A highly conserved protein of unknown function in Sinorhizobium meliloti affects sRNA regulation similar to Hfq

Citation: Pandey, S. P. et al. "A Highly Conserved Protein of Unknown Function in Sinorhizobium Meliloti Affects sRNA Regulation Similar to Hfq." Nucleic Acids Research 39.11 (2011): 4691–4708. Web.

As Published: http://dx.doi.org/10.1093/nar/gkr060

Publisher: Oxford University Press

Persistent URL: http://hdl.handle.net/1721.1/72961

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Published online 15 February 2011 *Nucleic Acids Research, 2011, Vol. 39, No. 11* ***4691–4708***
doi:10.1093/nar/gkr060

A highly conserved protein of unknown function in *Sinorhizobium meliloti* affects sRNA regulation similar to Hfq

Shree P. Pandey[1], Brenda K. Minesinger[1], Janesh Kumar[2] and Graham C. Walker[1,*]

[1]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139-4307 and [2]Laboratory of Cellular and Molecular Neurophysiology, Porter Neuroscience Research Center, National Institute of Child Health and Human Development, Bethesda, MD, USA

Received August 12, 2010; Revised January 12, 2011; Accepted January 20, 2011

ABSTRACT

The SMc01113/YbeY protein, belonging to the UPF0054 family, is highly conserved in nearly every bacterium. However, the function of these proteins still remains elusive. Our results show that SMc01113/YbeY proteins share structural similarities with the MID domain of the Argonaute (AGO) proteins, and might similarly bind to a small-RNA (sRNA) seed, making a special interaction with the phosphate on the 5′-side of the seed, suggesting they may form a component of the bacterial sRNA pathway. Indeed, eliminating *SMc01113/YbeY* expression in *Sinorhizobium meliloti* produces symbiotic and physiological phenotypes strikingly similar to those of the *hfq* mutant. Hfq, an RNA chaperone, is central to bacterial sRNA-pathway. We evaluated the expression of 13 target genes in the *smc01113* and *hfq* mutants. Further, we predicted the sRNAs that may potentially target these genes, and evaluated the accumulation of nine sRNAs in WT and *smc01113* and *hfq* mutants. Similar to *hfq, smc01113* regulates the accumulation of sRNAs as well as the target mRNAs. AGOs are central components of the eukaryotic sRNA machinery and conceptual parallels between the prokaryotic and eukaryotic sRNA pathways have long been drawn. Our study provides the first line of evidence for such conceptual parallels. Furthermore, our investigation gives insights into the sRNA-mediated regulation of stress adaptation in *S. meliloti*.

INTRODUCTION

In a symbiotic event, rhizobial bacteria, e.g. *Sinorhizobium meliloti,* colonize the legume roots to fix biological nitrogen. This symbiosis is extremely important not only from agricultural, economical and environmental perspectives, but also the interaction of rhizobial bacteria with its legume host is a valuable model system to study chronic infection processes (1). The rhizobial bacteria alternate between a free living phase in the soil/rhizophere and a symbiotic phase within the host plant cells where they differentiate into nitrogen fixing bacteroids (2). Symbiosis comprises a complex series of events that begin with the infection process, which is initiated with the exchange of signals between the bacteria and the plant hosts (3,4). The signal components include the NOD factors, secreted proteins, polysaccharides, flavonoids, ligand-gated-ion channels, the calcium-calmodulin dependent protein kinases and calcium spiking (3,5). Invasion of plant roots by bacteria involves the development of an infection thread and cytoskeleton remodeling. During the invasion, bacteria encounter a plant innate immune response- a burst of reactive-oxygen-species (ROS) (4). Further along, the infection thread is targeted to the cortical tissues, where the rhizobia are endocytosed resulting in formation of symbiosome, followed by their differentiation into nitrogen-fixing bacteroids. Adaptation to the extracellular environment of the host cells requires lipopolysaccharide (LPS) and non-LPS factors such as those needed for coping with ion-stress (3,4). Bacteria enclosed within the symbiosome membrane encounter very low oxygen levels, where they can express enzymes of the nitrogenase complex and begin nitrogen fixation (4). Fixed carbon provided by plants in the form of

*To whom correspondence should be addressed. Tel: +1 617 253 6716; Fax: +1 617 253 2643; Email: gwalker@mit.edu

The authors wish it to be known that, in their opinion, the second and third authors should be regarded as joint Second Authors.

dicarboxylic acids is metabolized to provide the energy necessary for nitrogen fixation.

Large-scale transcriptional, proteomic and metabolic rearrangements underlie the complex transition from the free-living state to the symbiotic state, as well as the changes when the free living bacteria adapt to changing environmental conditions (2,6–9). For example, a 'base-line proteomic output' evaluated proteins from rhizobia grown in several conditions: an accumulation of a total of 2224 proteins, derived from 810 distinct genes and resulting in alteration of 53 metabolic pathways was determined (8). Analyses comparing transcription during microoxic versus oxic conditions and between free-living versus symbiotic states in *S. meliloti* revealed hundreds of genes, including several potential regulators, being differentially regulated (2). The microoxic and bacteroid transcriptomes partially overlapped, suggesting complex modes of gene regulation in different environments. Furthermore, the development of a 'dual-genome symbiosis chip' facilitated identification of changes in approximately 5000 transcripts during symbiosis and when *S. meliloti* is grown in different environments (9). Thus, complex transcriptional rearrangements occur in *S. meliloti* during its adaptation to environmental stresses and symbiosis. How these large scale transcriptional and proteomic responses are coordinated remains largely unknown. The magnitude of responses suggests that transcriptional regulation by certain proteins may not be the only mechanism of gene regulation.

RNA silencing has emerged as a fundamental regulatory process affecting many layers of endogenous gene expression; non-coding small-RNAs (sRNAs) appear to be important regulators of gene expression during stress responses in both eukaryotes and bacteria. It is only relatively recently that the pervasiveness, importance and roles of the sRNA-mediated regulation in bacteria have become apparent. They modulate transcription, translation, mRNA stability and genomic integrity via several mechanisms including changes in RNA conformation, protein binding, base pairing with target RNAs and interactions with DNA (10). Most known sRNAs regulate gene expression by base pairing with their target mRNAs that is facilitated by the RNA chaperone, Hfq; Hfq is implicated as a central player in the sRNA-mediated regulations (11–14).

Although much is understood about the role of sRNAs in modulating gene expression in the model bacteria, *Escherichia coli* and *Salmonella* (10,15), comparatively less is known about the possible role of the sRNA-mediated interaction in the model symbiotic bacteria, *S. meliloti*. A few studies have recently appeared that use comparative genomics and prediction approaches to identify sRNAs in the *S. meliloti* genome (16–18). Although expression of few of the predicted sRNAs was verified in these studies, the physiological roles of most of these sRNAs, their possible targets and the pathways that may be under their control remain largely unknown. In our and related recent studies, mutation of *hfq* resulted in *S. meliloti* that are highly compromised in their symbiotic ability (19–22), suggesting that the sRNA-pathways form a central component of the gene regulation during symbiosis and stress adaptation. Characterization of the symbiotically deficient *hfq* mutant provided important clues in understanding the underlying molecular mechanisms of regulation during symbiosis and stress adaptation. Nevertheless, hardly any specifics for Hfq-dependent sRNA–mRNA interactions are known in *S. meliloti*.

A screen that we had conducted to identify oxidative-stress-compromised symbiotically deficient mutants led us to identify *SMc01113* gene as essential for symbiosis (23,24). *SMc01113* encodes a protein of unknown function that is strongly conserved in most bacteria (24); *Escherichia coli* ortholog is named YbeY. The SMc01113/YbeY proteins belong to the UPF0054 family. While the crystal structures of the YbeY/UPF0054 members from *E. coli*, *Aquifex aeolicus* and *Thermotoga maritima* have been determined and suggest that YbeY is a metalloprotein (25–27), the cellular and biochemical function of this protein remains elusive. It has been annotated as a heat shock protein, as a peptidase or even as a protease (25,28,29). Recently, in *E. coli*, YbeY has been reported to be involved in translational regulation during high-temperature adaptation, presumably by affecting ribosome activity (28,29).

In the current investigation, we explore the possible cellular function of the *S. meliloti SMc01113* gene that appears to be central to symbiosis. Our sequence and structure analysis suggested SMc01113′s similarity to a domain of Argonaute (AGO) proteins, which are central to sRNA mediated interaction in eukaryotes (30,31). AGOs are multi-domain-proteins comprising of N, PAZ, MID and PIWI domains: the N- (or the amino-terminal) domain projects a basic surface face for putative RNA recognition; the PAZ domain contains a specific binding pocket that anchors the characteristic 2-nt 3′-overhang that results from digestion of RNAs by RNase III (a step in the processing of sRNAs); the PIWI domain shows extensive homology to RNase H, and these are often referred to as 'slicers'; the MID domain resides between the PAZ and PIWI domains and binds the characteristic 5′-phosphates of smRNAs, thus anchoring them onto the Ago protein; the MID domain has also been implicated in protein–protein interactions (32,33). AGO proteins constitute the key components of the 'RNA-induced-silencing-complex' (RISC) that modulates sRNA-dependent gene expression, and are mostly found in eukaryotes. This suggested the possibility that *SMc01113* may form a component of, or be involved in, the interactions mediated by sRNA-pathways. We used a comparative approach with *hfq* as a reference to investigate this hypothesis. Our results suggest a possible role of *SMc01113* in interactions mediated by the sRNA-pathway.

MATERIALS AND METHODS

Strains, media and culture conditions

Strains used in this study are detailed in the Supplementary Table S2. *Sinorhizobium meliloti* strains were grown aerobically at 30°C in the complex LB medium containing 5 ml/l of 0.5M $CaCl_2$ and 25 ml/l of

$MgSO_4$ (LB-MC) to an optical density at 600 (OD_{600}) of 1.5–2; they were then inoculated in the LB-MC or rhizobial minimal medium (RMM) with glucose at an OD_{600} of 0.1 (19,20). For selection of *S. meliloti*, streptomycin (500 μg/ml), gentamycin (50 μg/ml) and neomycin (200 μg/ml) were used. Bacto agar was used at 1.5% for solid medium.

Plant nodulation assay and fitness estimates

Nodulation assays for plants were conducted as previously described (19). Briefly, the alfalfa (*Medicago sativa* cv. Iroquois) seeds were surface sterilized and germinated on 0.8% agar in water in dark for 3 days; seedlings were placed on the top of a petri dish with 25 ml Jensen's agar, and inoculated with 1 ml appropriate strain of *S. meliloti* grown to saturation in LB-MC. Plates were wrapped in aluminium foil and incubated at 25°C for 4 weeks, after which fitness parameters such as plant height, number of pink/white nodules, number of green leaves and plant dry weight were recorded.

Physiological characterization following challenge with stresses

Bacterial response to various stresses for WT, *smc01113* and *hfq* mutants *S. meliloti* was evaluated by analysis of the growth curves in liquid LB-MC or RMM. The affect of mutating *SMc01113* on swarming behavior was evaluated on swarming plates by measuring the diameter of bacterial spreading as previously described (19) and compared to the WT and *hfq* mutant. Similarly, the affect of heat shock (50°C for 35–40 min) was evaluated on the WT, *smc01113* and *hfq* mutants as previously described (19,20). Ability to respond to various other environmental stresses was evaluated by growing the WT, *smc01113* and *hfq* mutants with or without Paraquat (100 mM), MMS (0.025%), ethanol (2.5%), SDS (0.1%), NaCl (50 mM) or cefotoxime (50 μg/ml) for 48 h, where the cells were incubated with shaking at 30°C. Aliquots were collected at different time intervals, OD_{600} was measured and residual growth was determined (19,20).

Sequence and structure analysis

Similarity between the SMc01113, YbeY and AGO protein sequences was determined by aligning the sequences with CLUSTALW (http://www.ebi.ac.uk/Tools/clustalw2/index.html). Structural homology search was performed using YbeY (PDB ID: 1XM5; chain A) as search probe on DALI server (34) and pair-wise comparisons were done using DALIlite (35). Structural similarities between the YbeY (1XM5) and the *A. aeolicus* AGO (2NUB) protein were also determined by super-imposing the YbeY structure (27) with the structures of complete AGO protein and MID/PIWI domain (32) using the Crystallographic Object-Oriented Toolkit [COOT; (36,37)]. YbeY structure was analyzed by Hotpatch server (http://hotpatch.mbi.ucla.edu/) (38) for the DNA/RNA binding properties and for finding functionally important patches on its surface. YbeY (PDB ID: 1XM5; chain A) and 4mer RNA (generated from PDB ID: 2BGG, chain R) were docked by using the ClusPro 2.0 docking server (http://cluspro.bu.edu) in order to model YbeY and RNA interactions. A single phosphate (PO_4) ion was manually docked in COOT (36,37). Electrostatic surface calculation and all figures were made by PYMOL. Sequences homologous to SMc01113/YbeY in other bacterial species were extracted after performing a non-redundant global BLAST search with the SMc01113. Homologous sequences were extracted from NCBI, multiple alignments were made and phylogenetic tree was constructed using CLUSTALW. Similarly, homologous sequences of the *S. meliloti* sRNAs were searched in closely related species (*S. medicae*, *Rhizobium* sp. NGR234, *R. leguminosarum* and *R. etli*) using the 'genome BLAST' tool at the NCBI (http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi). The homologous sRNA sequences were extracted, multiple alignments were made and phylogenetic trees were constructed using CLUSTALW.

Expression analysis by quantitative real-time PCR

The *smc01113* and *hfq* mutants and the WT *S. meliloti* were grown aerobically at 30°C in the LB-MC medium to an OD_{600} of 1.5–1.8; they were then inoculated in RMM with 0.5% glucose, and grown for 24 or 48 h, after which the cells were harvested by centrifugation. Total RNA was isolated by TRI reagent (T 9424; Sigma, St Louis, MO, USA) using manufacturer's protocol. SYBR green assays were developed using Express One-step SYBRGreenER kit (#11780-200; Invitrogen; http://www.invitrogen.com) following the manufacturer's protocol. All the quantitative real-time PCR (qPCR) assays were performed with 50 ng total RNA. Three to four biological replicates were used for each genotype; each biological replicate was used twice on the qPCR plate. The $2^{-\Delta\Delta CT}$ method (relative quantitation) was used for data analysis. Briefly, the C_t (threshold cycle) values of target genes were normalized to the endogenous reference gene ($\Delta C_t = C_{t\ \text{target}} - C_{t\ \text{reference}}$) and compared with those values obtained from the calibrator ($\Delta\Delta C_t = \Delta C_{t\ \text{sample}} - \Delta C_{t\ \text{calibrator}}$). To simplify data interpretation, expression levels in control samples (WT, Rm1021) were fixed to 1 and relative expression levels were calculated with respect to this reference value in the *smc01113* and *hfq* mutants. A list of primers used in this study is available in the Supplementary Table S3. All the gene-specific primers were designed with Primer3 software (http://frodo.wi.mit.edu/primer3/), and all the assays were run on the 'AB7500 fast real-time PCR system' (Applied Biosystems; http://www.appliedbiosystems.com/absite/us/en/home.html).

Bioinformatic analysis of sRNAs-target interactions

In order to determine if the *S. meliloti* sRNAs targeted the genes differentially regulated in the *smc01113* and *hfq* mutants, and informed by the fact that sRNAs regulate the gene expression by mostly base pairing with their target mRNA sequences, we used the concept of 'seed pairing', already used in some other bacteria (39–42). Complete coding sequences, obtained from the Rhizobase database (http://genome.kazusa.or.jp/rhizobase/), were

used for genes differentially regulated in the *smc01113* and *hfq* mutants. Using custom-written perl scripts, 'seed-pair analysis' of sRNAs was conducted for the target genes: small 'seed' sequences, >7 nt, starting from the first as well as the second nucleotide at the 5′-end of the sRNA, with a floating window of 1 nt, were generated. Then, sequences from the 3′-ends of the target mRNA was compared for perfect complementarities with sRNA-seed-sequences from the 5′-ends. These seeds were mapped to the sequences of the target genes for 'Watson–Crick' base pairing, and a hit-map was generated (Supplementary Table S1). Most *trans*-acting sRNAs have the seeds >9 nt (10,13,43), therefore we used a length of 10 or more perfectly matching nucleotides as a cut-off. To test the statistical significance, similar analysis of predicting sRNA-complementary motifs was performed on a set of unrelated and randomized mRNA sequences [from plants; (44)]. Significantly low number of hits were obtained for seeds >10 nt (Fisher's exact test, $P \leq 0.05$).

In order to determine the conservation of the sRNA-target interaction sites for specific sRNAs and target mRNAs, we used sequences from three closely related species: *S. meliloti*, *S. medicae* and *Rhizobium sp.* NGR234. We extracted the longest 'seed' sequence from the target as well as sRNA. Sequences homologous to these 'seed sequences' were extracted from the aforementioned multiple alignments of the sRNAs and targets. Separate conservation profiles for sRNAs and targets were generated using the 'WebLogo' program (45,46). It is advantageous to use this methodology of determining the conservation as it gives frequencies of conservation of a given nucleotide at a particular position, consensus sequence and the total sequence conservation.

Over-expression of sRNAs

For over-expression of sra16 and sra35, genomic DNA corresponding to the appropriate sra coding sequence was PCR amplified from isolated *S. meliloti* Rm1021 genomic DNA and with engineered 5′-Xho*I* and 3′-Kpn*I* restriction sites. These DNA fragments were then cloned into Xho*I*/Kpn*I*-digested pMSO3 (47) under the control of the constitutive *Trp* promoter. Transfer of pMS03-sra16 and pMS03-sra35 into *S. meliloti* (SRA16 and SRA35) was achieved by tri-parental mating followed by selection of *S. meiloti* harboring the plasmid on LB/MC containing 100 μg/ml spectinomycin and 500 μg/ml streptomycin. WT Rm1021 strains were also grown with streptomycin. Over-expression of sra16 and 35, and its effects on their predicted targets were evaluated by qPCR analyses after growing the cells to late exponential/early stationary phase. We evaluated the affect of pMSO3 vector in the Rm1021 by similarly generating control Rm1021 strains with empty vector. No significant difference in growth or gene-expression was observed between the WT and the control WT strain harboring empty vector, which confirmed that presence of the vector had no effects on bacteria (data not shown). Similarly, over-expression of sra35 had no affect on growth of *S. meliloti*; sra16-over-expressing strain grew marginally better than WT or WT with empty vector controls (data not shown).

Yeast two-hybrid analysis

For cloning, XL1Blue *E. coli* were plated on LB media supplemented with 100 μg/ml of ampicillin and incubated at 37°C. Yeast were grown on synthetic complete medium containing 2% dextrose (SCD) and lacking the appropriate nutrient (48). Plates were supplemented with 2% agar. Yeast cultures and plates were incubated at 30°C.

Plasmids were constructed by PCR amplification of the targeted gene from *E. coli* genomic DNA (strain MC4100) using oligonucleotides containing homology to the gene of interest flanked by either *Eco*RI for the 5′-primer or *Bgl*II for the 3′-primer (Supplementary Table S3). The 5′-primer lacked homology to the start codon of the amplified ORF to insure correct fusion to the Gal4 binding domain (BD) or activation domain (AD). After amplification, the amplicon was digested with *Eco*RI and *Bgl*II, and ligated to *Eco*RI and *Bgl*II-digested pGBD-C2 and pGAD-CD plasmids (Supplementary Table S2). Correct clones were confirmed by DNA sequencing.

To check for protein interactions, the budding yeast strain PJ69-4A (49) (*MAT***a** *trp1-901 leu2-3, 112 ura3-52 his3-200 gal4*Δ *gal80*Δ *LYS2*::*GAL1-HIS3 GAL2-ADE2 met2*::*GAL7-lacZ*) was transformed sequentially using a standard lithium acetate protocol (50). In the first round, PJ69-4A was transformed with plasmids pGBD-C2, pM122, pGAD-C2 and pM130 selecting for either Trp^+ (pGBD-C2, pM122) or Leu^+ transformants (pGAD-C2 and pM130) on SCD–Trp or SCD–Leu, respectively. Colonies from the first round were subsequently transformed with the second plasmid containing the putative interacting protein and selected for on SCD–Leu–Trp. Finally, multiple Leu^+Trp^+ colonies from each transformation were pooled and resuspended in 100 μl of ddH_2O. Ten microliter of the resuspended colonies were spotted on SCD–Leu–Trp to confirm growth and SCD–Leu–Trp–Ade to select for protein–protein interactions. Plates were incubated at 30°C for 5 days before interpretation of data.

RESULTS

SMc01113 shows sequence and structural similarities to the AGO proteins

SMc01113 homologs are found in nearly every bacterium including symbiotic and pathogenic bacteria [Supplementary Figure S1; (24,27)]. SMc01113 and its *E. coli* homolog, YbeY, show amino acid similarity of >70% (Figure 1A). Expression of the *E. coli* homolog complements the symbiotic defects of the *smc01113* mutant of *S. meliloti* (24,51). Sequence and structural similarities of the YbeY to AGO protein were readily apparent when we performed homology studies with the YbeY. Sequence alignment of SMc01113/YbeY showed similarities with AGO protein sequences from

A

```
YbeY_Ecoli      MSQVILDLQLACEDNSGLPEE---SQFQTWLNAVIPQFQEE---------SEVTIRVVDT 48
SMc01113        --MTALDIQISVEAGDWPPEDELQSFCERVLEAAADFLAREENQPLPAQAAELSLVFTDD 58
                   . **:*:: * ..  **:   *  :  *:*.   : .*         :*::: ..*

YbeY_Ecoli      AESHSLNLTYRGKDKPTNVLSFP-FEVPPG-MEMSLLGDLVICRQVVEKEAQEQGKPLEA 106
SMc01113        QSIRAINAEWRGQDKATNVLSFPAFPVTPGRMPGPMLGDIVVAHETLRREAAELEKPFDA 118
                 . :::*  :**:**.******* * *.** *  .:***:*:.::.:.:** *  **::*

YbeY_Ecoli      HWAHMVVHGSLHLLGYDHIEDDEAEEMEALETEIMLALGYEDPYIAEKE- 155
SMc01113        HLTHLLVHGFLHLFGYDHIEDDEAERMEGLETRILARLGLSDPYGDQPPH 168
                * :*::*** ***:***********.**.***.*:  ** .***  :

Mid_AaAgo       NLRKFLELCRPFVKKDVLS--VEIISVSVYKKLEWRKE-EFLKELINFLKNKGIKLKIKG 57
YbeY_Ecoli      MSQVILDLQLACEDNSGLPEESQFQTWLNAVIPQFQEESEVTIRVVDTAESHSLNLTYRG 60
                  : :*:*  .  .:. *.   :: :       ::::* *.  .:::  :.:.::*. :*

Mid_AaAgo       K---SLILAQTREEAKEKLIPVINKIKDVDLVIVFLEEYPKVDPYKSFLLYDFVKRELLK 114
YbeY_Ecoli      KDKPTNVLSFPFEVPPGMEMSLLGDLVICR-QVVEKEAQEQGKPLEAHWAHMVVHGSLH- 118
                *   : :*: . * .    :.::.:.      :*  *   : .* ::.  : .*: .*

Mid_AaAgo       KMIPSQVILNRTLKNENLKFVLLNVAEQVLAKTGNIPYKL 154
YbeY_Ecoli      -LLGYDHIEDD--EAEEMEALETEIMLALGYEDPYIAEKE 155
                 ::  : * :   : *::: :  ::   :  :   *. *
```

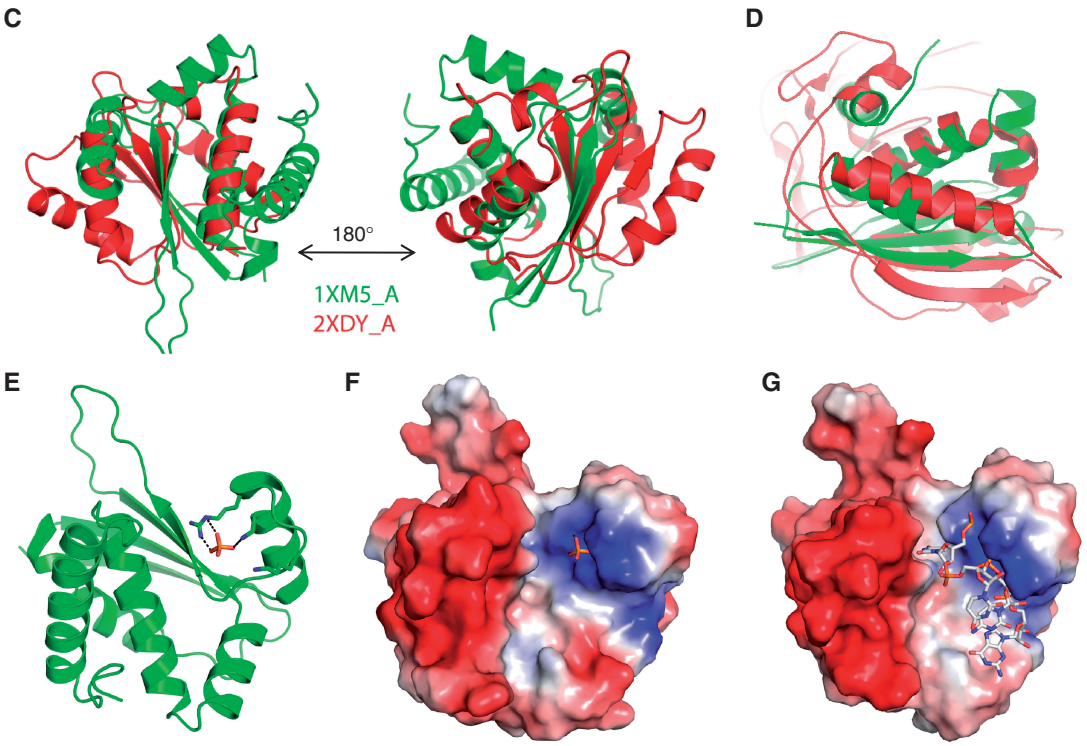

Figure 1. *Sinorhizobium meliloti* ORF, *SMc01113*/*ybey*, shares similarities with Ago protein. (**A**) Alignment of the protein sequences of the *S. meliloti* and *E. coli* homologs, SMc01113 and YbeY. (**B**) Alignment of protein sequences of the YbeY and Ago-Mid domain. Structural alignment of Ago-Mid domain (red) and YbeY (green) from *N. crassa* (**C**) and *A. aeolicus* AGO (2NUB) proteins (**D**). Manually docked PO4 shows potential interactions with the Arg59, K61 residues (**E**) in the positively charged cavity (**F**). (**G**) The lowest energy docked 4mer RNA (Cluspro 2.0 web server) onto YbeY surface fits nicely into the protein cavity, with negatively charged RNA backbone aligned towards the positively charged protein surface.

Neurospora crassa, *A. Aeolicus* and other species. In particular, SMc01113/YbeY showed sequence similarity with the MID domain of the AGO protein (Figure 1). The MID domain contains the binding site for the 5′-end of the eukaryotic sRNAs, by binding to the PO_4 group of a nucleic acid. The structure of *E. coli* YbeY (27) has been solved, as have the structures of AGO proteins from *A. aeolicus* and the *Thermus thermophilus* (32,33). Recently, the crystal structure of a eukaryotic MID-domain, recognizing the 5′-terminal phosphate of a guide RNA, has also been solved, and it highly resembles the previous structures of MID domains from other species (52).

A structural homology search with YbeY (PDB ID: 1XM5) by using the DALI server (34) against the full PDB database and PDB90 (34) resulted in multiple hits corresponding to a wide variety of domains with RNA-binding, enzymatic and regulatory functions. Among all the DALI hits, the MID domain of *N. crassa* AGO (PDB code: 2XDY), *Pyrococcus furious* AGO (PDB code: 1Z26) and the MID domain from human AGO (PDB ID: 3LUJ) show statistically significant *Z*-scores of 3.6, 3.2 and 2.9, respectively. Further, a pair-wise structure alignment using DAliLite (35) yielded a significant *Z*-score of 4.7 for YbeY (PDB ID: 1XM5) and the *N. crassa* AGO MID domain (PDB ID: 2XDY). Indeed, there are 83 equivalent C-alpha positions between YbeY and the *N. crassa* AGO MID domain with a root mean square deviation (RMSD) of 3.4 Å (Figure 1C). Further, we tested the structural similarities of YbeY with the MID-domain of the *A. Aeolicus* AGO protein; structural superimposition showed RMSD of 2.6 Å for 72 aligned residues indicating structural similarity (Figure 1D).

Encouraged by these results, we further explored the potential of YbeY to bind RNA. YbeY and a 4mer RNA (52) were docked by using the fully automated, web-based program ClusPro version 2.0 (53,54) with default parameters (Figure 1E–G). The ClusPro docking server yielded 69 top-scoring solutions based on the balanced, electrostatic-favored, hydrophobic-favored and van der Waals plus electrostatic-favored coefficients (53,54). The 4mer RNA fits nicely into the protein cavity (Figure 1G) and the negatively charged RNA backbone is aligned towards the positively charged protein surface of the cleft. Most strikingly, in the top-scoring model (Figure 1E–G) the free PO_4 group of RNA interacts with the Arg59, Lys61 (Figure 1E–F). The docking results suggest a probable RNA binding site in the YbeY protein. We further subjected the YbeY structure to analysis of surface binding properties by Hotpatch webserver (38). HotPatch finds unusual patches on the surface of proteins, and statistically computes how unusual they are (patch rareness), and how likely each patch is to be of functional importance [functional confidence (FC)]. The statistical analysis is done by comparing protein's surface against the surfaces of a large set of proteins whose functional sites are known. We analyzed the YbeY structure to find surface patches with putative DNA/RNA binding properties. DNA/RNA-binding refers to proteins that interact with oligonucleotides either catalytically or non-catalytically. Interestingly, Hotpatch identified three patches with DNA/RNA binding properties and they fall in the same protein cavity where we have docked the RNA 4mer (Supplementary Figure S1C and D). A fourth functionally important patch, corresponding to a His-triad (H114, H118, H124) and possibly coordinating a metal ion, was identified (Supplementary Figure S1D). This may provide a hydrolytic function to YbeY. These observations raised the intriguing possibility that SMc01113/YbeY may be involved in the process of sRNA-mediated interactions, as are the AGO proteins.

The *S. meliloti smc01113 mutant* shows symbiotic and physiological phenotypes highly similar to those of *hfq* mutant

Parallels between the prokaryotic and eukaryotic sRNA machinery have been noted (13), with the RISC complex serving as the protein scaffold for pairing of the sRNA with its target in eukaryotes, while the Hfq protein has been proposed to play a conceptually analogous role in bacteria. Members of biological pathways deregulated when the RNA chaperone, Hfq, is mutated, are most likely targets of sRNA-machinery (15). We reasoned that if *SMc01113* is involved in sRNA-mediated interactions, physiological and molecular loss-of-function phenotypes of *smc01113*/*ybey* should be comparable to those of other established components of sRNA-pathway such as *hfq*. To begin with exploring the hypothesis that SMc01113 might participate in the processes related to sRNA-mediated interactions, we compared the *smc01113* mutant phenotypes to that of an *hfq* mutant. Loss of Hfq (19) as well as SMc01113 (24) makes *S. meliloti* similarly symbiotically deficient, resulting in high fitness costs for both the interacting partners (Supplementary Figure S2). Swarming is an important multicellular phenomenon characterized by the coordinated and rapid movement of bacteria across semisolid surfaces, which may help the bacteria to adapt to extra- and intra- cellular niches (55). Both, *hfq* and *smc01113* mutants were equally defective in their ability to swarm (Figure 2A, Supplementary Figure S3).

We evaluated the susceptibility of the *hfq* and *smc01113* mutants to several stresses (Supplementary Figure S3) and observed that they displayed strikingly similar phenotypes (Figure 2). *Escherichia coli* YbeY has been implicated with adaptation to heat-shock (29); analogously, the loss of *hfq* or *smc01113* makes *S. meliloti* similarly susceptible to heat shock (Figure 2). Further, the *smc01113* and *hfq* mutants displayed similar sensitivities to agents like SDS and ethanol that may change the surface potential of the membranes and affect cellular osmoregulation (Figure 2). Resistance to oxidative stress, resulting in generation of reactive-oxygen-species is critical to *S. meliloti* (56). We evaluated the sensitivity of *smc01113* and *hfq* mutants to oxidative stress (Paraquat): both were similarly sensitive (Figure 2). We have previously reported that both the *smc01113* mutant (24) and the *hfq* mutant (20) are sensitive to hydrogen peroxide. The *smc01113* and *hfq* mutants were similarly sensitive to a DNA damaging agent, methyl methanesulfonate (MMS) (Figure 2). Additionally, we tested the sensitivity of the two mutants to the β-lactam antibiotic Cefotaxime, which affects bacterial cell wall synthesis, and to elevated salt concentration. Both the mutants were similarly susceptible to Cefotaxime treatment, but not to NaCl (Figure 2). Taken together, these observations indicate it is plausible that Hfq and SMc01113 play roles that are physiologically related.

SMc01113 and Hfq regulate similar molecular targets

To test this hypothesis of functionally related roles for these two proteins, we assayed mRNA accumulation patterns in the *smc01113* mutant for the genes that are

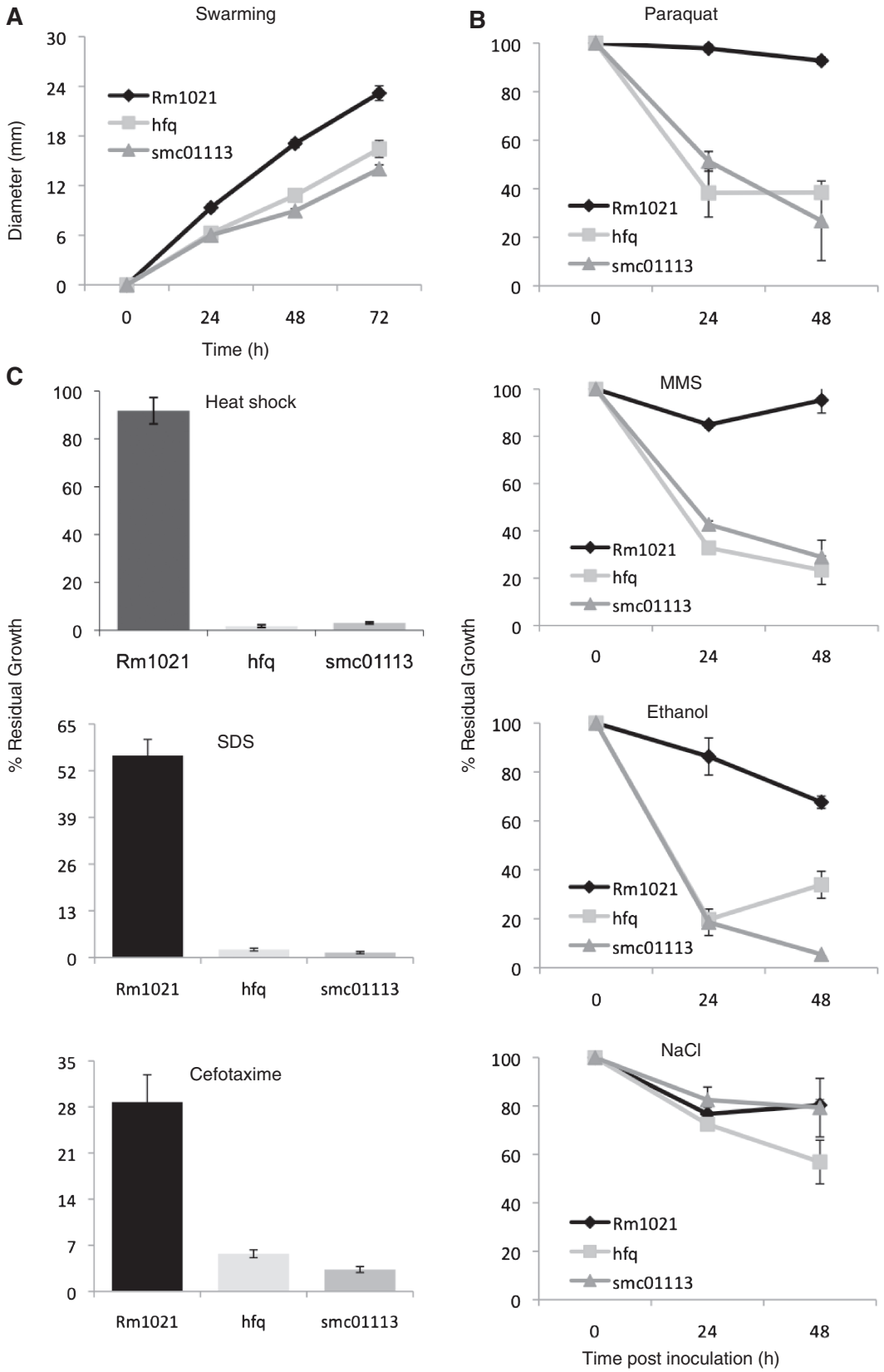

Figure 2. Loss of *hfq* or *SMc01113* makes *S. meliloti* similarly sensitive to environmental stresses. (**A**) *smc01113* mutant swarms as less efficiently as the *hfq* mutants than the wild-type strain. Strains were plated on the swarming plates (0.3% agar) and incubated at 30°C. Results of swarming for the three strains were evaluated by measuring colony progression every day for 3 days. (**B** and **C**) show sensitivity of the *smc01113* and the *hfq* mutants towards variety of stresses. Strains were grown with or without stress agents as described in the text and OD_{600} were measured, and percent residual growth (B) after 24 and 48 h (for paraquat, MMS, ethanol and NaCl) or (C) 48 h (for heat shock, SDS and cefotaxime) was calculated.

differentially regulated in the *hfq* mutant and observed striking overlaps (Figures 3 and 4). We recently identified proteins significantly affected in their expression in the *S. meliloti hfq* mutant (20). Of the 55 differentially accumulated proteins in the *hfq* proteome, 30 are annotated genes. Furthermore, we found that the activity of KatB and SodC was reduced in the *hfq* mutant compared to a wild-type (WT) strain (20). In order to understand how loss of SMc01113 affected the transcriptional accumulation of the targets that are under Hfq control, we performed qPCR analyses for 13 of the genes (namely, *sodC*, *katB*, *agpA*, *atpD*, *ppiA*, *ehuB*, *cysK1*, *livK*, *dppA1*, *dppA2*, *aapJ*, *glnA* and *thtR*), which are associated with diverse processes such as transport and oxidative/superoxide resistance. Most of these genes have been confirmed to show differential accumulation in *hfq* mutants in independent studies, in *S. meliloti*, *E. coli*, *Vibrio cholerae* and *Salmonella* (20–22) at RNA or protein levels in independent studies compared to WT Rm1021 strain, accumulation of eight of these genes was similarly down-regulated (Figure 3) and four similarly up-regulated (Figure 4) in both mutants, whereas one gene showed a differential accumulation pattern between the *hfq* and *smc01113* mutants.

Compared to the WT Rm1021 strain, the genes involved in protection from oxidative/superoxide stresses such as *katB* (catalase B), *sodC* (superoxide dismutase C), *cysK1* (Cysteine synthase A) were down-regulated in both

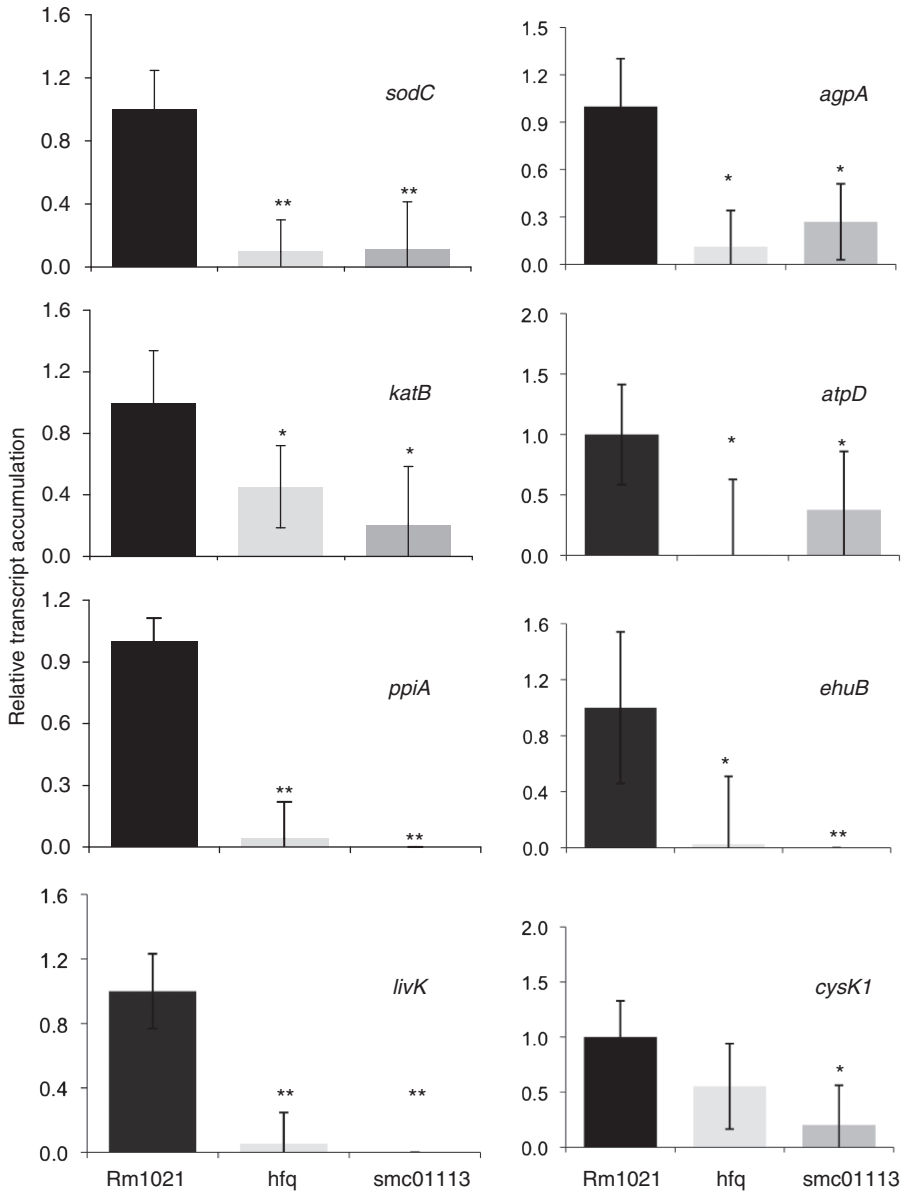

Figure 3. Loss of *hfq* or *SMc01113* similarly down-regulates target transcript accumulation. WT and the *smc01113* and the *hfq* mutants were grown in the RMM till exponential (*ppiA*, *katB* and *cysK1*) or stationary phase (*sodC*, *agpA*, *livK*, *ehuB* and *atpD*), RNA was extracted and qPCR was performed. Transcript accumulation in the WT strain was fixed to 1 and relative transcript abundance to WT in the two mutants was evaluated. 'Single' and 'double asterisk' shows significant difference at $P \leq 0.05$ and $P \leq 0.01$, respectively.

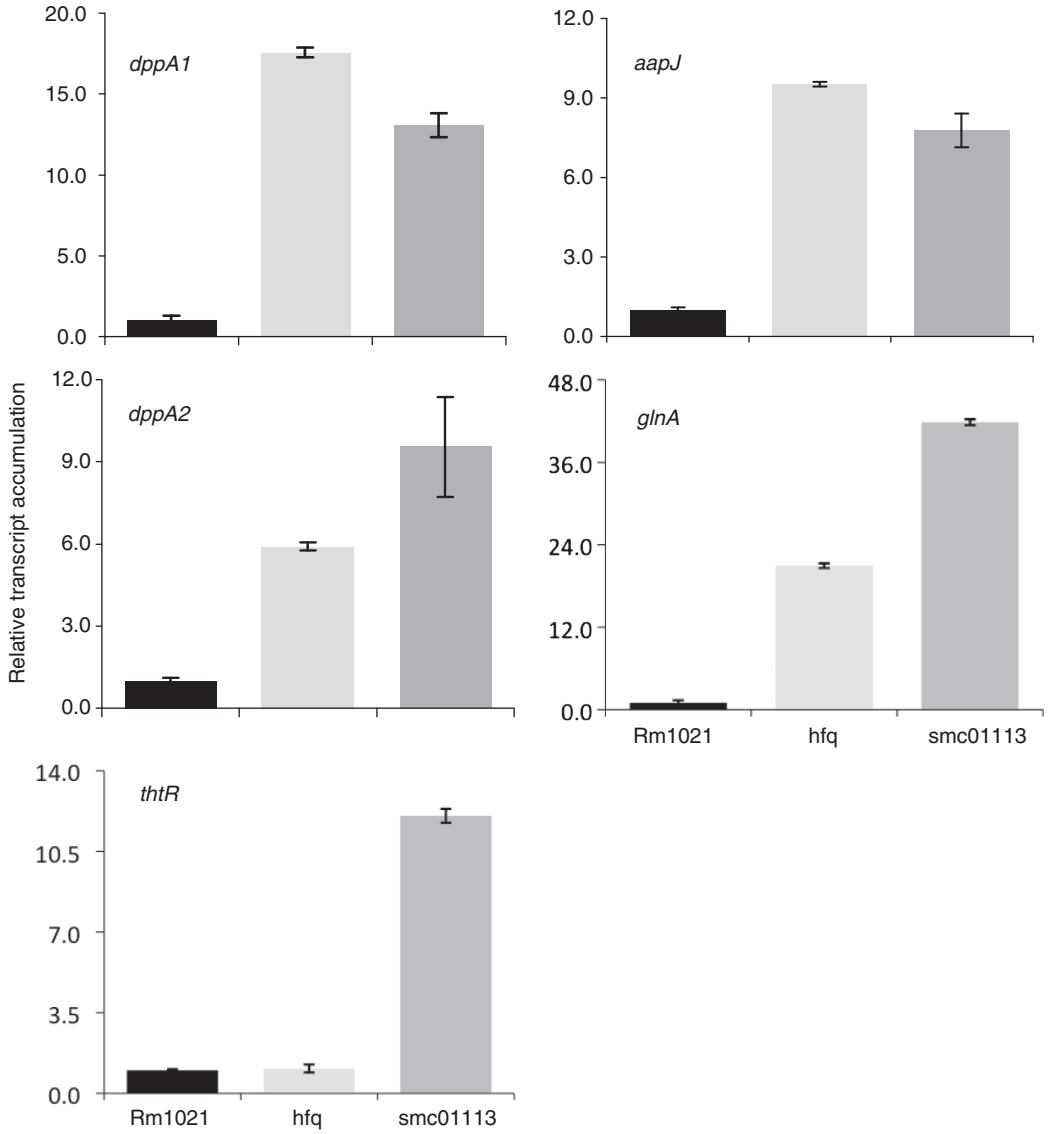

Figure 4. Up-regulated target transcript accumulation in the *S. meliloti* mutated for *hfq* or *SMc01113*. WT and the *smc01113* and the *hfq* mutants were grown in the RMM till stationary phase, RNA was extracted and qPCR was performed. Transcript accumulation in the WT strain was fixed to 1 and relative transcript abundance to WT in the two mutants was evaluated.

the *smc01113* and *hfq* mutants (Figure 3). Similarly, the expression levels of *agpA* [α-galactoside-binding protein; periplasmic binding protein gene required for the utilization of sugars (57)], *atpD* (F0F1 ATP synthase beta subunit), *ppiA* [peptidyl-prolyl *cis–trans* isomerise; involved in proper folding of newly synthesized proteins, resistance to antibiotics and in cell cycle (58,59)] and *ehuB* [Ectonine binding protein; belonging to ATP-binding cassette transporters, a central to osmoprotection under hyper-osmotic stress in *S. meliloti* and provides substrate specificity (60,61)] were down-regulated in both the mutants, when compared to WT Rm1021 (Figure 3). Somewhat surprisingly, the expression of *livK* (which was up-regulated in the *hfq* proteome) was transcriptionally down-regulated in both the *smc01113* and *hfq* mutants (Figure 3).

Transporters involved in peptide [*dppA1* and *dppA2;* (62)], amino acids and solute transport [*aapJ;* (63)] were up-regulated in both of the *smc01113* and *hfq* mutants when compared to their transcript levels in WT Rm1021 (Figure 4). Similarly, *glnA* (glutamine synthetase 1, a key enzyme in nitrogen assimilation) was up-regulated in both the mutants (Figure 4). However, *thtR* (a sulfotransferase) mRNA showed a difference in accumulation pattern between the *hfq* and *smc01113* mutants. No difference in mRNA levels was observed between the WT and *hfq* mutant, whereas *smc01113* mutant showed enhanced *thtR* mRNA levels, compared to WT and *hfq* mutant (Figure 4). Taken together, these observations suggest that bacterial SMc01113/YbeY proteins play regulatory roles that are highly related, but not completely identical to, those played by the RNA chaperone, Hfq.

In silico identification of possible sRNAs and their binding sites in the target genes differentially regulated in *hfq* and *smc01113* mutants

Above-mentioned results prompted us to search for the sRNAs that may regulate the altered expression of the above said targets in *hfq* and *smc01113* mutants. Our hypothesis suggested that we would observe similar patterns of accumulation of sRNAs between the *hfq* and *smc01113* mutants. We, therefore, undertook a computational approach to identify possible sRNAs that might target the 13 genes tested above.

sRNA-mediated regulation of gene expression falls into two broad categories (10). First is by the *cis*-encoded anti-sense sRNAs, which have extensive complementarity (>75 bp) with their targets, and second, by the *trans*-encoded sRNAs, which have more limited complementarity. The *trans*-encoded sRNAs have 10–25 bp complementary 'seed sequences' to their target mRNAs in discontinuous patches, where only a core of nucleotides seems to be critical for regulation. These sRNAs regulate translation by targeting regions in the 5′ UTRs, inside and upstream of the ribosome-binding sites (RBS)/Shine–Dalgarno sequences, occluding or sequestering RBS, as well as binding within the coding sequences of the targets largely leading to their degradation (13,43,64,65). The *trans*-encoded sRNAs are functionally analogous to eukaryotic microRNAs (miRNAs), often require Hfq for their action (11,13), may pair targets in the coding region (43) and thus could be the preferred mode of regulation of the differentially accumulated targets identified in *hfq* and *smc01113* mutants. The principle of seed pairing has been extensively and effectively used for determining sRNA–mRNA interaction in eukaryotes (39,41,44,66,67) as well as in some prokaryotes (40,42). In general, the longer the seed sequence (and thus the greater the perfect complementation), the higher the chances of an effective sRNA-target interaction.

Informed by these considerations, we performed a bioinformatic analysis to identify *S. meliloti* sRNAs (17) that could potentially bind to the genes differentially regulated in the *hfq* and *smc01113* mutants. Our analysis revealed that these genes have a large number of potential sRNA binding sites (Supplementary Table S1). Genes with multiple sRNA-binding sites of >10 nt, for specific sRNAs were identified; e.g. *dppA1* had three different complementary sites for the sRNA, sra16 (Figure 6B). It was strikingly apparent that a single sRNA has the potential to target more than one gene, and that a single gene can be targeted by several sRNAs (Supplementary Table S1; Figure 6A). For instance, *sodC* had a 13-nt sRNA complementary site for one sRNA, 12-nt seeds for five sRNAs, and 11-nt seeds for eight sRNAs (Supplementary Table S1). The potential for single sRNA to target multiple mRNAs is also striking. For example, our analysis indicated that sra35 potentially targets three genes, sra03 may target six genes and sra16 may target 10 genes (Supplementary Table S1, Figure 6A). Similar observations of 'one-to-many' sRNA–mRNA interactions, common in eukaryotes [e.g. (39)], have been previously reported in some other bacterial species, and would have profound effects on cellular physiology and stress adaptation (43,65,68–70). Using a combination of such criteria of longest seed length and potential to target several genes, we selected 9 sRNAs for their expression analysis in WT, a *smc01113* mutant and an *hfq* mutant.

Expression of sRNAs is similarly affected in *S. meliloti smc01113* and *hfq* mutants

Using qPCR analyses, we next investigated the relationship between changes in sRNA accumulation in *smc01113* and *hfq* mutants, and the changes in the various potential mRNA targets. Of the nine sRNAs tested, the accumulation of eight sRNAs (sras 03, 11, 16, 45, 47, 51, 63 and 65) was similarly reduced in both the *hfq* and *smc01113* mutants compared to the WT Rm1021 (Figure 5). In contrast, sra35 was strongly up-regulated in both *hfq* and *smc01113* mutants relative to the WT Rm1021 strain (Figure 5). The pattern of the variation of sRNAs relative to particular mRNA target indicated both, inverse correlations (up-regulation of target corresponding to down-regulation of the sRNAs and *vice versa*) as well as direct correlations (change of target and the targeting sRNAs in the same direction; Figure 6A).

sRNA-mediated-interactions in bacteria results in both the inhibition and the stimulation of translation (10,13,43,64,65,69). Inverse correlations of expression of sRNAs and their potential targets (red/green regions in Figure 6A) suggest inhibitory action of these sRNAs on their targets; conversely similarly-regulated sRNA-target pairs (yellow regions, Figure 6A) indicate a stimulatory action of sRNAs in these interactions. Thus, these patterns suggest the existence of complex interactions between the sRNAs and their targets that are necessary for modulating and fine tuning the expression of genes to optimum levels.

To gain further insight into the biological relevance of the putative sRNA-target interactions described above (Figure 6A, Supplementary Table S1), we performed phylogenetic comparisons of the sRNAs and sRNA-target-motifs in mRNAs found in our *in silico* analysis in multiple rhizobium species. Homologous sequences to several of these differentially regulated sRNAs were identified in three or more related *Rhizobium* species (Supplementary Figures S4 and S5), suggesting their wide organismic relevance in symbiosis. Some of these sRNA sequences have been reported in other α-proteobacterial species (17). Conservation in mRNA-targets with complementarity to the seed (2–7 nt) of the miRNAs in closely related species has been used as one of the main criteria of identifying miRNA-target interactions in eukaryotes (41,71–73). On the other hand, sRNAs/targets in viruses are not well conserved (74). When we performed an analysis of conservation of the complementary sites in the mRNA as well as in sRNAs in the three closely related rhizobial bacteria, we found that the motifs of sRNAs as well as their complementary sequences in the targets were fairly well conserved (Figure 6C). SNPs in both sRNAs as well as mRNAs in other rhizobial species were seen (Figure 6C). Single-base substitutions from T to G/C (and *vice versa*) were seen in few cases.

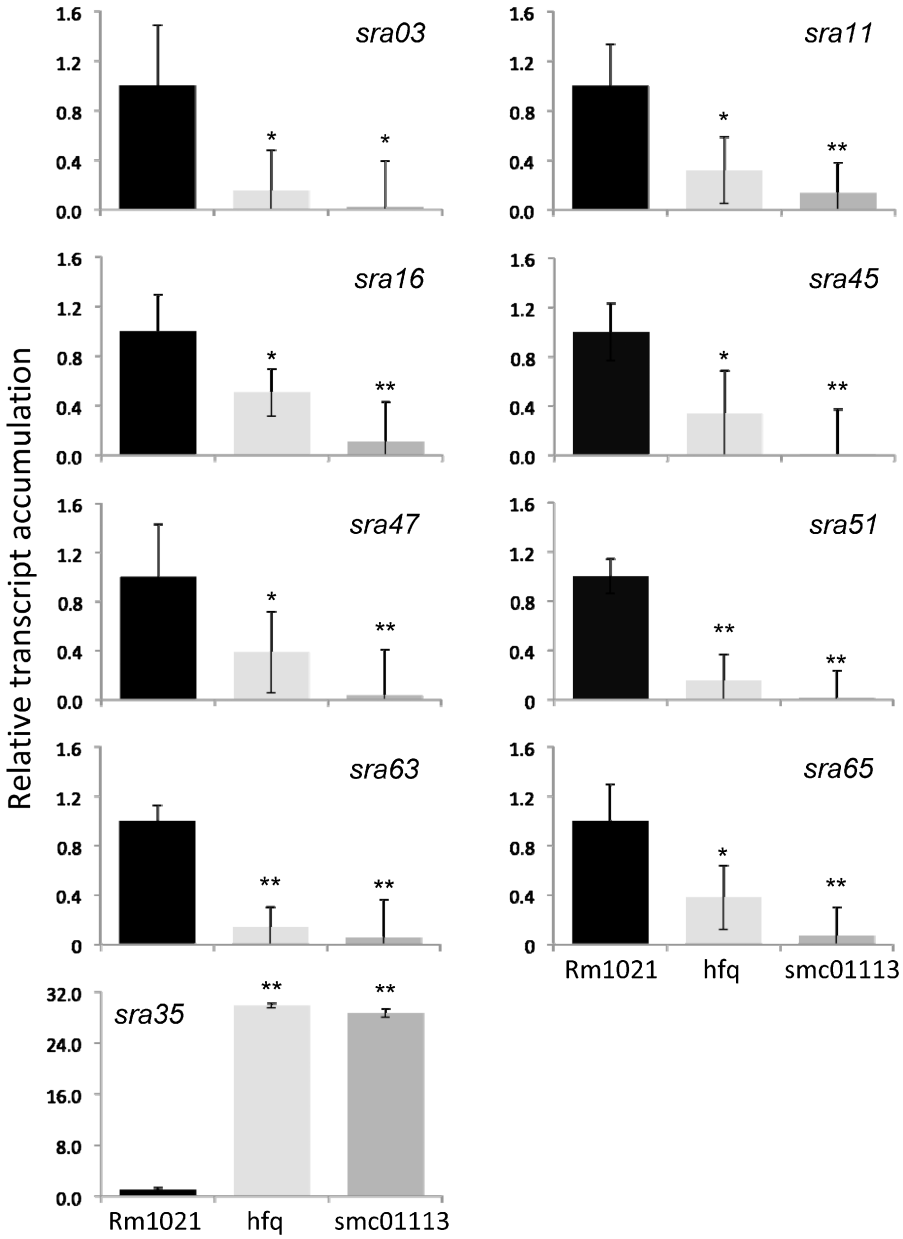

Figure 5. Loss of *hfq* or *SMc01113* similarly deregulates sRNA accumulation. The WT, *smc01113* and *hfq* mutants, were grown in the RMM till exponential (*sras03*, *11*, *16*, *45*, *47*, *51* and *65*) or stationary phase (*sra35* and *sra63*), RNA was extracted and qPCR was performed. Transcript accumulation in the WT strain was fixed to 1 and relative transcript abundance to WT in the two mutants were evaluated as in Figures 3 and 4. 'Single' and 'double asterisk' shows significant difference at $P \leq 0.05$ and $P \leq 0.01$, respectively.

Such mismatches in these sRNA–mRNA motifs may be analogous to eukaryotic miRNA target sites with seeds having G:U base pairs or single nucleotide bulges. These results suggest a large degree of conservation of these sites among closely related rhizobia and are thus interesting targets for future study of sRNA–mRNA interaction in rhizobia species.

Over-expression of sRNAs

In order to evaluate our predictions of sRNA-targets, we over-expressed sra16 and sra35 (Figure 7A and B) and evaluated the effects on the targets we initially predicted (Figure 6). Sra16 was down-regulated whereas sra35 was up-regulated in *hfq*- and *smc01113*- mutants, compared to WT Rm1021 (Figure 5). We tested the accumulation of 4 of the 10 predicted targets of sra16 and all the three predicted targets of sra35 (Figure 6). Putative target sites for sra16 and sra35 did not overlap with the ribosome binding sites. Three sra16-targets (*glnA*, *dppA2* and *katB*) were down-regulated in the sra16-over-producing strain (SRA16; Figure 7C, E and G) compared to WT, whereas the fourth (*dppA1*) did not show a difference between the two strains (data not shown). Over-expression of sra35 (SRA35; Figure 7) resulted in up-regulation of *glnA* (Figure 7D), whereas two other targets (*ppiA* and *atpD*) showed a trend of down-regulation (Figure 7F

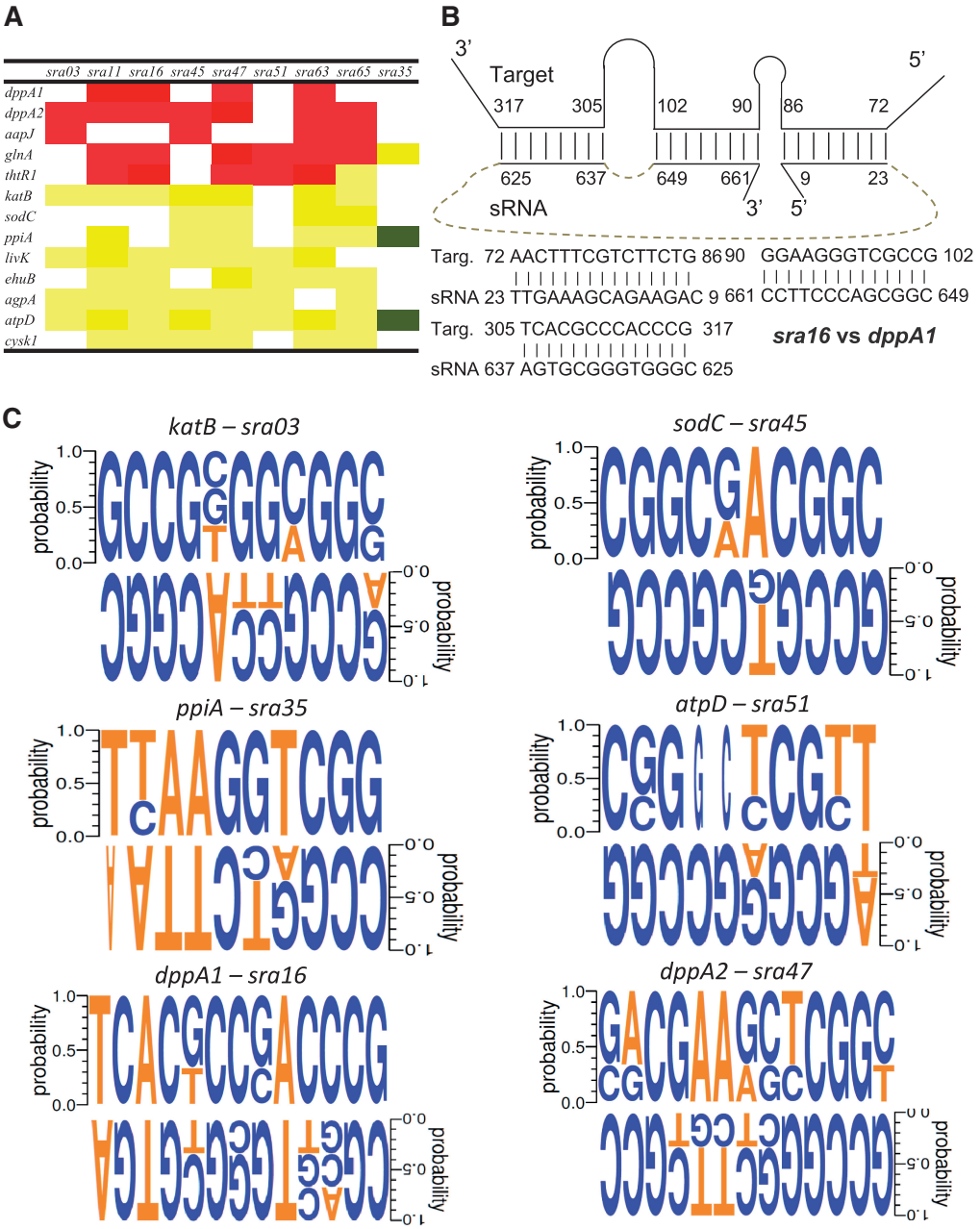

Figure 6. sRNA-target relation in *S. meliloti* and conservation of motifs in sRNAs and their possible targets in other related rhizobium species. (**A**) Correlation between the direction of accumulation of the sRNAs and their target in the *S. meliloti* strains mutated for *hfq* and *SMc01113*, as derived from Figures 3–5. Red and green shows inverse correlation, and yellow shows that the sRNAs and their predicted targets show accumulation in the similar direction. (**B**) Schematic representation of the three *sra16* binding sites in the *dppA1* gene. Sra16 was down-regulated whereas dppA1 was up-regulated in the *hfq* and *smc01113* mutants. (**C**) Conservation of the target-sRNA motif sequences in the three related rhizobium species. Weblogos were generated as described in the text.

and H). These results show that our predictions were identifying biologically relevant sRNA–mRNA interactions and that individual sRNAs have multiple targets. Sras16 and 35 differently regulated the accumulation of a common target, *glnA*: whereas *glnA* was down-regulated in bacteria over-expressing sra16, over-expression of sra35 up-regulated *glnA* expression, providing evidence that multiple sRNAs may regulate a given target differently, and that this regulation of expression of genes by sRNAs is complex. Thus, we were able to largely confirm the results of our predictions of sRNA-targets as well as target-sRNA relationships.

Potential interactions of SMc01113/YbeY with other components of sRNA machinery/RNA-degradation

Since our data suggested that the function of SMc01113/YbeY is related to Hfq, we next asked whether SMc01113/YbeY physically interacts with the Hfq protein or with

other members of the sRNA machinery and RNA-degradation. The sRNA-machinery and the components of RNA-degradosome are well characterized in *E. coli* (75). RNase E (*rne*) has been implicated as an integral component of the sRNA-machinery and degrades targets of sRNAs (10,13,76,77). RNase III (*rnc*), whose eukaryotic analogs are the DICER/dicer-like (DCL) proteins, participates in sRNA-mediated interaction, along with the RNase E, by performing initial cleavage (78,79). The PIWI domain of the AGO proteins has the RNase H- (*rnhB*) type of catalytic domain (32). The two other RNases comprising important component of the RNA-degradosome during stress adaptation are RNase R and PNPase; both of these have overlapping functions (80). Informed by the fact that sRNAs can determine which proteins bind to Hfq (81), we asked if we could detect an interaction of YbeY with Hfq and other RNases in the absence of bacterial sRNAs. We therefore conducted a directed yeast two-hybrid screen to determine if YbeY interacts physically with any of these components. However no significant interaction with any of these was observed regardless of whether YbeY was fused with either the *GAL4* activation or binding domains (Figure 8). This suggests that either (i) there is no physical interaction between YbeY/SMc01113 with Hfq and other RNases, or (ii) the interaction of YbeY/SMc01113 with these components is highly transient and/or may require additional facilitator proteins. Transient interaction with YbeY, mediated by other proteins would be consistent with the fact that there is a large difference in the number of Hfq protein molecules per cell [approximately 30–60000; (82)] relative to YbeY/SMc01113 (approximately 1000; Walker,G.C., unpublished data) molecules in a cell. The involvement of other co-factors in modulating protein–protein interactions during sRNA regulation is conceivable. For example, it has recently became apparent that the RraA regulatory protein modulates the helicase and RNA-binding activities of the RNA-degradosome (83). Furthermore, the association of the AGO–siRNA complex with AGO/binding proteins e.g. Giw1p for correct localization has been known (84). We also generated a *S. meliloti* double mutant of *hfq* and *smc01113*, which showed highly reduced growth and enhanced doubling time over both the single mutants (Figure 8). As discussed more fully below, these

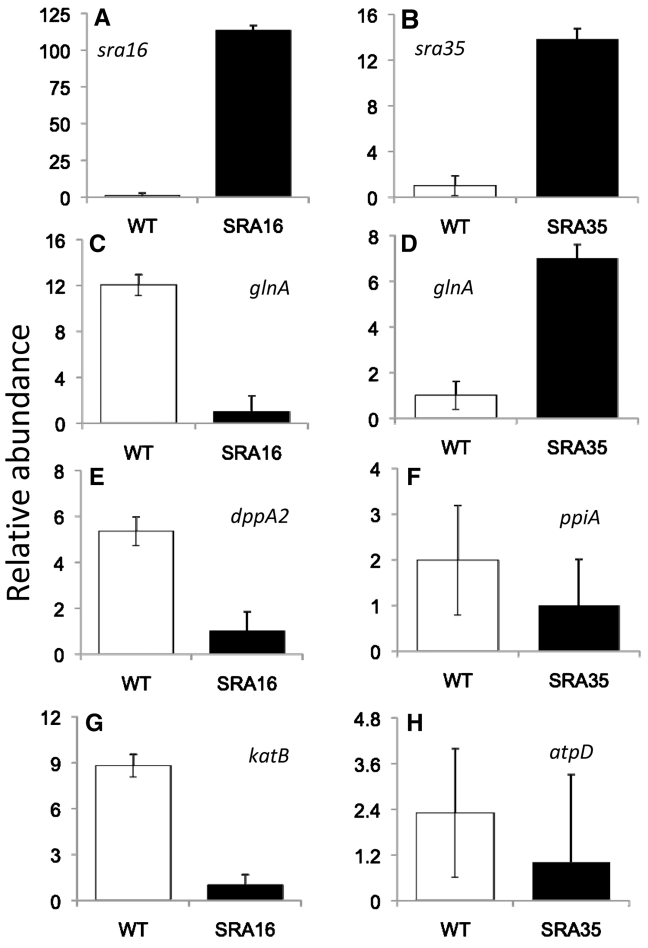

Figure 7. Over-expression of sRNAs affect accumulation of their predicted targets. (**A** and **B**) show the levels of over-expression of sra16 and 35, respectively. (**C**, **E** and **G**) show down-regulation of sra16 targets. (**D**, **F** and **H**) show the changes in putative sra35 targets. Sras 16 and 35 differently regulate the accumulation of common target, *glnA* (C and D). sRNA-targeting regions were predicted in the coding-region.

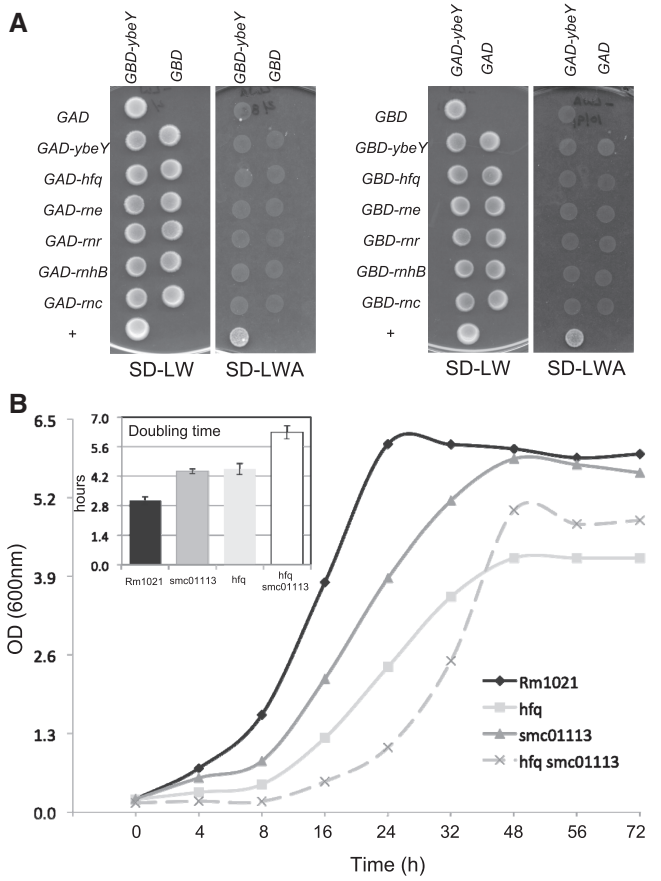

Figure 8. (**A**) Directed yeast two hybrid of *YbeY*. No interaction was detected. (**B**) Growth of the WT, *smc01113* and *hfq* single, and *smc01113 hfq* double mutant in LB-MC. Insert shows doubling time of WT and each of the three mutants in LB-MC.

observations are consistent with SMc01113/YbeY and Hfq playing functionally related roles in sRNA regulation but with YbeY playing certain Hfq-independent roles (81) as well.

DISCUSSION

Taken together, our results suggest that an important, previously unrecognized function for the SMc01113/YbeY proteins, members of the UPF0054 family which is present in most bacterial species, is to participate in sRNA regulation. The SMc01113/YbeY proteins show intriguing structural similarities to the MID domain of the AGO proteins, which has a capacity to capture the 5′-phosphate of sRNA seed regions. Loss of SMc01113 results into perturbed sRNA accumulation, deregulated mRNA levels and consequently, physiological and symbiotic defects. Furthermore, there are striking similarities in the symbiotic, physiological and molecular profiles of the *smc01113* and *hfq* mutants. Taken together, our results suggest that SMc01113 homologs may form a component of sRNA-pathway and facilitate effective outcomes of the sRNA–mRNA interaction: in particular, we suggest that SMc01113/YbeY proteins could help in recognizing the prokaryotic sRNAs during their Hfq mediated interaction with the target mRNAs. In addition, this study also gives insights into the fundamentals of sRNA-mediated regulatory basis of stress adaptation in *S. meliloti* by identifying multiple sRNA-target interactions. Although sRNAs have been identified in a few genomic studies of *S. meliloti* (16–18), little is known about their functions, targets and regulation. Understanding regulatory mechanisms in *S. meliloti* are of utmost importance from agricultural, economical, environmental perspectives as well as for their implications for other symbiotic and pathogenic interactions.

An estimated 82 sRNAs were originally predicted in *S. meliloti* (16–18), whereas the most recent genome-wide survey puts the number at 1125 [which included *trans*-encoded sRNAs, antisense RNAs, putative mRNA leader and sense transcripts, which may be degradation products of mRNAs; (85)]. Genome-wide searches in *E. coli* estimate that the number of sRNAs is ~2% of the number of protein coding genes (77,86). In bacteria, a major class of these sRNAs, as in eukaryotes, act by base pairing with the target mRNAs. In response to environmental stress, the expression of several sRNAs may be stabilized by the RNA chaperone Hfq, which facilitates the *trans*-annealing of the sRNAs to their target mRNA in an anti-sense manner (70,87). This is followed by the recognition of the sRNA-target complex by a multi-ribonucleoprotein complex, resulting in mRNA degradation (11,75,88) or mRNA stabilization, facilitated ribosome assembly and enhanced translation (10,13,64,69). However, these processes are still poorly understood, as little is known regarding the assembly of the sRNA–mRNA complex, except that Hfq is involved (11,13,81).

Conceptually, striking parallels exist between the prokaryotic and the eukaryotic sRNA-pathways in terms of synthesis of sRNAs, the presentation of sRNA and mRNAs onto the protein scaffold and the eventual outcomes of sRNA–mRNA interactions (13). Both, the sRNAs and the mRNAs bind to the Hfq protein, but at different sites (87). However, mechanistically, little is known about how the prokaryotic sRNAs are recognized and loaded on to the Hfq protein scaffold during their interaction with the target mRNAs. This is better understood in eukaryotes, where the si/miRNAs are loaded into AGOs, and then guided and assembled on to the targets by the RISC. The recognition of target is governed by the 'seed sequence' of the guide RNA strand associated with the MID domain of the AGO proteins that generates 'an enhanced affinity anchor site promoting fidelity in target recognition' (89). This stabilizes and guides the assembly of the active RISC (89). The structure of the MID domain associated with the dsRNA and ssDNA reveals that the MID-domain recognizes the 5′-phosphate specifically and anchors sRNAs on to the AGO/RISC (32,52). The 5′-end of the guide strand (of sRNA) could bind to the AGO in a stacked helical confirmation, both in the absence and presence of the partner mRNA strand, facilitating easy access for recognition of the mRNA target (32,52). Moreover, the MID domain has been implicated in protein–protein interactions: interactors such as Tas3 in *S. pombe* form a so-called 'Ago-hook' that binds the MID domain (90). Whether the binding of the protein interactors and sRNAs occur simultaneously, or, is mutually exclusive remains unclear. Different AGO family members have diverse functions in a given organism (91). The MID domains are central to the functional distinctions that define the various AGO protein family members (92). Proteins that perform analogous functions in the rhizobial or other bacterial sRNA-pathways have remained unknown. Our structural analysis studies suggest that SMc01113/YbeY and their bacterial homologs might perform analogous functions of anchoring the seed sRNAs, which can help the sRNA access the mRNA. SMc01113/YbeY proteins have >50% sequence similarity to MID domains, and furthermore have similar structural folds to those of both the prokaryotic and eukaryotic MID domains [Figure 1; (52)]. Moreover, our structural analysis suggests that they similarly have a capacity to bind to the sRNA seeds, and regulate accumulation of sRNAs and target mRNAs (Figures 1, 3–5). Our analyses suggest that, similarly to the corresponding residues of the MID domain, R59 and K61 may undergo key interactions with a phosphate at 5′-end of an RNA sequence. The structure of YbeY suggests that it could accommodate an internal phosphate at the 5′-end of the seed sequence instead of requiring that the 5′-phosphate be at the 5′-terminus as well as a 5′-phosphate at the terminus of a sRNA. This suggests that SMc01113/YbeY homologs have a function that is analogous to the functions of AGO MID domain, facilitating the recognition of the sRNA, thereby guiding it to the target mRNA/Hfq and the assembly of active complex.

YbeY differs from the AGO MID domain in one interesting aspect; a metal coordinated ion, postulated to be zinc (27), is present on the other side of the cleft opposite R59 and K61 (Supplementary Figure S1D), along with a

highly conserved His-triad (H114, H118 and H124; Supplementary Figure S1). These could act catalytically in a hydrolytic reaction. In contrast to speculations of YbeY being a peptidase or a protease (28), our docking of the RNA and analysis of YbeY structure suggests that it is plausible that a phosphodiester bond could be located close enough to the presumed coordinated water for hydrolysis to occur. Thus, it is possible that YbeY could contribute catalytically, like an RNase, to RNA cleavage events after binding to a guide RNA. We hypothesize a model where a sRNA substrate would bind to the positively charged surface of YbeY with help of a phosphate that plausibly interacts with K59, L61, L63 and other residues on the positively charged surface (Figure 1, Supplementary Figure S1) followed by a metal-coordinated hydrolysis of phosphodiester bond by the His-triad (Supplementary Figure S1). Such metal-dependent RNA-binding and RNase activity is established for nucleases for instance in *V. vulnificus* and *Xenopus laevis* (93,94). Indeed, the 5′-binding pocket in AGO, present at the junction of the MID and PIWI domains [the two domains now regarded as 'MID/PIWI lobe'; (95)], involves a metal ion coordinated to the C-terminal carboxylate of the AGO polypeptide and the first (5′) and third phosphates of the guide strand (95). It is conceivable that YbeY family proteins may have been evolutionarily like primitive AGO proteins, where catalysis and RNA binding was present in one structure. With the evolution of complex genetic switches, the catalytic capacity got separated into more, RNaseH-like PIWI of AGOs, whereas MID got specialized for RNA binding and protein–protein interaction functions. Indeed, YbeY shares some structural similarities with PIWI (e.g. for PIWI of 2NUB AGO, two structures superimpose with a significant RMSD of 3.06 Å; data not shown).

Our proposal that SMc01113/YbeY proteins participate in the sRNA-pathway strengthens the parallels between the AGO/RISC functions and the bacterial sRNA–mRNA assembly process that have been previously noted (13,81). Furthermore, as in the eukaryotes, our computational analysis suggests a 'one-to-many' sRNA-target relationship in *S. meliloti*. Thus, this study gives the first line of evidence that supports more than a conceptual similarity (13) between the prokaryotic and eukaryotic RNAi machinery. The involvement of other analogous facilitator proteins (currently unknown), however, cannot be ruled out. The next step will be to work out the biochemical properties of the SMc01113/YbeY protein so that we can better understand the mechanism of such action.

For most of the phenotypes we tested, along with the underlying molecular targets and sRNAs, the *smc01113* and *hfq* mutants displayed close similarity. Nevertheless, certain observations indicate that SMc01113 and Hfq also act independently. The difference in *thtR* levels in *smc01113* and *hfq* mutants is consistent with this possibility, as is the reduced growth of *smc01113 hfq* double mutant compared to the single mutants in a rich LB-MC medium. Recent work from Rasouly *et al.* (28) and from our lab in *E. coli* rRNAs (51) suggests an additional function for YbeY as the deletion of *ybeY* from *E. coli* results in defects in ribosome assembly and activity as well as in attenuation of processing of ends of rRNAs (51). Interestingly, mutating the R59 residue, which our docking studies suggest may participate in coordinating the PO_4 in a sRNA seed (Figure 1E–G) on a multicopy plasmid did not affect rRNA processing but did affect the growth rate. The possibility that a member of sRNA-pathway could potentially be involved in maturation of rRNA-ends is not at all surprising. For example, RNase E, central to sRNA pathway (11,13,77,81), is also necessary for maturation of ends of rRNA (96,97). In another example, PNPase is indispensable for 3′-end maturation of 23S rRNA, is a component for mRNA degradation is sRNAs' context, and is also a key regulator of sRNAs, (81,98). Even Hfq interacts with at least 24 ribosomal proteins (81), and also with RNase E and PNPase (81); in *E. coli*, YbeY shows genetic interactions with RNase E and PNPase (51). This overlap in function and sharing of components between sRNA-pathway and rRNA maturation and ribosome assembly may form a more general strategy of efficient post-transcriptional gene regulation by coupling translation and transcription (81). This warrants further investigation.

Our observations are also important because they verify the differential expression of genes in *S. meliloti hfq* proteome at the transcriptional level. Along with that, we have also identified several sRNAs that are Hfq-dependent. Mapping of these differentially accumulated sRNAs/mRNAs, the conservation of sRNA–mRNA complementary motifs in rhizobial species closely related to *S. meliloti* (Figure 6), and change in accumulation of predicted targets in bacteria over-expressing sRNAs (Figure 7) indicate functions for these sRNAs, a possible mode of regulation of these genes and a mechanistic relationship between the sRNAs and the targets. We have here further generated additional support for biological relevance of some *S. meliloti* sRNA–mRNA relationships (Figure 7) and identified other such candidate relationships for further investigation.

Our results also offer insights into the regulatory basis of stress adaptation in *S. meliloti*. Pathways such as transport system, oxidative stress resistance and nitrogen metabolism are coordinated and controlled by the sRNA-machinery. Along with identifying a new component, this study places the sRNA-pathway squarely in the centre of stress adaptation and regulation of plasticity of *S. meliloti* as the bacteria undergoes ecosystem shifts during its course of survival.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank the members of the Walker lab for the critical reading of the article. G.C.W. is an American Cancer Society Research Professor.

FUNDING

National Institutes of Health grant GM31030 (to G.C.W.); MIT Center for Environmental Health Sciences NIEHS P30 ES002109; National Institutes of Health fellowship F32 GM078966 (to B.K.M.). Funding for open access charge: National Institute of Health grant GM31030.

Conflict of interest statement. None declared.

REFERENCES

1. LeVier,K., Phillips,R.W., Grippe,V.K., Roop,R.M. and Walker,G.C. (2000) Similar requirements of a plant symbiont and a mammalian pathogen for prolonged intracellular survival. *Science*, **287**, 2492–2493.
2. Becker,A., Berges,H., Krol,E., Bruand,C., Ruberg,S., Capela,D., Lauber,E., Meilhoc,E., Ampe,F., de Bruijn,F.J. *et al*. (2004) Global changes in gene expression in Sinorhizobium meliloti 1021 under microoxic and symbiotic conditions. *Mol. Plant Microbe Interact*., **17**, 292–303.
3. Gibson,K.E., Kobayashi,H. and Walker,G.C. (2008) Molecular determinants of a symbiotic chronic infection. *Ann. Rev. Genet*., **42**, 413–441.
4. Jones,K.M., Kobayashi,H., Davies,B.W., Taga,M.E. and Walker,G.C. (2007) How rhizobial symbionts invade plants: the Sinorhizobium-Medicago model. *Nat. Rev. Microbiol*., **5**, 619–633.
5. Deakin,W.J. and Broughton,W.J. (2009) Symbiotic use of pathogenic strategies: rhizobial protein secretion systems. *Nat. Rev. Microbiol*., **7**, 312–320.
6. Natera,S.H., Guerreiro,N. and Djordjevic,M.A. (2000) Proteome analysis of differentially displayed proteins as a tool for the investigation of symbiosis. *Mol. Plant Microbe Interact*., **13**, 995–1009.
7. Ampe,F., Kiss,E., Sabourdy,F. and Batut,J. (2003) Transcriptome analysis of Sinorhizobium meliloti during symbiosis. *Genome Biol*., **4**, R15.
8. Djordjevic,M.A., Chen,H.C., Natera,S., Van Noorden,G., Menzel,C., Taylor,S., Renard,C., Geiger,O. and Weiller,G.F. (2003) A global analysis of protein expression profiles in Sinorhizobium meliloti: discovery of new genes for nodule occupancy and stress adaptation. *Mol. Plant Microbe Interact*., **16**, 508–524.
9. Barnett,M.J., Toman,C.J., Fisher,R.F. and Long,S.R. (2004) A dual-genome Symbiosis Chip for coordinate study of signal exchange and development in a prokaryote-host interaction. *Proc. Natl Acad. Sci. USA*, **101**, 16636–16641.
10. Waters,L.S. and Storz,G. (2009) Regulatory RNAs in bacteria. *Cell*, **136**, 615–628.
11. Aiba,H. (2007) Mechanism of RNA silencing by Hfq-binding small RNAs. *Curr. Opin. Microbiol*., **10**, 134–139.
12. Brennan,R.G. and Link,T.M. (2007) Hfq structure, function and ligand binding. *Curr. Opin. Microbiol*., **10**, 125–133.
13. Gottesman,S. (2005) Micros for microbes: non-coding regulatory RNAs in bacteria. *Trends Genet*., **21**, 399–404.
14. Storz,G., Opdyke,J.A. and Zhang,A.X. (2004) Controlling mRNA stability and translation with small, noncoding RNAs. *Curr. Opin. Microbiol*., **7**, 140–144.
15. Vogel,J. (2009) A rough guide to the non-coding RNA world of Salmonella. *Mol. Microbiol*., **71**, 1–11.
16. del Val,C., Rivas,E., Torres-Quesada,O., Toro,N. and Jimenez-Zurdo,J.I. (2007) Identification of differentially expressed small non-coding RNAs in the legume endosymbiont Sinorhizobium meliloti by comparative genomics. *Mol. Microbiol*., **66**, 1080–1091.
17. Ulve,V.M., Sevin,E.W., Cheron,A. and Barloy-Hubler,F. (2007) Identification of chromosomal alpha-proteobacterial small RNAs by comparative genome analysis and detection in Sinorhizobium meliloti strain 1021. *BMC Genom*., **8**, 467.
18. Valverde,C., Livny,J., Schluter,J.P., Reinkensmeier,J., Becker,A. and Parisi,G. (2008) Prediction of Sinorhizobium meliloti sRNA genes and experimental detection in strain 2011. *BMC Genom*., **9**, 416.
19. Barra-Bily,L., Pandey,S.P., Trautwetter,A., Blanco,C. and Walker,G.C. (2010) The Sinorhizobium meliloti RNA chaperone Hfq mediates symbiosis of S. meliloti and alfalfa. *J. Bacteriol*., **192**, 1710–1718.
20. Barra-Bily,L., Fontenelle,C., Jan,G., Flechard,M., Trautwetter,A., Pandey,S.P., Walker,G.C. and Blanco,C. (2010) Proteomic alterations explain phenotypic changes in Sinorhizobium meliloti lacking the RNA chaperone Hfq. *J. Bacteriol*., **192**, 1719–1729.
21. Gao,M., Barnett,M.J., Long,S.R. and Teplitski,M. (2010) Role of the *Sinorhizobium meliloti* global regulator Hfq in gene regulation and symbiosis. *Mol. Plant Microbe Interact*., **23**, 355–365.
22. Torres-Quesada,O., Oruezabal,R.I., Peregrina,A., Jofre,E., Lloret,J., Rivilla,R., Toro,N. and Jimenez-Zurdo,J.I. (2010) The *Sinorhizobium meliloti* RNA chaperone Hfq influences central carbon metabolism and the symbiotic interaction with alfalfa. *BMC Microbiol*., **10**, 71.
23. Davies,B.W. and Walker,G.C. (2007) Identification of novel *Sinorhizobium meliloti* mutants compromised for oxidative stress protection and symbiosis. *J. Bacteriol*., **189**, 2110–2113.
24. Davies,B.W. and Walker,G.C. (2008) A highly conserved protein of unknown function is required by *Sinorhizobium meliloti* for symbiosis and environmental stress protection. *J. Bacteriol*., **190**, 1118–1123.
25. Oganesyan,V., Busso,D., Brandsen,J., Chen,S., Jancarik,J., Kim,R. and Kim,S.H. (2003) Structure of the hypothetical protein AQ_1354 from Aquifex aeolicus. *Acta Crystallogr. D Biol. Crystallogr*., **59**, 1219–1223.
26. Penhoat,C.H., Li,Z., Atreya,H.S., Kim,S., Yee,A., Xiao,R., Murray,D., Arrowsmith,C.H. and Szyperski,T. (2005) NMR solution structure of Thermotoga maritima protein TM1509 reveals a Zn-metalloprotease-like tertiary structure. *J. Struct. Funct. Genomics*, **6**, 51–62.
27. Zhan,C., Fedorov,E.V., Shi,W., Ramagopal,U.A., Thirumuruhan,R., Manjasetty,B.A., Almo,S.C., Fiser,A., Chance,M.R. and Fedorov,A.A. (2005) The ybeY protein from *Escherichia coli* is a metalloprotein. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun*., **61**, 959–963.
28. Rasouly,A., Davidovich,C. and Ron,E.Z. (2010) The heat shock protein YbeY is required for optimal activity of the 30S ribosomal subunit. *J. Bacteriol*., **192**, 4592–4596.
29. Rasouly,A., Schonbrun,M., Shenhar,Y. and Ron,E.Z. (2009) YbeY, a heat shock protein involved in translation in Escherichia coli. *J. Bacteriol*., **191**, 2649–2655.
30. Mallory,A.C., Hinze,A., Tucker,M.R., Bouche,N., Gasciolli,V., Elmayan,T., Lauressergues,D., Jauvion,V., Vaucheret,H. and Laux,T. (2009) Redundant and specific roles of the ARGONAUTE proteins AGO1 and ZLL in development and small RNA-directed gene silencing. *PLoS Genet*., **5**, e1000646.
31. Nowotny,M. and Yang,W. (2009) Structural and functional modules in RNA interference. *Curr. Opin. Struct. Biol*., **19**, 286–293.
32. Yuan,Y.R., Pei,Y., Ma,J.B., Kuryavyi,V., Zhadina,M., Meister,G., Chen,H.Y., Dauter,Z., Tuschl,T. and Patel,D.J. (2005) Crystal structure of A. aeolicus argonaute, a site-specific DNA-guided endoribonuclease, provides insights into RISC-mediated mRNA cleavage. *Mol. Cell*, **19**, 405–419.
33. Wang,Y., Juranek,S., Li,H., Sheng,G., Tuschl,T. and Patel,D.J. (2008) Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. *Nature*, **456**, 921–926.
34. Holm,L. and Rosenstrom,P. (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res*., **38(Suppl.)**, W545–W549.
35. Holm,L. and Park,J. (2000) DaliLite workbench for protein structure comparison. *Bioinformatics*, **16**, 566–567.
36. Emsley,P. and Cowtan,K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr*., **60**, 2126–2132.
37. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr*., **66**, 486–501.
38. Pettit,F.K., Bare,E., Tsai,A. and Bowie,J.U. (2007) HotPatch: a statistical a pproach to finding biologically relevant features on protein surfaces. *J. Mol. Biol*., **369**, 863–879.

39. Brennecke,J., Stark,A., Russell,R.B. and Cohen,S.M. (2005) Principles of microRNA-target recognition. *PLoS Biol.*, **3**, e85.
40. Busch,A., Richter,A.S. and Backofen,R. (2008) IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. *Bioinformatics*, **24**, 2849–2856.
41. Lewis,B.P., Burge,C.B. and Bartel,D.P. (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell*, **120**, 15–20.
42. Pichon,C. and Felden,B. (2008) Small RNA gene identification and mRNA target predictions in bacteria. *Bioinformatics*, **24**, 2807–2813.
43. Pfeiffer,V., Papenfort,K., Lucchini,S., Hinton,J.C. and Vogel,J. (2009) Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nat. Struct. Mol. Biol.*, **16**, 840–846.
44. Pandey,S.P., Shahi,P., Gase,K. and Baldwin,I.T. (2008) Herbivory-induced changes in the small-RNA transcriptome and phytohormone signaling in Nicotiana attenuata. *Proc. Natl Acad. Sci. USA*, **105**, 4559–4564.
45. Crooks,G.E., Hon,G., Chandonia,J.M. and Brenner,S.E. (2004) WebLogo: a sequence logo generator. *Genome Res.*, **14**, 1188–1190.
46. Schneider,T.D. and Stephens,R.M. (1990) Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res.*, **18**, 6097–6100.
47. Bernett,M.J., Oke,V. and Long,S.R. (2000) New genetic tools for use in the Rhizobiaceae and other bacteria. *Biotechniques*, **29**, 240–245.
48. Sherman,F. (1991) Getting started with yeast. *Methods Enzymol.*, **194**, 3–21.
49. James,P., Halladay,J. and Craig,E.A. (1996) Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics*, **144**, 1425–1436.
50. Gietz,R.D., Schiestl,R.H., Willems,A.R. and Woods,R.A. (1995) Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast*, **11**, 355–360.
51. Davies,B.W., Köhrer,C., Jacob,A.I., Simmons,L.A., Zhu,J., Aleman,L.M., Rajbhandhari,U.L. and Walker,G.C. (2010) Role of Escherichia coli YbeY, a highly conserved protein, in rRNA processing. *Mol. Microbiol.*, **78**, 506–518.
52. Boland,A., Tritschler,F., Heimstadt,S., Izaurralde,E. and Weichenrieder,O. (2010) Crystal structure and ligand binding of the MID domain of a eukaryotic Argonaute protein. *EMBO Rep.*, **11**, 522–527.
53. Comeau,S.R., Gatchell,D.W., Vajda,S. and Camacho,C.J. (2004) ClusPro: an automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics*, **20**, 45–50.
54. Kozakov,D., Brenke,R., Comeau,S.R. and Vajda,S. (2006) PIPER: an FFT-based protein docking program with pairwise potentials. *Proteins*, **65**, 392–406.
55. Nogales,J., Dominguez-Ferreras,A., Amaya-Gomez,C.V., van Dillewijn,P., Cuellar,V., Sanjuan,J., Olivares,J. and Soto,M.J. (2010) Transcriptome profiling of a Sinorhizobium meliloti fadD mutant reveals the role of rhizobactin 1021 biosynthesis and regulation genes in the control of swarming. *BMC Genomics*, **11**, 157.
56. Chang,C., Damiani,I., Puppo,A. and Frendo,P. (2009) Redox changes during the legume-rhizobium symbiosis. *Mol. Plant*, **2**, 370–377.
57. Gage,D.J. and Long,S.R. (1998) alpha-Galactoside uptake in Rhizobium meliloti: isolation and characterization of agpA, a gene encoding a periplasmic binding protein required for melibiose and raffinose utilization. *J. Bacteriol.*, **180**, 5739–5748.
58. Justice,S.S., Hunstad,D.A., Harper,J.R., Duguay,A.R., Pinkner,J.S., Bann,J., Frieden,C., Silhavy,T.J. and Hultgren,S.J. (2005) Periplasmic peptidyl prolyl cis-trans isomerases are not essential for viability, but SurA is required for pilus biogenesis in Escherichia coli. *J. Bacteriol.*, **187**, 7680–7686.
59. Shaw,P.E. (2002) Peptidyl-prolyl isomerases: a new twist to transcription. *EMBO Rep.*, **3**, 521–526.
60. Jebbar,M., Sohn-Bosser,L., Bremer,E., Bernard,T. and Blanco,C. (2005) Ectoine-induced proteins in Sinorhizobium meliloti include an Ectoine ABC-type transporter involved in osmoprotection and ectoine catabolism. *J. Bacteriol.*, **187**, 1293–1304.
61. Hanekop,N., Hoing,M., Sohn-Bosser,L., Jebbar,M., Schmitt,L. and Bremer,E. (2007) Crystal structure of the ligand-binding protein EhuB from Sinorhizobium meliloti reveals substrate recognition of the compatible solutes ectoine and hydroxyectoine. *J. Mol. Biol.*, **374**, 1237–1250.
62. Carter,R.A., Yeoman,K.H., Klein,A., Hosie,A.H., Sawers,G., Poole,P.S. and Johnston,A.W. (2002) dpp genes of Rhizobium leguminosarum specify uptake of delta-aminolevulinic acid. *Mol. Plant Microbe Interact.*, **15**, 69–74.
63. Walshaw,D.L. and Poole,P.S. (1996) The general L-amino acid permease of Rhizobium leguminosarum is an ABC uptake system that also influences efflux of solutes. *Mol. Microbiol.*, **21**, 1239–1252.
64. Hammer,B.K. and Bassler,B.L. (2007) Regulatory small RNAs circumvent the conventional quorum sensing pathway in pandemic Vibrio cholerae. *Proc. Natl Acad. Sci. USA*, **104**, 11145–11149.
65. Sharma,C.M., Darfeuille,F., Plantinga,T.H. and Vogel,J. (2007) A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites. *Genes Dev.*, **21**, 2804–2817.
66. Pandey,S.P. and Baldwin,I.T. (2008) Silencing RNA-directed RNA polymerase 2 increases the susceptibility of Nicotiana attenuata to UV in the field and in the glasshouse. *Plant J.*, **54**, 845–862.
67. Doench,J.G. and Sharp,P.A. (2004) Specificity of microRNA target selection in translational repression. *Genes Dev.*, **18**, 504–511.
68. Guillier,M. and Gottesman,S. (2008) The 5′ end of two redundant sRNAs is involved in the regulation of multiple targets, including their own regulator. *Nucleic Acids Res.*, **36**, 6781–6794.
69. Prevost,K., Salvail,H., Desnoyers,G., Jacques,J.F., Phaneuf,E. and Masse,E. (2007) The small RNA RyhB activates the translation of shiA mRNA encoding a permease of shikimate, a compound involved in siderophore synthesis. *Mol. Microbiol.*, **64**, 1260–1273.
70. Soper,T.J. and Woodson,S.A. (2008) The rpoS mRNA leader recruits Hfq to facilitate annealing with DsrA sRNA. *RNA*, **14**, 1907–1917.
71. Friedman,R.C., Farh,K.K., Burge,C.B. and Bartel,D.P. (2009) Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.*, **19**, 92–105.
72. Gu,J., Fu,H., Zhang,X. and Li,Y. (2007) Identifications of conserved 7-mers in 3'-UTRs and microRNAs in Drosophila. *BMC Bioinformatics*, **8**, 432.
73. Kheradpour,P., Stark,A., Roy,S. and Kellis,M. (2007) Reliable prediction of regulator targets using 12 Drosophila genomes. *Genome Res.*, **17**, 1919–1931.
74. Provost,P., Barat,C., Plante,I. and Tremblay,M.J. (2006) HIV-l and the microRNA-guided silencing pathway: an intricate and multifaceted encounter. *Virus Res.*, **121**, 107–115.
75. Carpousis,A.J. (2007) The RNA degradosome of Escherichia coli: an mRNA-degrading machine assembled on RNase E. *Annu. Rev. Microbiol.*, **61**, 71–87.
76. Kime,L., Jourdan,S.S., Stead,J.A., Hidalgo-Sastre,A. and McDowall,K.J. (2009) Rapid cleavage of RNA by RNase E in the absence of 5'-monophosphate stimulation. *Mol. Microbiol*, **76**, 590–604.
77. Gottesman,S. (2004) The small RNA regulators of Escherichia coli: roles and mechanisms. *Ann. Rev. Microbiol.*, **58**, 303–328.
78. Afonyushkin,T., Vecerek,B., Moll,I., Blasi,U. and Kaberdin,V.R. (2005) Both RNase E and RNase III control the stability of sodB mRNA upon translational inhibition by the small regulatory RNA RyhB. *Nucleic Acids Res.*, **33**, 1678–1689.
79. Portier,C., Dondon,L., Grunberg-Manago,M. and Regnier,P. (1987) The first step in the functional inactivation of the Escherichia coli polynucleotide phosphorylase messenger is a ribonuclease III processing at the 5′ end. *EMBO J.*, **6**, 2165–2170.
80. Chen,C. and Deutscher,M.P. (2010) RNase R is a highly unstable protein regulated by growth phase and stress. *RNA*, **16**, 667–672.
81. Lee,T. and Feig,A.L. (2009) The RNA–Protein Complexes of *E. coli* Hfq: Form and Function. *Non-Protein Coding RNAs*, Vol. 13. Springer, Berlin Heidelberg, pp. 249–271.

82. Kajitani,M., Kato,A., Wada,A., Inokuchi,Y. and Ishihama,A. (1994) regulation of the Escherichia-coli hfq gene encoding the host factor for phage q-beta. *J. Bacteriol.*, **176**, 531–534.
83. Gorna,M.W., Pietras,Z., Tsai,Y.C., Callaghan,A.J., Hernandez,H., Robinson,C.V. and Luisi,B.F. (2010) The regulatory protein RraA modulates RNA-binding and helicase activities of the E. coli RNA degradosome. *RNA*, **16**, 553–562.
84. Noto,T., Kurth,H.M., Kataoka,K., Aronica,L., DeSouza,L.V., Siu,K.W., Pearlman,R.E., Gorovsky,M.A. and Mochizuki,K. (2010) The Tetrahymena argonaute-binding protein Giw1p directs a mature argonaute-siRNA complex to the nucleus. *Cell*, **140**, 692–703.
85. Schluter,J.P., Reinkensmeier,J., Daschkey,S., Evguenieva-Hackenberg,E., Janssen,S., Janicke,S., Becker,J.D., Giegerich,R. and Becker,A. (2010) A genome-wide survey of sRNAs in the symbiotic nitrogen-fixing alpha-proteobacterium Sinorhizobium meliloti. *BMC Genom.*, **11**, 245.
86. Hershberg,R., Altuvia,S. and Margalit,H. (2003) A survey of small RNA-encoding genes in Escherichia coli. *Nucleic Acids Res.*, **31**, 1813–1820.
87. Link,T.M., Valentin-Hansen,P. and Brennan,R.G. (2009) Structure of Escherichia coli Hfq bound to polyriboadenylate RNA. *Proc. Natl Acad. Sci. USA*, **106**, 19292–19297.
88. Masse,E., Salvail,H., Desnoyers,G. and Arguin,M. (2007) Small RNAs controlling iron metabolism. *Curr. Opin. Microbiol.*, **10**, 140–145.
89. Parker,J.S., Parizotto,E.A., Wang,M., Roe,S.M. and Barford,D. (2009) Enhancement of the seed-target recognition step in RNA silencing by a PIWI/MID domain protein. *Mol. Cell*, **33**, 204–214.
90. Till,S., Lejeune,E., Thermann,R., Bortfeld,M., Hothorn,M., Enderle,D., Heinrich,C., Hentze,M.W. and Ladurner,A.G. (2007) A conserved motif in Argonaute-interacting proteins mediates functional interactions through the Argonaute PIWI domain. *Nat. Struct. Mol. Biol.*, **14**, 897–903.
91. Correa,R.L., Steiner,F.A., Berezikov,E. and Ketting,R.F. (2010) MicroRNA-directed siRNA biogenesis in Caenorhabditis elegans. *PLoS Genet.*, **6**, e1000903.
92. Djuranovic,S., Zinchenko,M.K., Hur,J.K., Nahvi,A., Brunelle,J.L., Rogers,E.J. and Green,R. (2010) Allosteric regulation of Argonaute proteins by miRNAs. *Nat. Struct. Mol. Biol.*, **17**, 144–150.
93. Hake,L.E., Mendez,R. and Richter,J.D. (1998) Specificity of RNA binding by CPEB: requirements for RNA recognition motifs and a novel Zinc finger. *Mol. Cell. Biol.*, **18**, 685–693.
94. Li,C., Hor,L., Chang,Z., Tsai,L., Yang,W. and Yuan,H.S. (2003) DNA binding and cleavage by the periplasmic nuclease Vvn: a novel structure with a known active site. *EMBO J.*, **22**, 4014–4025.
95. Parker,J.S. (2010) How to slice: snapshots of Argonaute in action. *Silence*, **1**, 3.
96. Klein,F. and Evguenieva-Hackenberg,E. (2002) RNase E is involved in 5′-end 23S rRNA processing in alpha-Proteobacteria. *Biochem. Biophys. Res. Commun.*, **299**, 780–786.
97. Li,Z., Pandit,S. and Deutscher,M.P. (1999) RNase G (CafA protein) and RNase E are both required for the 5′ maturation of 16S ribosomal RNA. *EMBO J.*, **18**, 2878–2885.
98. Andrade,J.M. and Arraiano,C.M. (2008) PNPase is a key player in the regulation of small RNAs that control the expression of outer membrane proteins. *RNA*, **14**, 543–551.