis. Proc Natl Acad Sci U S A 117:23823–23834. https://doi.org/10.1073/PNAS.2009094117/-/DCSUPPLEMENTAL
176. Gosai J, Anandhan S, Bhattacharjee A, Archana G (2020) Elucidation of quorum sensing components and their role in regulation of symbiotically important traits in Ensifer nodulating pigeon pea. Microbiological Research 231:126354. https://doi.org/10.1016/j.micres.2019.126354
177. Keating SM, Bornstein BJ, Finney A, Hucka M (2006) SBMLToolbox: An SBML toolbox for MATLAB users. Bioinformatics 22:1275–1277. https://doi.org/10.1093/bioinformatics/btl111
178. Bornstein BJ, Keating SM, Jouraku A, Hucka M (2008) LibSBML: An API library for SBML. Bioinformatics 24:880–881. https://doi.org/10.1093/bioinformatics/btn051
179. Heirendt L, Arreckx S, Pfau T, et al (2019) Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. Nature Protocols 14:639–702. https://doi.org/10.1038/s41596-018-0098-2
180. DiCenzo GC, Checcucci A, Bazzicalupo M, et al (2016) Metabolic modelling reveals the specialization of secondary replicons for niche adaptation in Sinorhizobium meliloti. Nature Communications 7:10.1038 & #x2F;ncomms12219-10.1038 & #x2F;ncomms12219. https://doi.org/10.1038/ncomms12219
181. Vallenet D, Calteau A, Cruveiller S, et al (2017) MicroScope in 2017: An expanding and evolving integrated resource for community expertise of microbial genomes. Nucleic Acids Research 45:D517–D528. https://doi.org/10.1093/nar/gkw1101

182. Nelson M, Guhlin J, Epstein B, et al (2018) The complete replicons of 16 Ensifer meliloti strains offer insights into intra- and inter-replicon gene transfer, transposon-associated loci, and repeat elements. Microbial Genomics 4:10.1099/mgen.0.000174. https://doi.org/10.1099/mgen.0.000174

183. Rivas E, Clements J, Eddy SR (2017) A statistical test for conserved RNA structure shows lack of evidence for structure in lncRNAs. Nat Meth 14:45–48. https://doi.org/10.1038/nmeth.4066 http://www.nature.com/nmeth/journal/v14/n1/abs/nmeth.4066.html#supplementary-information

184. Davalos M, Fourment J, Lucas A, et al (2004) Nitrogen regulation in Sinorhizobium meliloti probed with whole genome arrays. FEMS MicrobiolLett 241:33–40

185. Arcondéguy T, Jack R, Merrick M (2001) P II Signal Transduction Proteins, Pivotal Players in Microbial Nitrogen Control . Microbiology and Molecular Biology Reviews 65:80–105. https://doi.org/10.1128/MMBR.65.1.80-105.2001/ASSET/3B6E39BD-BDF0-4034-9FB2-6E71BD5560A0/ASSETS/GRAPHIC/MR0110007004.JPEG

186. Dixon R, Kahn D (2004) Genetic regulation of biological nitrogen fixation. Nat Rev Microbiol 2:621–631. https://doi.org/10.1038/nrmicro954

187. Hagberg KL, Yurgel SN, Mulder M, Kahn ML (2016) Interaction between Nitrogen and Phosphate Stress Responses in Sinorhizobium meliloti. Frontiers in Microbiology 0:1928. https://doi.org/10.3389/FMICB.2016.01928

188. Ronson CW, Nixon BT, Albright LM, Ausubel FM (1987) Rhizobium meliloti ntrA (rpoN) gene is required for diverse metabolic functions.

Journal of Bacteriology 169:2424–2431. https://doi.org/10.1128/JB.169.6.2424-2431.1987

189. Labes M, Rastogi V, Watson R, Finan TM (1993) Symbiotic nitrogen fixation by a nifA deletion mutant of Rhizobium meliloti: the role of an unusual ntrC allele. Journal of Bacteriology 175:2662–2673. https://doi.org/10.1128/JB.175.9.2662-2673.1993
190. Patriarca EJ, Merrick MJ, Iaccarino M (1998) Down-Regulation of the Regulatory System: A Mechanism to Uncouple Nitrogen Fixation and Assimilation in Bacteroids. 119–120. https://doi.org/10.1007/978-94-011-5159-7_38
191. Yurgel SN, Kahn ML (2008) A mutant GlnD nitrogen sensor protein leads to a nitrogen-fixing but ineffective Sinorhizobium meliloti symbiosis with alfalfa. Proc Natl Acad Sci U S A 105:18958–18963. https://doi.org/10.1073/pnas.0808048105
192. Ferro-Luzzi Ames G, Nikaido K (1985) Nitrogen regulation in Salmonella typhimurium. Identification of an ntrC protein-binding site and definition of a consensus binding sequence. The EMBO Journal 4:539–547. https://doi.org/10.1002/J.1460-2075.1985.TB03662.X
193. Hervás AB, Canosa I, Santero E (2010) Regulation of glutamate dehydrogenase expression in Pseudomonas putida results from its direct repression by NtrC under nitrogen-limiting conditionsm mi_7329 305..319. https://doi.org/10.1111/j.1365-2958.2010.07329.x
194. Patriarca EJ, Chiurazzi M, Manco G, et al (1992) Activation of the Rhizobium leguminosarum glnII gene by NtrC is dependent on upstream DNA sequences. Molecular and General Genetics MGG 1992 234:3 234:337–345. https://doi.org/10.1007/BF00538692

195. Alvarez-Morales A, Dixon R, Merrick M (1984) Positive and negative control of the glnA ntrBC regulon in Klebsiella pneumoniae. EMBO J 3:501–507. https://doi.org/10.1002/J.1460-2075.1984.TB01837.X

196. Pini F, Frage B, Ferri L, et al (2013) The DivJ, CbrA and PleC system controls DivK phosphorylation and symbiosis in Sinorhizobium meliloti. Mol Microbiol 90:54. https://doi.org/10.1111/MMI.12347

197. Pini F, de Nisco NJ, Ferri L, et al (2015) Cell Cycle Control by the Master Regulator CtrA in Sinorhizobium meliloti. PLOS Genetics 11:e1005232. https://doi.org/10.1371/JOURNAL.PGEN.1005232

198. Cheng J, Sibley CD, Zaheer R, Finan TM (2007) A Sinorhizobium meliloti minE mutant has an altered morphology and exhibits defects in legume symbiosis. Microbiology (N Y) 153:375–387. https://doi.org/10.1099/MIC.0.2006/001362-0

199. de Nisco NJ, Abo RP, Wu CM, et al (2014) Global analysis of cell cycle gene expression of the legume symbiont Sinorhizobium meliloti. Proc Natl Acad Sci U S A 111:3217–3224. https://doi.org/10.1073/PNAS.1400421111

200. Sourjik V, Muschler P, Scharf B, Schmitt R (2000) VisN and VisR are global regulators of chemotaxis, flagellar, and motility genes in Sinorhizobium (Rhizobium) meliloti. Journal of Bacteriology 182:782–788. https://doi.org/10.1128/JB.182.3.782-788.2000/ASSET/68DB6562-2C6C-49BA-85D0-FCCF475E6E1D/ASSETS/GRAPHIC/JB0301201004.JPEG

201. Baaziz H, Compton KK, Hildreth SB, et al (2021) McpT, a Broad-Range Carboxylate Chemoreceptor in Sinorhizobium meliloti. Journal of Bacteriology 203:. https://doi.org/10.1128/JB.00216-21

202. Rotter C, Mühlbacher S, Salamon D, et al (2006) Rem, a new transcriptional activator of motility and chemotaxis in Sinorhizobium

meliloti. Journal of Bacteriology 188:6932–6942. https://doi.org/10.1128/JB.01902-05/ASSET/A5D715C3-2E33-4127-8C75-0CB08DACC56D/ASSETS/GRAPHIC/ZJB0190661190009.JPEG

203. Szeto WW, Nixon BT, Ronson CW, Ausubel FM (1987) Identification and characterization of the Rhizobium meliloti ntrC gene: R. meliloti has separate regulatory pathways for activation of nitrogen fixation genes in free-living and symbiotic cells. Journal of Bacteriology 169:1423–1432. https://doi.org/10.1128/JB.169.4.1423-1432.1987
204. Browning DF, Butala M, Busby SJW (2019) Bacterial Transcription Factors: Regulation by Pick "N" Mix. Journal of Molecular Biology 431:4067–4077. https://doi.org/10.1016/J.JMB.2019.04.011
205. Patriarca EJ, Riccio A, Taté R, et al (1993) The ntrBC genes of Rhizobium leguminosarum are part of a complex operon subject to negative regulation. Molecular Microbiology 9:569–577. https://doi.org/10.1111/J.1365-2958.1993.TB01717.X
206. Patriarca EJ, Riccio A, Colonna-Romano S, et al (1994) DNA binding activity of NtrC from Rhizobium grown on different nitrogen sources. FEBS Letters 354:89–92. https://doi.org/10.1016/0014-5793(94)01094-3
207. Balbontin R, Fiorini F, Figueroa-Bossi N, et al (2010) Recognition of heptameric seed sequence underlies multi-target regulation by RybB small RNA in Salmonella enterica. Mol Microbiol 78:380–394
208. Papenfort K, Bouvier M, Mika F, et al (2010) Evidence for an autonomous 5' target recognition domain in an Hfq-associated small RNA. Proc Natl Acad Sci USA 107:20435–20440. https://doi.org/1009784107 [pii] 10.1073/pnas.1009784107
209. Shao Y, Feng L, Rutherford ST, et al (2013) Functional determinants of the quorum-sensing non-coding RNAs and their roles in target regulation. EMBO J 32:2158–2171. https://doi.org/10.1038/emboj.2013.155

210. Papenfort K, Vogel J (2009) Multiple target regulation by small noncoding RNAs rewires gene expression at the post-transcriptional level. Res Microbiol 160:278–287. https://doi.org/10.1016/j.resmic.2009.03.004
211. Fröhlich KS, Papenfort K, Berger AA, Vogel J (2012) A conserved RpoS-dependent small RNA controls the synthesis of major porin OmpD. Nucleic Acids Research 40:3623–3640. https://doi.org/10.1093/NAR/GKR1156
212. Rice JB, Vanderpool CK (2011) The small RNA SgrS controls sugar–phosphate accumulation by regulating multiple PTS genes. Nucleic Acids Research 39:3806–3819. https://doi.org/10.1093/NAR/GKQ1219
213. Amaya-Gómez C v., Hirsch AM, Soto MJ (2015) Biofilm formation assessment in Sinorhizobium meliloti reveals interlinked control with surface motility. BMC Microbiology 15:1–14. https://doi.org/10.1186/S12866-015-0390-Z/TABLES/1
214. Bernabéu-Roda L, Calatrava-Morales N, Cuéllar V, Soto MJ (2015) Characterization of surface motility in Sinorhizobium meliloti: regulation and role in symbiosis. Symbiosis 67:79–90. https://doi.org/10.1007/S13199-015-0340-4/FIGURES/7
215. Waters LS, Storz G (2009) Regulatory RNAs in Bacteria. Cell 136:615–628. https://doi.org/10.1016/J.CELL.2009.01.043
216. Vriezen JAC, de Bruijn FJ, Nüsslein K (2007) Responses of rhizobia to desiccation in relation to osmotic stress, oxygen, and temperature. Applied and Environmental Microbiology 73:3451–3459. https://doi.org/10.1128/AEM.02991-06/ASSET/5CC14C7F-50C9-488B-8439-B242B5EC4BAE/ASSETS/GRAPHIC/ZAM0110778380003.JPEG
217. Wei W, Jiang J, Li X, et al (2004) Isolation of salt-sensitive mutants from Sinorhizobium meliloti and characterization of genes involved in salt tolerance. https://doi.org/10.1111/j.1472-765X.2004.01577.x

218. Campbell GRO, Sharypova LA, Scheidle H, et al (2003) Striking Complexity of Lipopolysaccharide Defects in a Collection of Sinorhizobium meliloti Mutants. Journal of Bacteriology 185:3853. https://doi.org/10.1128/JB.185.13.3853-3862.2003

219. Sibley CD, MacLellan SR, Finan TM (2006) The Sinorhizobium meliloti chromosomal origin of replication. Microbiology (N Y) 152:443–455. https://doi.org/10.1099/MIC.0.28455-0/CITE/REFWORKS

220. Latch JN, Margolin W (1997) Generation of buds, swellings, and branches instead of filaments after blocking the cell cycle of Rhizobium meliloti. Journal of Bacteriology 179:2373–2381. https://doi.org/10.1128/JB.179.7.2373-2381.1997

221. Ledermann R, Schulte CCM, Poole PS (2021) How rhizobia adapt to the nodule environment. Journal of Bacteriology 203:539–559. https://doi.org/10.1128/JB.00539-20/SUPPL_FILE/JB.00539-20-S0001.PDF

222. de Bruijn FJ, Rossbach S, Schneider M, et al (1989) Rhizobium meliloti 1021 has three differentially regulated loci involved in glutamine biosynthesis, none of which is essential for symbiotic nitrogen fixation. Journal of Bacteriology 171:1673–1682. https://doi.org/10.1128/JB.171.3.1673-1682.1989

223. Torres MJ, Rubia MI, Bedmar EJ, Delgado MJ (2011) Denitrification in Sinorhizobium meliloti. Biochemical Society Transactions 39:1886–1889. https://doi.org/10.1042/BST20110733

224. Torres MJ, Rubia MI, de La Peña TC, et al (2014) Genetic basis for denitrification in Ensifer meliloti. BMC Microbiology 14:1–10. https://doi.org/10.1186/1471-2180-14-142/FIGURES/4

225. Cabrera JJ, Sánchez C, Gates AJ, et al (2011) The nitric oxide response in plant-associated endosymbiotic bacteria. Biochemical Society Transactions 39:1880–1885. https://doi.org/10.1042/BST20110732
226. Huala E, Ausubel FM (1989) The central domain of Rhizobium meliloti NifA is sufficient to activate transcription from the R. meliloti nifH promoter. Journal of Bacteriology 171:3354–3365. https://doi.org/10.1128/JB.171.6.3354-3365.1989
227. Mendoza A, Leija A, Martinez-Romero E, et al (1995) The enhancement of ammonium assimilation in Rhizobium etli prevents nodulation of Phaseolus vulgaris. Molecular Plant-Microbe Interactions 8:584–592. https://doi.org/10.1094/MPMI-8-0584
228. Mendoza A, Valderrama B, Leija A, Mora J (1998) NifA-dependent expression of glutamate dehydrogenase in Rhizobium etli modifies nitrogen partitioning during symbiosis. Molecular Plant-Microbe Interactions 11:83–90. https://doi.org/10.1094/MPMI.1998.11.2.83
229. Osburne MS, Signer ER (1980) Ammonium assimilation in Rhizobium meliloti. Journal of Bacteriology 143:1234–1240. https://doi.org/10.1128/JB.143.3.1234-1240.1980
230. Barnett MJ, Fisher RF, Jones T, et al (2001) Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti pSymA megaplasmid. Proc Natl Acad Sci U S A 98:9883–9888. https://doi.org/10.1073/PNAS.161294798/SUPPL_FILE/2947FIG2.JPG
231. Barnett MJ, Long SR (2015) The Sinorhizobium meliloti SyrM regulon: effects on global gene expression are mediated by syrA and nodD3. J Bacteriol. https://doi.org/10.1128/jb.02626-14
232. Dusha I, Austin S, Dixon R (1999) The upstream region of the nodD3 gene of Sinorhizobium meliloti carries enhancer sequences for the

transcriptional activator NtrC. FEMS Microbiology Letters 179:491–499. https://doi.org/10.1111/J.1574-6968.1999.TB08768.X

233. Dusha I, Kondorosi A (1993) Genes at different regulatory levels are required for the ammonia control of nodulation in Rhizobium meliloti. MGG Molecular & General Genetics 240:435–444. https://doi.org/10.1007/BF00280398
234. Dusha I, Bakos A, Kondorosi A, et al (1989) The Rhizobium meliloti early nodulation genes (nodABC) are nitrogen-regulated: Isolation of a mutant strain with efficient nodulation capacity on alfalfa in the presence of ammonium. Molecular and General Genetics MGG 1989 219:1 219:89–96. https://doi.org/10.1007/BF00261162
235. Groisman EA (2016) Feedback Control of Two-Component Regulatory Systems. http://dx.doi.org/101146/annurev-micro-102215-095331 70:103–124. https://doi.org/10.1146/ANNUREV-MICRO-102215-095331
236. Sun X, Zhulin I, Wartell RM (2002) Predicted structure and phyletic distribution of the RNA-binding protein Hfq. Nucleic Acids Res 30:3662–3671
237. Romby P, Charpentier E (2010) An overview of RNAs with regulatory functions in gram-positive bacteria. Cell Mol Life Sci 67:217–237. https://doi.org/10.1007/s00018-009-0162-8
238. Smirnov A, Wang C, Drewry LL, Vogel J (2017) Molecular mechanism of mRNA repression in trans by a ProQ-dependent small RNA. EMBO J 36:1029–1045. https://doi.org/10.15252/embj.201696127
239. Westermann AJ, Venturini E, Sellin ME, et al (2019) The Major RNA-Binding Protein ProQ Impacts Virulence Gene Expression in Salmonella enterica Serovar Typhimurium. MBio 10:. https://doi.org/10.1128/mbio.02504-18

240. Melamed S, Adams PP, Zhang A, et al (2020) RNA-RNA Interactomes of ProQ and Hfq Reveal Overlapping and Competing Roles. Mol Cell 77:411–425.e7. https://doi.org/10.1016/j.molcel.2019.10.022
241. Olejniczak M, Storz G (2017) ProQ/FinO-domain proteins: Another ubiquitous family of RNA matchmakers? Mol Microbiol doi: 10.1111/mmi.13679. https://doi.org/10.1111/mmi.13679
242. Windbichler N, Schroeder R (2006) Isolation of specific RNA-binding proteins using the streptomycin-binding RNA aptamer. Nat Protoc 1:637–640. https://doi.org/10.1038/nprot.2006.95
243. Rieder R, Reinhardt R, Sharma C, Vogel J (2012) Experimental tools to identify RNA-protein interactions in Helicobacter pylori. RNA Biology 9:520–531. https://doi.org/10.4161/rna.20331
244. Said N, Rieder R, Hurwitz R, et al (2009) In vivo expression and purification of aptamer-tagged small RNA regulators. Nucleic Acids Res 37:e133–e133. https://doi.org/10.1093/nar/gkp719
245. Windbichler N, von Pelchrzim F, Mayer O, et al (2008) Isolation of small RNA-binding proteins from E. coli: evidence for frequent interaction of RNAs with RNA polymerase. RNA Biol 5:30–40
246. Babitzke P, Romeo T (2007) CsrB sRNA family: sequestration of RNA-binding regulatory proteins. Curr Opin Microbiol 10:156–163. https://doi.org/10.1016/j.mib.2007.03.007
247. Jørgensen MG, Thomason MK, Havelund J, et al (2013) Dual function of the McaS small RNA in controlling biofilm formation. Genes Dev 27:1132–1145. https://doi.org/10.1101/gad.214734.113
248. Müller P, Gimpel M, Wildenhain T, Brantl S (2019) A new role for CsrA: promotion of complex formation between an sRNA and its mRNA target in Bacillus subtilis. RNA Biology 16:972–987. https://doi.org/10.1080/15476286.2019.1605811

249. Milojevic T, Sonnleitner E, Romeo A, et al (2013) False positive RNA binding activities after Ni-affinity purification from Escherichia coli. RNA Biology 10:1066–1069. https://doi.org/10.4161/rna.25195

250. Nisa-Martinez R, Jimenez-Zurdo JI, Martinez-Abarca F, et al (2007) Dispersion of the RmInt1 group II intron in the Sinorhizobium meliloti genome upon acquisition by conjugative transfer. Nucleic Acids Res 35:214–222. https://doi.org/10.1093/nar/gkl1072

251. Haddad N, Saramago M, Matos RG, et al (2013) Characterization of the biochemical properties of Campylobacter jejuni RNase III. Biosci Rep 33:e00082. https://doi.org/10.1042/BSR20130090

252. Beljantseva J, Kudrin P, Andresen L, et al (2017) Negative allosteric regulation of Enterococcus faecalis small alarmone synthetase RelQ by single-stranded RNA. Proc Natl Acad Sci U S A 114:3726–3731. https://doi.org/10.1073/pnas.1617868114

253. Hauryliuk V, Atkinson GC (2017) Small Alarmone Synthetases as novel bacterial RNA-binding proteins. RNA Biology 14:1695–1699. https://doi.org/10.1080/15476286.2017.1367889

254. Baltz AG, Munschauer M, Schwanhäusser B, et al (2012) The mRNA-Bound Proteome and Its Global Occupancy Profile on Protein-Coding Transcripts. Molecular Cell 46:674–690. https://doi.org/10.1016/J.MOLCEL.2012.05.021

255. Castello A, Fischer B, Eichelbaum K, et al (2012) Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell 149:1393–1406. https://doi.org/10.1016/J.CELL.2012.04.031

256. Matia-Gonzalez AM, Laing EE, Gerber AP (2015) Conserved mRNA-binding proteomes in eukaryotic organisms. Nat Struct Mol Biol 22:1027–1033. https://doi.org/10.1038/nsmb.3128

http://www.nature.com/nsmb/journal/v22/n12/abs/nsmb.3128.html#supplementary-information

257. Walden WE, Selezneva AI, Dupuy J, et al (2006) Structure of dual function iron regulatory protein 1 complexed with ferritin IRE-RNA. Science (1979) 314:1903–1908. https://doi.org/10.1126/SCIENCE.1133116/SUPPL_FILE/WALDEN-SOM.PDF
258. Nagy E, Rigby WFC (1995) Glyceraldehyde-3-phosphate dehydrogenase selectively binds AU-rich RNA in the NAD(+)-binding region (Rossmann fold). J Biol Chem 270:2755–2763. https://doi.org/10.1074/JBC.270.6.2755
259. Nagy E, Henics T, Eckert M, et al (2000) Identification of the NAD+-Binding Fold of Glyceraldehyde-3-Phosphate Dehydrogenase as a Novel RNA-Binding Domain. Biochemical and Biophysical Research Communications 275:253–260. https://doi.org/10.1006/BBRC.2000.3246
260. Holmqvist E, Li L, Bischler T, et al (2018) Global Maps of ProQ Binding In Vivo Reveal Target Recognition via RNA Structure and Stability Control at mRNA 3′ Ends. Molecular Cell 70:971–982.e6. https://doi.org/10.1016/j.molcel.2018.04.017
261. dos Santos RF, Arraiano CM, Andrade JM (2019) New molecular interactions broaden the functions of the RNA chaperone Hfq. Curr Genet 65:1313–1319. https://doi.org/10.1007/S00294-019-00990-Y
262. MacLean AM, Finan TM, Sadowsky MJ (2007) Genomes of the symbiotic nitrogen-fixing bacteria of legumes. Plant Physiology 144:615–622
263. Bobrovskyy M, Vanderpool CK (2013) Regulation of bacterial metabolism by small RNAs using diverse mechanisms. Annual Review of Genetics 47:209–232. https://doi.org/10.1146/annurev-genet-111212-133445

264. Robledo M, García-Tomsig NI, Jiménez-Zurdo JI (2020) Riboregulation in nitrogen-fixing endosymbiotic bacteria. Microorganisms 8:384. https://doi.org/10.3390/microorganisms8030384

265. Georg J, Lalaouna D, Hou S, et al (2020) The power of cooperation: Experimental and computational approaches in the functional characterization of bacterial sRNAs. Molecular Microbiology 113:603–612

266. Vogel J, Wagner EG (2007) Target identification of small noncoding RNAs in bacteria. Current Opinion in Microbiology 10:262–270. https://doi.org/10.1016/j.mib.2007.06.001

267. Lalaouna D, Eyraud A, Devinck A, et al (2019) GcvB small RNA uses two distinct seed regions to regulate an extensive targetome. Molecular Microbiology 111:473–486. https://doi.org/doi:10.1111/mmi.14168

268. Sergiev P v., Golovina AY, Osterman IA, et al (2016) N6-Methylated Adenosine in RNA: From Bacteria to Humans. J Mol Biol 428:2134–2145. https://doi.org/10.1016/J.JMB.2015.12.013

269. Deng X, Chen K, Luo GZ, et al (2015) Widespread occurrence of N6-methyladenosine in bacterial mRNA. Nucleic Acids Res 43:6557–6567. https://doi.org/10.1093/NAR/GKV596

270. Brenes-Álvarez M, Minguet M, Vioque A, Muro-Pastor AM (2020) NsiR1, a small RNA with multiple copies, modulates heterocyst differentiation in the cyanobacterium Nostoc sp. PCC 7120. Environmental Microbiology 22:3325–3338. https://doi.org/10.1111/1462-2920.15103

271. Ionescu D, Voss B, Oren A, et al (2010) Heterocyst-specific transcription of NsiR1, a non-coding RNA encoded in a tandem array of direct repeats in cyanobacteria. J Mol Biol 398:177–188. https://doi.org/10.1016/J.JMB.2010.03.010

272. Mitschke J, Vioque A, Haas F, et al (2011) Dynamics of transcriptional start site selection during nitrogen stress-induced cell differentiation in Anabaena sp. PCC7120. Proc Natl Acad Sci U S A 108:20130–20135. https://doi.org/10.1073/PNAS.1112724108/SUPPL_FILE/SD01.XLSX

273. Ren B, Wang X, Duan J, Ma J (2019) Rhizobial tRNA-derived small RNAs are signal molecules regulating plant nodulation. Science (1979) 365:919–922. https://doi.org/10.1126/science.aav8907

274. Massé E, Escorcia FE, Gottesman S (2003) Coupled degradation of a small regulatory RNA and its mRNA targets in Escherichia coli. Genes Dev 17:2374–2383. https://doi.org/10.1101/gad.1127103

275. Beisel CL, Storz G (2010) Base pairing small RNAs and their roles in global regulatory networks. FEMS Microbiol Rev 34:866–882. https://doi.org/10.1111/j.1574-6976.2010.00241.x

276. Bossi L, Figueroa-Bossi N (2016) Competing endogenous RNAs: a target-centric view of small RNA regulation in bacteria. Nat Rev Micro 14:775–784. https://doi.org/10.1038/nrmicro.2016.129

277. Denham EL (2020) The sponge RNAs of bacteria – How to find them and their role in regulating the post-transcriptional network. Biochimica et Biophysica Acta - Gene Regulatory Mechanisms 1863:194565