Three Ribosomal Operons of Escherichia coli Contain Genes Encoding Small RNAs That Interact With Hfq and CsrA in vitro

Stenum, Thomas Søndergaard; Kongstad, Mette; Holmqvist, Erik; Kallipolitis, Birgitte; Svenningsen, Sine Lo; Sørensen, Michael Askvad

Published in:
Frontiers in Microbiology

DOI:
10.3389/fmicb.2021.625585

Publication date:
2021

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY

Citation for published version (APA):
Stenum, T. S., Kongstad, M., Holmqvist, E., Kallipolitis, B., Svenningsen, S. L., & Sørensen, M. A. (2021). Three Ribosomal Operons of Escherichia coli Contain Genes Encoding Small RNAs That Interact With Hfq and CsrA in vitro. Frontiers in Microbiology, 12, [625585]. https://doi.org/10.3389/fmicb.2021.625585
Three out of the seven ribosomal RNA operons in *Escherichia coli* end in dual terminator structures. Between the two terminators of each operon is a short sequence that we report here to be an sRNA gene, transcribed as part of the ribosomal RNA primary transcript by read-through of the first terminator. The sRNA genes (rrA, rrB and rrF) from the three operons (rrnA, rrnB and rrnD) are more than 98% identical, and pull-down experiments show that their transcripts interact with Hfq and CsrA. Deletion of rrA, B, F, as well as overexpression of rrB, only modestly affect known CsrA-regulated phenotypes like biofilm formation, pgaA translation and glgC translation, and the role of the sRNAs *in vivo* may not yet be fully understood. Since RrA, B, F are short-lived and transcribed along with the ribosomal RNA components, their concentration reflects growth-rate regulation at the ribosomal RNA promoters and they could function to fine-tune other growth-phase-dependent processes in the cell. The primary and secondary structure of these small RNAs are conserved among species belonging to different genera of Enterobacteriales.

**Keywords:** sRNA, CsrA, Hfq, ribosomal RNA operon, dual terminators

**INTRODUCTION**

Bacterial small regulatory RNAs (sRNA) are major post-transcriptional regulators of gene expression. Mechanistically, the majority of these sRNAs act by base paring to complementary sequences in mRNA targets, thereby altering translation initiation rates and/or mRNA stability (Wagner and Romby, 2015). Association rates between sRNAs and their target RNAs are often strongly increased by the presence of the homohexameric RNA chaperone Hfq, which binds both RNAs and facilitates base pairing (Santiago-Frangos and Woodson, 2018). Many sRNAs are involved in the rapid reorganization of bacterial gene expression as a response to various types of stresses (recently reviewed in Holmqvist and Wagner, 2017). However, sRNAs that are expressed in the absence of an acute stress have also been described, including anti-toxin sRNAs (reviewed in Brantl and Jahn, 2015), the sRNAs ChiX and Spot 42 which regulate different aspects of carbohydrate metabolism (Møller et al., 2002; Rasmussen et al., 2009; Beisel and Storz, 2011) and MgrR (Moon and Gottesman, 2009), a regulator of lipopolysaccharide composition. The sRNAs...
CarB and CsrC are expressed in response to the accumulation of end-metabolism products at the entry to stationary phase (Lawhon et al., 2002; Gonzalez Chavez et al., 2010). In contrast to base-pairing sRNAs, CarB/C act by sequestering a single protein target, the global regulator CsrA. This small (7 kDa) homodimeric RNA-binding protein acts by binding at or close to ribosome binding sites (RBS) in a myriad of different mRNAs (Potts et al., 2017). Targets of CsrA include mRNAs encoding proteins involved in carbon metabolism (Liu et al., 1995), biofilm formation (Jackson et al., 2002), motility (Wei et al., 2001), quorum sensing, and virulence (Altier et al., 2002). CarB and CsrC function by mimicking CsrA targets and carry ~18 and nine motifs for CsrA binding, respectively (Liu et al., 1997; Weilbacher et al., 2003). As a consequence, they antagonize CsrA by sequestering it away from its lower-affinity mRNAs targets, thereby decreasing the effective concentration of CsrA. More recent, similar activities on CsrA have been described for two additional sRNAs in *E. coli*: McAS (Jørgensen et al., 2013) and GadY (Parker et al., 2017), both of which are believed to contain two binding sites for CsrA. Unlike CarB/C, both McAS and GadY also regulate gene expression independent of CsrA (Opdyke et al., 2004; Jørgensen et al., 2012; Thomason et al., 2012).

While the majority of characterized sRNAs from *E. coli* are transcribed from intergenic regions (IGRs) under the control of a dedicated promoter, several reports suggest that a substantial number of sRNAs are generated from 5′ or 3′ untranslated regions (UTRs) by RNase-dependent mRNA processing (Kawano et al., 2005; Chao et al., 2012; Miyakoshi et al., 2015). Additionally, the glyW-cysT-leuZ transcript, which is processed to give rise to tRNA<sup>β</sup><sup>W</sup>, tRNA<sup>cys</sup>T and tRNA<sup>leu</sup>Z, also generates the sRNA 3′ETS<sup>leu</sup>Z (Lalouna et al., 2015). This sRNA, which is generated by RNase E-dependent processing, base-pairs to two other sRNAs, RyhB and RybB. The pairing neutralizes transcriptional noise from the ryhB and rybb genes and counteracts potential regulatory outcomes of inadvertent expression of the corresponding sRNAs. The 5′-ETS<sup>leu</sup>Z is the first functional tRNA-derived fragment (tRF) described in bacteria. However, numerous tRFs have been reported in eukaryotes, where they control multiple different cellular processes, including genome stability (Martinez et al., 2017; Schorn et al., 2017), cell-cell signaling (Baglio et al., 2015), response to viral infection (Yeung et al., 2009) and stress responses (Emara et al., 2010; Saikia et al., 2014).

In the present study, we have investigated the family of so-called tRNA-linked repeats (TLRs) from *E. coli*. The TLRs are a class of sequences located in tRNA or ribosomal RNA (rRNA) operons. Since the first description of TLRs in 1978 (Egan and Landy, 1978), a total of 22 TLR genes have been identified (Rudd, 1999), which are distributed between ten different loci on the *E. coli* K-12 chromosome, each locus harboring one to five TLRs. A striking feature common to all TLRs is that 18–19 bp of their 3′-end is identical to the 3′-end of the tRNA or rRNA gene that is located immediately upstream of the TLR (Figure 1). Regarding TLR functionality, one of the TLRs found in the pre-tRNA transcript tyrT-tyrV was initially reported to be involved in recovery from amino acid starvation ( Böl sl and Kersten, 1991). However, this claim was later retracted as the phenotype was shown to originate from a nearby open reading frame ( Böl sl and Kersten, 1994). Thus far, the TLRs have no known function. In the following, we present evidence that the three TLRs located downstream from rRNA operon A (rrA), B (rrB) and D (rrF), respectively, are transcribed, processed, and bind the post-transcriptional regulators Hfq and CsrA. We present evidence that these novel sRNAs may act as regulators to fine-tune CsrA activity.

**MATERIALS AND METHODS**

**Culture Growth and Media**

The study was carried out in *E. coli* K-12 MAS1081 (MG1655 rph<sup>+</sup> gatC<sup>+</sup> glpR<sup>+</sup>). All strains used in the study are listed in **Supplementary Table 1**. Unless otherwise noted, all cultures were grown in MOPS minimal medium (Neidhardt et al., 1974) at 37°C shaking at 160 rpm and were grown exponentially for at least ten generations before start of the experiment to obtain balanced growth. Antibiotics were added as described for each experiment.

**RNA Purification and Northern Blotting**

Culture aliquots were harvested into 1/4 vol ice-cold stop solution (95% ethanol, 5% phenol) (Bernstein et al., 2002). Subsequently RNA was purified using hot phenol and flash freezing in liquid nitrogen as in Fessler et al. (2020). Briefly: Stopped culture aliquots were centrifuged 2 min at 20,000 g and resuspended in 0.1 vol cold 0.3 M sucrose, 0.01 M NaOAc pH 4.5 followed by addition of 0.1 vol 2% SDS 0.01 M NaOAc pH 4.5. Phenol (saturated with water) was added to the liquid phase at a 1:1 ratio, the tubes were vortexed and incubated 3 min at 65°C. After freezing 15 sec in liquid N<sub>2</sub> and centrifugation at 20,000 g for 5 min, the water phase was transferred to new tubes and the phenol extraction was repeated. If the RNA was used in an enzymatic reaction after purification, a chloroform extraction step was included. The RNA was ethanol-precipitated, washed by 96% ethanol, air dried at room temperature and dissolved in 10 mM NaOHAc, 1 mM EDTA. For northern blots, RNA was mixed 1:1 with loading buffer (0.1 M NaOAc (pH 5.0), 8 M urea, 0.05% (w/v) bromophenol blue and 0.05% (w/v) xylene cyanol) and size-separated on denaturing 0.4 mm thick polyacrylamide gels using 1 × TBE buffer (90 mM Tris, 90 mM boric acid and 2 mM EDTA). The RNA was electroblotted onto Hybond-N membranes (GE Healthcare) (1.5 V/cm, 1.5 h) in 10 mM NaOHAc, 1 mM EDTA. For northern blots, RNA was mixed 1:1 with loading buffer (0.1 M NaOAc (pH 5.0), 8 M urea, 0.05% (w/v) bromophenol blue and 0.05% (w/v) xylene cyanol) and size-separated on denaturing 0.4 mm thick polyacrylamide gels using 1 × TBE buffer (90 mM Tris, 90 mM boric acid and 2 mM EDTA). The RNA was electroblotted onto Hybond-N membranes (GE Healthcare) (1.5 V/cm, 1.5 h) in 40 mM Tris-acetate (pH 8.1), 2 mM EDTA. After UV-crosslinking (0.12 J/cm<sup>2</sup>) the membranes were pre-hybridized (1 h, rotating at 42°C) in hybridization solution (0.9 M NaCl, 0.05 M NaH<sub>2</sub>PO<sub>4</sub> (pH 7.7), 5 mM EDTA, 5 × Denhardt’s solution (0.1% BSA, 0.1% Ficoll 400, 0.1% polyvinylpyrrolidone), 0.5% (w/v) SDS and 100 mg/ml sheared, denatured salmon sperm DNA). Probe hybridization was done by adding 30 pmol of oligo-DNA, 5′-end labeled with <sup>32</sup>P (overnight, rotating at 42°C). Subsequently, membranes were washed several times with 0.3 M NaCl, 30 mM sodium citrate, 0.1% SDS at 42°C. Radioactive signals were quantified on a PhosphorImager (Typhoon-GE Healthcare) using ImageQuant software as previously described.
Promega) and left overnight rotating at 4 
µ
 tubes containing 300 cleared lysate, and the remaining lysate was transferred to fresh 
60 min). Total RNA was prepared from an aliquot of the 
NaCl, 5% glycerol), lysed by sonication and centrifuged (20,000 
suspended in 2 ml lysis buffer (50 mM Tris-HCl pH 7.5, 50 mM 
cells were pelleted and washed in medium without biotin, re-
washed four times in lysis buffer as above and RNA was harvested 
biotin, at 37 
◦
µ
15 

were grown exponentially in MOPS medium supplemented with; 
production of sufficient BirA was achieved without induction of 
brA 
of Hfq_bio we introduced the plasmid pBirA where 
purifying Hfq_bio (Kay et al., 2009). To ensure full biotinylation 
charged RNA. Hfq_bio was biotinylated by the biotin ligase 
reduce the risk that the tag might unspecifically bind negatively 
hfq 
allele was tagged with a biotinylation sequence 
strain, 
700 ng for total 

by washing several times with 98 
C, 15 mM NaCl, 1.5 mM 

and incubating at RT for 60 min. Samples were phenol/CHCl 
were used for analysis by northern blotting.

S1 Nuclease Mapping  
Briefly, a [\gamma-^32\text{P}]-ATP end-labeled DNA oligo antisense to the TLR area of interest was hybridized to total RNA and single-stranded overhangs were removed by addition of S1 nuclease. The resulting fragments were visualized on a denaturing polyacrylamide gel. If the synthetic oligo extends beyond the end of the target RNA, the position of the end can be determined by the number of nucleotides removed from the DNA oligo. S1 nuclease was used to map the RrB transcript of a strain harboring the plasmid pTSS1. RNA was harvested by hot phenol 1 h after IPTG induction. One pmol of 5'-end [\text{ ^^32P} ] -labeled probe was hybridized to 30 
µg of total RNA from the strain of interest. Hybridization was done in 50% formamide, 20 mM HEPES, 0.5 mM EDTA, 0.2 M NaCl, 0.05% (w/v) SDS and performed overnight in a thermocycler (68°C for 10 min, then the temperature was lowered to 54°C and decreased 1°C every 30 min until reaching 20°C). Digestion was performed by adding 300 
µl 0.28 M NaCl, 50 mM NaOAc pH 4.6, 4.5 mM ZnSO 
along with 300 
µl S1 nuclease (Thermo Fisher Scientific) and incubating at RT for 60 min. Samples were phenol/CHCl 
extracted, ethanol precipitated, size separated by electrophoresis on 7 M urea, 10% polyacrylamide sequencing gels and detected by autoradiography.

Circular RACE Mapping  
Circular RACE was used to map the isoform of RrB enriched on Hfq. Briefly, RNA purified from the Hfq_bio purification 

(Ørsen, 2001; Stenum et al., 2017) and in case of very low signals (e.g., Figure 2C) the signal found in the estimated position of a band was used. Before re-probing, membranes were stripped by washing several times with 98 
C, 15 mM NaCl, 1.5 mM 

for Hfq-bound RNA from the Hfq_Bio strain and 700 ng for total RNA samples), was used for analysis by northern blotting.

S1 Nuclease Mapping  
 Briefly, a [\gamma-^32\text{P}]-ATP end-labeled DNA oligo antisense to the TLR area of interest was hybridized to total RNA and single-stranded overhangs were removed by addition of S1 nuclease. The resulting fragments were visualized on a denaturing polyacrylamide gel. If the synthetic oligo extends beyond the end of the target RNA, the position of the end can be determined by the number of nucleotides removed from the DNA oligo. S1 nuclease was used to map the RrB transcript of a strain harboring the plasmid pTSS1. RNA was harvested by hot phenol 1 h after IPTG induction. One pmol of 5'-end [\text{ ^^32P} ] -labeled probe was hybridized to 30 
µg of total RNA from the strain of interest. Hybridization was done in 50% formamide, 20 mM HEPES, 0.5 mM EDTA, 0.2 M NaCl, 0.05% (w/v) SDS and performed overnight in a thermocycler (68°C for 10 min, then the temperature was lowered to 54°C and decreased 1°C every 30 min until reaching 20°C). Digestion was performed by adding 300 
µl 0.28 M NaCl, 50 mM NaOAc pH 4.6, 4.5 mM ZnSO 
along with 300 
µl S1 nuclease (Thermo Fisher Scientific) and incubating at RT for 60 min. Samples were phenol/CHCl 
extracted, ethanol precipitated, size separated by electrophoresis on 7 M urea, 10% polyacrylamide sequencing gels and detected by autoradiography.

Circular RACE Mapping  
Circular RACE was used to map the isoform of RrB enriched on Hfq. Briefly, RNA purified from the Hfq_bio purification.
experiment was circularized using RNA ligase, reverse transcribed using random hexamer primers, PCR amplified twice using nested sets of specific primers (cRACE rrB-2 1F + cRACE rrB-2 1R and cRACE rrB-2 2F + cRACE rrB-2 2R, see Supplementary Table 3). The PCR-library was sequenced on an Illumina Mi-seq by 300 bp paired-end sequencing. The resulting sequences were merged and subsequently listed by abundance. All reads that could not be merged were left out of the analysis.

RNA Stability During Rifampicin Treatment

Cultures were grown exponentially at 37°C, shaking at 160 rpm for at least 10 generations in MOPS medium supplemented with 0.2% glucose and 10 µg/ml uracil. At an OD600 of 0.7, 3 × 15 ml culture aliquots were collected (time 0 min) and rifampicin was added to the remaining culture to a final concentration of 100 µg/ml. Aliquots of 15 ml were collected at 2.5, 5, 10, and 20 min post rifampicin treatment. Five percent spike-in culture overexpressing tRNAselC was added to each sample aliquot as in Stenum et al. (2017). RNA was harvested using TRI-reagent (Sigma), as described by the manufacturer. The RNA was size-separated on polyacrylamide gels, northern blotted and probed as described above. RrBshort transcript levels were normalized to the tRNAselC level for each sample.

Structure Probing

The structure of RrB was investigated in vitro with and without Hfq.

Reactions (10 µl) containing 0.1 pmol of 5’-end [γ-32P] ATP-labeled transcript and 50 nM unlabeled E. coli tRNA were incubated with different concentrations of hexameric Hfq (Hfq6) (3, 1.5, and 0 µM) at 37°C for 100 min along with the relevant cleavage buffer. For Pb2+ cleavage: 1 × Structural Probing Buffer (Ambion AM2237), Pb2+ was added to a final concentration of 10 mM and samples incubated 1 min at 37°C. Control (C1) was without Hfq and Pb2+. RNaseIII cleavage: 1 × Short Cut MnCl2 buffer (New England Biolabs) and 0.002 units Short Cut RNaseIII (New England Biolabs), samples were incubated 20 min at 37°C. Control (C2) was without Hfq and RNaseIII. Control T1: 1 × Structural Probing Buffer (Ambion AM2237), sample was incubated at 95°C for 1 min, transferred to 37°C for 1 min. After addition of 0.05 U RNase T1 (Ambion AM2237), sample was incubated at 95°C for 5 min. OH ladder: 1 × Alkaline Hydrolysis Buffer (Ambion AM2237), sample was incubated at 95°C for 5 min. All samples were cooled by addition of 200 µl ice-cold H2O and transferred to ice. The digested RNA was phenol extracted, ethanol precipitated, resuspended in 1 × loading buffer II (Ambion AM2237), and separated on an 8% polyacrylamide/urea/TBE gel. Radioactive signal from the dried gel was visualized on a PhosphorImager (Typhoon -GE Healthcare). Hfq protein used for all in vitro experiments was a kind gift from Anders Boysen.
FIGURE 3 | Determination of 5’ and 3’ ends of RrB by S1 nuclease analysis and of RrB<sub>short</sub> by circular RACE. In the S1 nuclease mapping analysis (A–C) the 32<sup>P</sup>-labeled probe DNA was visualized by autoradiography of 10% poly-acrylamide sequencing gels. (A) mapping of the 5’-end using probe rrB68-upstrm [shown in panel (D)] and total RNA from an IPTG-induced culture of MG1655 + pRrB over-expressing truncated rrfB and rrB. The S1-protected fragment (lane 4) is 13 nt shorter than the untreated probe (lane 1). The exact number of nucleotides removed was determined by loading two labeled oligos as size markers [lane 2 (rrB54-upstrm) and 3 (rrB50-upstrm), probe rrB68-upstrm truncated by 14 and 18 nt respectively]. (B) probe rrB3′-map40-20 was hybridized to total RNA from an IPTG-induced culture of Δ22TLR + pRrfB-RrB over-expressing rrfB and rrB. The S1 protected band was shortened by 7 nt (lane 3) compared to the untreated probe (lane 1). Lane 2 shows a similar experiment using total RNA from the IPTG-induced strain Δ22TLR harboring empty vector (pJFR1). (C) Probe rrB3′-map30-35 was hybridized to total RNA from MG1655 + pRrfB-RrB over-expressing rrfB and rrB. The S1 protected band was shortened by 22 nt (lane 3) compared to the untreated probe (lane 1). Lane 2 shows a similar experiment using total RNA from the strain Δ22TLR harboring empty plasmid. (D) Map of the probes used for the experiments shown in panels (A–C). Top line represents the genomic map, rrB (5S) is shown in red, the intergenic sequences in gray and the annotated version of rrB in blue. Probe sequences are presented as green lines and probes shown in transparent colors were also used for mapping but detected no ends. The vertical broken lines denote the two ends detected in this study. (E) Circular RACE mapping of the RrB sequences that co-precipitated with Hfq. The ten most abundant RrB<sub>short</sub> sequences detected by circular RACE and deep sequencing are aligned to the genomic sequence. The size of the characters correlates with their relative abundance, which is also stated as a percentage of total merged reads (n = 4722). Bold characters highlight the two terminators T1 and T2 and the direct repeat, which is also found at the 3’-end of the mature 5S RNA is highlighted in red (see Figure 1).

MS2 Affinity Purification, RNA-Seq and MS-MS
Affinity purification of MS2-tagged RNAs was done either with in vivo expressed RNA or in vitro transcripts that were added to the cell lysate. The 115 bp RrB<sub>short</sub> sequence, mapped as the most abundant variant pulled down with Hfq (Figure 3) was cloned into the plasmid pH1S21 using PCR amplifying RrB<sub>short</sub> with either Nhel or XbaI restriction sites on the ends.
Purified PCR fragments were restriction digested with either NheI or XbaI and ligated into the pNS21 plasmid cut with the corresponding restriction enzyme, resulting in RRB\textsuperscript{short} 3’-fused to MS2 (NheI digestion) or RRB\textsuperscript{short} 5’-fused to MS2 (XbaI digestion), respectively. The transcripts from the resulting plasmids are terminated by the \texttt{rrfA} terminator originally found in \textit{Vibrio} species. Transcription from the plasmids is under control of the \texttt{P_{ElaC-O-1}} promoter. As the strains used for affinity purification harbor only the chromosomal copy of \textit{laci}, expression from these plasmids should be constitutive. This was verified by northern blotting (Supplementary Figure 4). Cell lysates of the strain \texttt{hfq-FLAG} (JVS814) harboring the different MS2-aptamer expressing plasmids, were prepared by growing cells in LB with 100 \(\mu\)g/ml ampicillin, to an OD\textsubscript{600} of 1.0. Cell pellets corresponding to 50 OD\textsubscript{600} units were resuspended in 800 \(\mu\)l of buffer A (20 mM Tris-HCl pH 8.0, 150 mM KCl, 1 mM MgCl\textsubscript{2}, 1 mM DTT) and lysed by addition of glass beads (0.1 mm), flash freezing in liquid N\textsubscript{2} and shaking at 30 Hz for 10 min. Lysates were cleared by centrifugation (30 min 16,000 \(g\)), flash freezing in liquid N\textsubscript{2} and thawing at 37 \(\circ\)C. Lysates corresponding to 2 and 0.5 OD\textsubscript{600} units of cell culture were used to prepare RNA and protein, respectively. Before addition of lystate, the affinity chromatography columns (Bio-Spin #732-6008, Bio-Rad) were prepared by adding 100 \(\mu\)l amylose resin (New England Biolabs, #E8021S), washing three times with 2 ml buffer A and subsequently adding 200 pmol \texttt{MS2-MBP} recombinant protein (a gift from the Jörg Vogel group). All steps of the affinity chromatography were done at 0.1 \(\mu\)l/\(\mu\)l, 30 min at 37 \(\circ\)C, gel-purified from denaturing polyacrylamide gels, dephosphorylated using calf intestinal alkaline phosphatase (30 min at 37 \(\circ\)C) and 5’ radio labeled using T4 polynucleotide kinase and \([\gamma^{32}\text{P}]\) ATP. After each step the RNA was phenol extracted once, chloroform extracted twice and ethanol precipitated. Binding reactions contained 100 mM KCl, 10 mM MgCl\textsubscript{2}, 2 mM DTT, 7.5% glycerol, 0.1 U SUPERase-IN RNase inhibitor (ThermoFisher), 2 ng total yeast RNA, 120 pM labeled RNA and 0–1,600 nM CsrA-3xFLAG in a 10 \(\mu\)l reaction. Samples were incubated 10 min at 37 \(\circ\)C and separated on 8% native polyacrylamide gels using 1x TBE as buffer. Radioactive signals were detected on a PhosphorImager (Typhoon -GE Healthcare). The CsrA-3xFLAG was purified from the strain CsrA-3xFLAG carrying the plasmid pBAD-RBS-csrA:3xFLAG as described by Jorgensen et al. (2013).

**Construction of Deletion Mutants**

For \(\lambda\) Red recombination, we used a MA51081-derivative strain (MAS1080) harboring the \(\lambda\) RED prophage \([\lambda\ cI\delta_{57} \Delta(cro-bioA)]\) imported from strain HME68 (Sawitzke et al., 2007) by P1 transduction. Deletion mutations were constructed in this strain as described (Sawitzke et al., 2007). For each deletion, a \texttt{cat-sacB} cassette was first inserted at the desired genomic region by selecting for chloramphenicol resistance and confirming by PCR. The cassette was then replaced by a DNA fragment designed to yield deletion of 107 bp, 103 bp and 103 bp for \texttt{rrnA}, \texttt{rrnB} and \texttt{rrnD}, respectively, by counter-selection of the cassette by sucrose tolerance and confirmation by PCR and DNA sequencing. The deletions roughly match the annotated TLR sequences, including all GGA motifs found in \texttt{RrA}, \texttt{B}, \texttt{F}. For reasons outside the scope of this study, the DNA fragments were constructed so that the TLR sequence was replaced with a tRNA gene in two out of the three deletion sites (see Supplementary Table 1).

To obtain an \textit{E. coli} mutant where the \(\lambda\) Red recombination enzymes had not been expressed, and thereby reduce the risk of undesired genome mutations, each mutated locus was then moved to an otherwise wildtype strain by P1 transduction first of each \texttt{cat-sacB} cassette and next of the locus carrying a deletion. P1 transduction was done as described by Miller (1972).

**Biofilm Measurements**

Biofilm was measured in microtiter plates using peg-lids (Nunc-TSP, cat. no. 445497) and crystal violet staining. Cultures were grown in 96-well flat-bottom microtiter plates. Ten microliter outgrown culture was used to inoculate each well containing 150 \(\mu\)l YT media (per liter: 8 g tryptone, 5 g yeast extract and 5 g NaCl) supplemented with 100 \(\mu\)l/ml ampicillin and 1 mM IPTG for strains carrying plasmids. The outer-most wells on the plates were not used, as the results from these were found to fluctuate more than average. Plates were incubated at 37 \(\circ\)C for 48 h (no
The normalized data was plotted against time and the area under glgC Fluorescence of GFP was measured from strains harboring the minimal media supplemented with 0.2% glucose, 15 µg/ml ampicillin and 1 mM IPTG for strains carrying plasmids. The plates were incubated approx. 16 h at 37°C before the spread of the bacteria was measured. Measurements were repeated several times with approximately 2 h in between to compensate for differences in the growth rates of different strains.

**Motility Measurements**

Mobility was measured on soft agar plates. One microliter outgrown culture was used to inoculate each plate by injection of 1 µl culture halfway into the agar in the center of the plate. Plates contained YT media (per liter: 8 g tryptone, 5 g yeast extract and 5 g NaCl) supplemented with 15 µg/ml ampicillin and 1 mM IPTG for strains carrying plasmids. The plates were incubated approx. 16 h at 37°C before the spread of the bacteria was measured. Measurements were repeated several times with approximately 2 h in between to compensate for differences in the growth rates of different strains.

**Translational Reporters**

The glgC-gfp in-frame fusion was constructed by replacing the NsiI-Nhel fragment of pXG10-SF (Supplementary Table 2) with the 5’UTR and first 30 nucleotides of the glgC coding sequence by restriction digestion and ligation. The pgaA-lacZ in-frame fusion was made by cloning the promoter of lacZ along with a Nhel restriction site into pGHB253-kan (Supplementary Table 2) by replacing the EcoRI-BamHI fragment. Subsequently, the 5’UTR and first 30 nucleotides of the pgaA coding sequence was inserted using BamHI-Nhel digestion. All primers used in the study are listed in Supplementary Table 3.

**Beta-Galactosidase Measurement**

Beta-Galactosidase (β-gal) activity was measured from overnight cultures harboring the translational fusion pgaA-lacZ, grown in test tubes in YT media (per liter: 8 g tryptone, 5 g yeast extract and 5 g NaCl) supplemented with 15 µg/ml ampicillin. The β-gal activity was measured using the fluorescent substrate 4-methylumbellifere b-D-galactopyranoside (MUG), as described (Li et al., 2018). OD600 of the cultures was measured and used to normalize the corresponding β-gal values.

**GFP Measurements**

Fluorescence of GFP was measured from strains harboring the glgC-gfp translational fusion. Cultures were grown in MOPS minimal media supplemented with 0.2% glucose, 15 µg/ml chloramphenicol and 100 µg/ml ampicillin when needed, at 37°C shaking at 300 rpm in 96-well black microtiter plates with clear, flat bottoms (Costar) in a plate reader (Synergy H1, Biotek). OD600 and GFP fluorescence (excitation 470 nm and emission 510 nm) was measured simultaneously every 10 min. The values for GFP were normalized to the corresponding values for OD600. The normalized data was plotted against time and the area under the curve was calculated as a measure of relative GFP expression. The same amount of data points were used for each strain and only data where cultures were growing exponentially was used in the analysis.

**RESULTS**

**Expression of RrA, B, F and Interaction With Hfq**

Three out of seven ribosomal RNA (rRNA) operons from *E. coli* contain a sequence from the TLR-family downstream from the SS gene (Supplementary Figure 1). These three sequences, named rrA, rrB and rrF respectively, are very similar in sequence (Figure 1). Furthermore, their location downstream of 5S, and their sequence, is highly conserved in many species of the Enterobacteriales, suggesting a specific role for these sequences and structures (Supplementary Figure 2).

We hypothesized that the sequences may function as sRNAs. To begin testing this hypothesis we first examined whether RNA molecules of defined sizes could be detected after processing of the primary *rrn* transcripts. Total RNA harvested from exponentially growing *E. coli* K-12 cultures was subjected to northern blot analysis. To detect RrA, RrB and RrF, here collectively referred to as RrA, B, F, we used a single probe, which is expected to detect all three sequences since they only differ at a few positions. As shown in Figure 2A (lane 1 and 2), a distinct band of 150–200 nt was observed, confirming that RrA, B, F can be detected after processing of the primary transcript. The size of the RrA, B, F RNA was estimated based on re-probing of the membrane for transcripts of known sizes (Figures 2A, B) and comparison to an RNA ladder (Figure 2C). Next, we asked whether the RrA, B, F RNAs interact with the RNA chaperone Hfq, like many well-characterized *E. coli* sRNAs (Bilusic et al., 2014). To enable pull-down of Hfq and analysis of co-purified RNAs, a biotinylation sequence (Beckett et al., 2008) was inserted in the chromosomal *hfq* gene. The C-terminally tagged Hfq protein (*Hfq_hbio*) is functional, as shown by intact repression of an mRNA target by an Hfq-dependent sRNA in the tagged strain (Supplementary Figure 3). After affinity purification of Hfq_hbio using a streptavidin resin, co-purificated RNA was analyzed by northern blot analysis. Detection of the known Hfq-binding sRNAs OxyS and ChiX only in the co-purificated RNA verified that Hfq-binding RNA had been enriched in the Hfq pull-down assay (Figure 2B). Interestingly, the pull-down revealed that a shorter form of RrA, B, F (RrA, B, F_short) was highly enriched on Hfq in addition to the band detected in total RNA (Figure 2A, lane 4). Further, to verify the identity of the bands on our northern blots found by the RrA, B, F probe, we made a blot containing RNA from a strain overexpressing RrB from pRrB, a strain deleted for *rrA, B, F*, and several samples from the wildtype strain harvested at different stages of growth (Figure 2C). The blot verifies the identity of the bands and shows that RrA, B, F are only detected during exponential growth, where the rRNA operons are actively transcribed. During stationary phase, the transcriptional activity of rRNA operons is absent or very low (Baracchini and Bremer, 1991) but the 5S, 16S and 23S RNAs...
Figure 3D

FIGURE 4 | RrA, B, F_short is a short-lived transcript. (A) Northern blot made from total RNA harvested from a culture of MG1655. The culture was grown in MOPS minimal media supplemented with 0.2% glucose, 10 mg/ml uracil. Rifampicin was added at 0 min to a concentration of 100 μg/ml. RNA was harvested at the indicated timepoints, right before and after rifampicin addition. Before RNA purification, spike-in cells overexpressing the 5S gene selC were added to each experimental sample (see section “Materials and Methods”). 8–10 μg RNA was loaded in each lane. The blot was probed for RrA, B, F and rRNA_selC as indicated. The arrow indicates the band used for quantification. The lane marked “C” only contain spike-in cells. (B) Quantification of northern blots. The half-life of RrA, B, F was estimated as the average half-life calculated from two independent experiments (Rep 1, Rep 2). Rep 2 is shown in panel (A). Independent biological replicates were obtained for all sample points except for the 10 min time point post rifampicin addition.

are stable molecules. Therefore the 5S rRNA detected during stationary phase in Figure 2C was transcribed during growth and serves as a qualitative loading control.

The ends of the longer RNA species were mapped using S1 nuclease protection assay (Berk and Sharp, 1977). As this assay was not sensitive enough to detect chromosomally expressed RrA, B, F, we used a strain overexpressing the 5S gene rrB along with rrB from a plasmid. The S1 protected fragments are displayed in Figures 3A–C. Several bands spaced with one nucleotide intervals are seen, which is a common observation when using S1 nuclease transcript mapping (Green and Roeder, 1980; Aiba et al., 1981; Brosius et al., 1982) and is probably due to “end-nibbling” (Shenk et al., 1975). Thus, we define the transcript ends based on the longest protected bands. For the 5′ end the probe was shortened by 13 nt (Figure 3A). For the 3′ end the two different probes were shortened by 7 nt (Figure 3B) and 22 nt (Figure 3C), respectively. They both predict the same end. The 5′ end of RrB was detected 3 nt downstream of the mature 3′ end of the 5S transcript (see Figures 1, 3A). This site has been described as the initial RNase E processing site of the pre-5S RNA (Roy et al., 1983; Li and Deutscher, 1995) suggesting that the 5′ end of RrB is generated by RNase E cleavage. The 3′ end of RrB was detected 13 nt downstream of its annotated 3′ end (Figure 1). This 3′ end was verified using two different probes (Figures 3B, C). These results predict that the most abundant RrA, B, F transcript in total RNA has a length of 173 nt, which is in good agreement with the northern blots shown in Figures 2A, C. We also carried out S1 analysis with probes antisense to the area between the two observed ends to detect any alternative transcript ends, but could not detect any (data not shown, the probes are shown in Figure 3D).

To determine the ends of the Hfq-enriched shorter version of RrA, B, F (RrA, B, F_short) we used the circular RACE method (McGrath, 2011), coupled with next generation sequencing. As this method is PCR-based it requires substantially less input RNA than the S1 nuclease assay. Circular RACE yielded a variety of sequences mapping to rrA, B, F, of which the ten most abundant are shown in Figure 3E. The most abundant reads predict the Hfq-associated RrA, B, F_short transcript to be 113–115 nt in length. This is in accordance with our observations from northern blots (Figure 2).

To detect the RrA, B, F_short fragment in the total RNA fraction, we carried out another northern blot analysis using substantially more total RNA per lane than the one presented in Figure 2A. This blot showed three different species of the RrA, B, F RNA (Figure 4A): the 115 nt RrA, B, F_short, the 173 nt fragment, and a longer version. We suspect that the longer version represents a processing intermediate that includes the T2 terminator sequence. Thus, the RrA, B, F_short is not uniquely seen in the RNA fraction co-precipitating with Hfq but can also be detected in the total RNA fraction.

A short half-life of the RrA, B, F RNA could explain its low abundance relative to the rRNA transcripts expressed from the same operons and the absence of signal from stationary phase cells (Figure 2C). We measured the half-life of RrA, B, F_short by monitoring the levels of RrA, B, F_short upon transcription initiation blocking by addition of rifampicin to a culture in balanced growth (Figure 4). Indeed, we found that the half-life of RrA, B, F_short was ~2.5 min, which is very short relative to the rRNA transcripts that are stable on the time scale of hours during exponential growth (Piir et al., 2011).

Taken together, we conclude that the RrA, B, F RNAs can be detected in exponentially growing cells as at least two transcripts of ~173 and ~115 nt that appear to interact with Hfq and that the shorter variant RrA, B, F_short has a half-life of about 2–3 min. The presence of RrA, B, F is therefore dependent on active rRNA transcription.

The Structure of RrA, B, F Is Conserved Within the Order of Enterobacteriales

In order to assess the conservation of RrA, B, F between species we did a BLAST search2 with rrB and the sequences

2https://blast.ncbi.nlm.nih.gov/Blast.cgi
surrounding the locus as input (see Supplementary Figure 2 and Supplementary Material). Both the sequence and the location of the rrb downstream of the gene encoding 5S was found to be conserved in at least 31 species from 12 genera, all belonging to the order of Enterobacteriales (Supplementary Figure 2). To evaluate the structural conservation of the RrA, B, F homologs we used the locARNA algorithm (Will et al., 2007, 2012; Raden et al., 2018) to predict a consensus structure (Figure 5). The consensus structure shows a high degree of structural conservation, particularly in the three stem loop structures named 2-4 in the figure, which could indicate a conserved function of the RNA.

We speculated that the enrichment of RrA, B, Fshort in RNA that had co-precipitated with Hfq could be due to a role for Hfq in chaperoning the correct folding of RrA, B, F, as shown for other RNAs (Geissmann and Touati, 2004; Soper and Woodson, 2008; Bordeau and Felden, 2014; Hoekzema et al., 2019). In this case, we would expect to see differences in the RrA, B, Fshort structure with and without Hfq. To experimentally investigate the structure of the 115 nt Rrbshort RNA, we performed RNA structure probing in the presence or absence of Hfq (Figure 6). As shown in Figure 1B, Rrshort and Rrbshort are identical in sequence, and differ by only two nucleotides from Rrfshort. We therefore expect RrA, B, Fshort to serve identical functions, and arbitrarily chose the 115 nt long Rrbshort from the rnrB operon for these experiments. An in vitro transcript of the RNA was incubated with increasing concentrations of Hfq, followed by exposure to the RNA cleavage agent Pb2+ that primarily hydrolyzes single-stranded RNA (Ciesiolka et al., 1998), or the endoribonuclease RNase III that predominantly hydrolyzes double-stranded RNA with a strong preference for helices of sufficient length (Robertson, 1990). Fragmented RNA was size-separated on a polyacrylamide gel and visualized by autoradiography (Figure 6A) In general, only subtle changes in the fragment pattern were observed upon addition of Hfq (Figure 6A). The most notable Hfq-induced change was found in the Pb2+-treated samples at the apparent single-stranded region at nt 63-64, 66-68. This sequence displayed reduced hydrolysis by Pb2+ in the presence of Hfq (Figure 6A). In contrast, nt 53-56 showed slightly enhanced Pb2+ cleavage, upon addition of Hfq. Finally, nt 48-50 displayed slightly enhanced RNase III cleavage after incubation with Hfq (Figure 6A). The change in cleavage pattern observed at nt 63-68 as a consequence of Hfq addition could be interpreted as direct binding of Hfq in the area, leading to protection from Pb2+-induced cleavage. Alternatively, the main outcome of interaction with Hfq may be a structural rearrangement that allows the region around nt 63-68 to engage in intramolecular base-pairing. The structure probing data do not allow us to distinguish between these two potential effects of Hfq. To pursue the hypothesis that Hfq may facilitate a structural rearrangement of the region, we used the algorithm Mfold (Zuker, 2003)\(^4\) to predict the most thermodynamically favorable Rrbshort structure with (Figure 6B) and without (Figure 6C) the constraint of a single-stranded region from nt 63-68 (open red triangles). The predicted structure

\[^4\]http://unafold.rna.albany.edu/?q=mfold

\(\text{FIGURE 5 | Conservation of sequence and structure of 30 Rb homologs found in Enterobacteria. The presented consensus structure was made using the algorithm LocARNA (Will et al., 2007, 2012; Raden et al., 2018) It is based on an alignment of 30 Rb homologs, from 30 different bacterial species, identified by BLAST search (see Supplementary Figure 2 and Supplementary Material to Figure 5). The base pairs of the locARNA structure are colored according to the level of structure and sequence conservation, as specified in the color legend. The hue indicates sequence conservation of the base pair (red: most conserved, purple: least conserved). The color saturation indicates the structural conservation of the base pair (High saturation: most conserved, low saturation: least conserved). Stem-loops are numbered as in Figure 6, Stem-loop 1 is not predicted to form in this consensus structure. Genera of the organisms included in the analysis are: Salmonella, Shigella, Klebsiella, Escherichia, Pectobacterium, Brenneria; the more distant: Citrobacter, Enterobacter, Providencia, Raoultella; and the most distant: Cedecea, Seratia and Hafnia.}\)
in Figure 6C, which shows the consensus stem-loop structure (stem loop #2) predicted by locARNA.

**RrB Binds to CsrA in vitro**

In order to identify potential interaction partners of RrA, B, F besides Hfq, we used MS2 affinity purification coupled with either mass spectrometry or RNA sequencing (Said et al., 2009; Corcoran et al., 2012; Lalouenna et al., 2015). Here, an RNA of interest is expressed as a fusion with an MS2-aptamer sequence, to allow purification of in vivo formed RNA-RNA or RNA-protein complexes using immobilized MS2 protein (Bardwell and Wickens, 1990; Said et al., 2009). We constructed plasmid-borne versions of RrB shorter, MS2-tagged at either the 5’-end or the 3’-end and affinity-purified the RNAs. Expression and purification of the tagged RNAs was verified by northern blotting (Supplementary Figure 4). RNAs co-purifying with RrB shorter were identified by deep RNA sequencing (RNA-seq) and co-purifying proteins were identified by mass spectrometry. The results were compared to results for affinity purification of the MS2 RNA alone. While no specific enrichment was reproducibly detected in the RNA-seq analysis (data not shown), several proteins were specifically enriched in pull downs with MS2-tagged RrB shorter, the most strongly enriched protein being the RNA-binding post-transcriptional regulator CsrA (Table 1). Somewhat surprising, Hfq was not found among the enriched proteins. We verified that RrB shorter binds Hfq in vitro by electrophoretic mobility shift assay (EMSA). The EMSA analysis showed that RrB shorter bound Hfq in vitro but with substantially lower affinity than the established Hfq-binding sRNA OxyS (Supplementary Figure 5). We will return to this point in section “Discussion.”

The canonical CsrA binding motif identified both in *S. typhimurium* (Holmqvist et al., 2016) and *E. coli* (Dubey et al., 2005; Potts et al., 2017) contains a GGA sequence located in the loop of a stem-loop structure. Interestingly, the RrA, B, F shorter sequences have five GGA sequences, two of which are predicted to be located in the loops of stem-loop structures by the mfold and LocaRNA structure prediction algorithms (Figures 5, 6). Intriguingly, the structure probing experiments shown in Figure 6A suggest that the potential CsrA binding motif in loop 2 is only formed after interaction with Hfq.

The machine-learning-based algorithm InvenireSRNA (Fakhry et al., 2017) is designed to predict sRNAs of the CsrB/C family. To gauge how RrA, B, F rank in this algorithm compared to other *E. coli* non-coding RNAs, we provided the algorithm with a total of 2902 sequences, including all annotated *E. coli* sRNA sequences, the 22 TLR sequences, all sequences identified as CsrA-binding peaks from CLIP-seq data (Potts et al., 2017),

| Protein | MS2-RrB | RrB-MS2 |
|---------|---------|---------|
| csrA    | 8.68    | 6.83    | 5.80    | 3.55    | 6.22    |
| gapA    | 1.04    | 1.88    | 6.03    | 3.49    | 3.11    |
| rluC    | 4.75    | 3.75    | 1.28    | 1.98    | 2.94    |
| rpmM    | 3.81    | 2.60    | 4.75    | 0.13    | 2.82    |
| rpsS    | 6.40    | 2.71    | 1.40    | 2.33    | 2.63    |
| sucA    | 2.79    | 1.85    | 2.62    | 2.33    | 2.40    |
| deaD    | -0.05   | 5.04    | 3.85    | 2.21    |
| rpsA    | 3.90    | 4.14    | 2.01    |
| gomA    | 7.72    | 1.93    |
| dnaK    | 1.01    | 2.84    | 3.56    | 1.85    |
| hq      | -1.81   | -2.29   | 4.91    | -4.06   |

RrB shorter was tagged with the MS2 sequence at either the 5’-end (MS2-RrB) or the 3’-end (RrB-MS2). The co-purifying proteins were identified and quantified by mass spec. The table presents the ten proteins most highly enriched on average, in these pull downs compared to pull downs using the MS2 transcript alone. The data for Hfq is also included. The numbers in the table represent the log2-transformed ratio of intensities between the sample and the MS2 transcript alone, after normalization to the total number of fragments analyzed in each sample. Color tones represent the degree of co-purification: red color (positive values) indicates relative more peptides co-purifying with RrB compared to the MS2 transcript alone, whereas blue color (negative values) indicates less co-purification.
and finally all intergenic regions of *E. coli* shorter than 1,000 bp. The results show that CsrB and CsrC both score high as expected (Table 2), while McaS and GadY both have low probability scores, suggesting that the features that make the latter acceptable binding partners for CsrA are not picked up by the algorithm. In accordance with the MS2-purification and structure prediction, RrA\textsubscript{short}, RrB\textsubscript{short} and RrF\textsubscript{short} are predicted to bind CsrA with high probability and, remarkably, rank among the 16 highest scoring sequences in the analysis (Table 2). Notably, the remaining 19 members of the TLR family located in tRNA operons all obtained a probability score of <0.01, suggesting RrA, B, F are unique among the TLR family in their affinity for CsrA.

To further validate the binding between RrB\textsubscript{short} and CsrA we conducted EMSAs. *In vitro* transcribed RrB\textsubscript{short} was incubated with increasing amounts of purified CsrA-3xFLAG and separated on non-denaturing gels. As seen in Figure 7A, an RrB\textsubscript{short}-CsrA complex is observed from a concentration of 50 nM CsrA, and a complete shift is seen at 100–200 nM CsrA. At higher concentrations of CsrA, several complexes with higher molecular weights are visible. The number of different band sizes suggests that RrB\textsubscript{short} has multiple binding sites for CsrA, which is expected based on our RNA structure predictions (Figures 5, 6). The binding relationship between CsrA and RrB\textsubscript{short} is therefore complex but a first approximation of an overall K\textsubscript{D} from our data in Figure 7A is around 100 nM. We repeated the binding experiment using a mutant version of RrB\textsubscript{short}, where three GGA motifs in the mutant RNA have been changed to “UUU” (number 1, 2 and 5 on Figure 6C), which includes the two motifs found in loops of stem structures. By only mutating the three GGA motifs found in single stranded regions, we do not expect this to have any consequences for the secondary structure of the RNA. The affinity for CsrA is clearly lower for this mutant compared to the wildtype RrB\textsubscript{short} as the mutant does not show any shift within the tested range of CsrA concentrations. Note that part of the RNA is degraded at the highest CsrA concentration. We attribute this to residual activity of the RNase that was added during CsrA purification, see materials and methods. To further probe the specificity of RrB\textsubscript{short} to CsrA we did competition binding experiments where a preformed CsrA-RrB\textsubscript{short} complex (CsrA at 200 nM, RrB\textsubscript{short} at 120 pM) was challenged with increasing concentrations of the unlabeled competitors RrB\textsubscript{short}, RrF\textsubscript{short}, 3GGA and CsrB (Figure 7B). CsrB is believed to have ~18 CsrA binding sites (Liu et al., 1997), a K\textsubscript{D} around 1 nM of binding to CsrA (Weibacher et al., 2003) and thus we expect it to be a more efficient competitor than RrB\textsubscript{short}. This is indeed what we observe, CsrB fully competes off the labeled RrB\textsubscript{short} at a concentration of 3.13 nM, whereas a concentration of 50 nM is needed in the self-competition with RrB\textsubscript{short}. RrF\textsubscript{short}, 3GGA show significantly less competition and is not able to fully compete the labeled RrB\textsubscript{short} off CsrA within the range of concentrations tested.

With these binding assays we clearly show that RrB\textsubscript{short} binds CsrA *in vitro* and that this interaction relies on the predicted binding sites located in single-stranded regions. The two remaining GGA motifs in the mutant RNA RrB\textsubscript{short}-3GGA do not efficiently bind CsrA on their own. We suspect that these low affinity GGA sites may require nearby high affinity sites in order to bind CsrA, this type of binding has previously been described for the CsrA homolog RsmE (Duss et al., 2014).

In combination, our results from MS2 affinity purification, structure prediction, the InvenireSRNA algorithm, and the EMSAs strongly suggest that RrB\textsubscript{short}, and likely RrA, B, F\textsubscript{short}, specifically interact with the post-transcriptional regulator CsrA.

### Phenotypic Effects of Altered RrA, B, F Levels

We next asked whether RrA, B, F might act by sequestering CsrA, similarly to the effects of CsrB, CsrC, GadY and McaS. To this end, we constructed an *E. coli* mutant deleted for the chromosomal copies of rra, B, F (see Materials and Methods). Additionally, to overexpress *rrb* in a fashion that would presumably allow normal processing of RrB\textsubscript{short}, we cloned the *rrb* sequence including the 3′-half of the adjacent 5S gene and both terminators (T1 and T2) into a modified pUC18 plasmid (pJFR1, see Supplementary Table 2) to make pRRB.

Putative effects of RrB expression on the activity of CsrA were investigated by monitoring several CsrA-regulated phenotypes: growth rate, biofilm formation, motility, and post-transcriptional repression of the CsrA-regulated genes *pgaA* (Wang et al., 2005) and *glgC* (Baker et al., 2002). As a positive control, we moved

---

**Table 2**: Transcripts predicted to regulate CsrA by the algorithm InvenireSRNA (Fakhry et al., 2017).

| RNA          | Probability score |
|--------------|-------------------|
| CsrB sRNA    | 0.9997            |
| CsrB chip-SEQ| 0.9977            |
| speA chip-SEQ| 0.9807            |
| glyQ_yaaB    | 0.9505            |
| setA_leuD    | 0.8794            |
| CmpX chip-SEQ| 0.872             |
| lhr_grxD     | 0.8634            |
| tatE_Ipa     | 0.8518            |
| CsrC sRNA    | 0.8435            |
| zapB chip-SEQ| 0.8308            |
| CsrC chip-SEQ| 0.7979            |
| ybhH_ybhl    | 0.7439            |
| glnD chip-SEQ| 0.7013            |
| yaeE_argP    | 0.6407            |
| RrA\textsubscript{short} | 0.6245 |
| RrF\textsubscript{short} | 0.5833 |
| McaS         | 0.005             |
| GadY         | 0.0003            |

The input into InvenireSRNA was all annotated sRNAs, the 22 TLR sequences, all sequences identified as peaks in a comprehensive CsrA CHIP-seq experiment (Potts et al., 2017) and all intergenic regions of *E. coli* shorter than 1,000 bp (2962 sequences in total). The table displays the 16 highest-scoring transcripts from the analysis, and the sRNAs McaS and GadY, along with the associated probability scores. Intergenic regions are named by the two genes flanking them, separated by “-”, and CHIP-seq peaks by addition of “chip-SEQ” to the gene name. Note that CsrB/C are present twice as they are included as full-length sequences from the list of sRNAs and as peak regions from the CHIP-seq data.
a previously characterized \textit{csrA::kan} allele (Romeo et al., 1993) into our wildtype strain (which is otherwise isogenic to the \textit{ΔrrA, B, F} strain). The \textit{csrA::kan} allele encodes a truncated version of CsrA which is known to have strongly reduced CsrA activity (Romeo et al., 1993). The results of these analyses are presented in \textbf{Figure 8}. In general, if RrA, B, F regulate CsrA activity in a manner similar to CsrB, we expect the phenotype of overexpression of \textit{rrB} to display a similar tendency to that of the \textit{csrA::kan} allele. Neither deletion nor overexpression of the \textit{rrA, B, F} significantly affected the growth rate in minimal glucose medium (\textbf{Figure 8}A). For comparison, \textit{E. coli} harboring a truncated CsrA have been reported to grow at approximately half the growth rate of the wildtype strain when glucose is the carbon source (Morin et al., 2016). The mutants lacking or overexpressing the \textit{rrA, B, F} were also tested for motility on soft agar plates. Neither overexpression nor deletion of the \textit{rrA, B, F} affected motility (\textbf{Figure 8}B). In contrast, the \textit{csrA::kan} strain was severely impaired for motility as previously described (Romeo et al., 1993), although we note that this phenotype in our hands was very variable between replicate experiments.

CsrA is a negative regulator of \textit{PGA} (poly-\textit{β}-1,6-\textit{N}-acetyl-d-glucosamine)-mediated biofilm formation (Wang et al., 2005). In agreement with this, we found that the \textit{csrA::kan} strain produces more biofilm than wildtype as measured by crystal violet staining (\textbf{Figure 8}C). The \textit{ΔrrA, B, F} strain also formed significantly more biofilm than the wildtype. However, overexpression of RrB also showed somewhat increased biofilm formation, although the difference from the strain carrying the empty vector control was not statistically significant.

In connection with biofilm, translation of the \textit{pgaABCD} mRNA, encoding the machinery for synthesis and transport of PGA, is known to be strongly repressed by CsrA (Wang et al., 2005). Indeed, we found > 20-fold higher activity of \textit{β}-galactosidase from a \textit{pgaA-lacZ} fusion in the presence of CsrB overexpression than in the wildtype strain (\textbf{Figure 8}D). A modest increase in \textit{β}-galactosidase activity was also observed in the \textit{ΔrrA, B, F} mutant. Again, overexpression of RrB resulted in a phenotype similar to that of the \textit{Δrr,AB,F} mutant, pointing to a surprisingly similar effect on the \textit{pgaA-lacZ} fusion of decreasing and increasing the \textit{RrA, B, F} levels. There is good correlation between our data on biofilm formation and our data on the \textit{pgaA-lacZ} fusion for all the strains (\textbf{Figures 8}C,D), as would be expected for CsrA-dependent regulation of biofilm (Wang et al., 2005).

Lastly, we tested the expression of GFP from a plasmid-borne translational \textit{glgC-gfp} fusion during exponential growth, the growth phase where RrA, B, F accumulated (\textbf{Figure 2}C). The \textit{glgCAP} operon encodes enzymes involved in glycogen metabolism, and is known to be repressed by CsrA (Baker et al., 2002). CsrB overexpression increases \textit{glgC-gfp} expression substantially, while we observed slightly more GFP signal both upon deletion and overexpression of \textit{rrA, B, F} (\textbf{Figure 8}E). In summary, we only observed minor phenotypic effects of changes in the expression of \textit{RrA, B, F}. The \textit{RrA, B, F} RNAs appear to modestly influence biofilm formation and expression from the CsrA-controlled \textit{pgaA} and \textit{glgC} mRNAs, but all three assays showed the same curious trend, namely that both deletion of \textit{rrA, B, F} and overexpression of \textit{rrB} resulted in a change consistent with modestly reduced CsrA activity. A plausible explanation for this unexpected similarity of phenotypes would be if RrB RNA expressed from the vector was processed into a form that exerted a different action on the tested properties than chromosomally expressed RrB. To address this hypothesis, we also tested expression of the \textit{glgC-gfp} fusion in the presence of the pKK3535...
plasmid that contains the entire rrnB operon, including rrB. Although the rRNA operon promoters are feedback-controlled and thus difficult to overexpress, this plasmid was expected to cause a relative increase in RrB levels, because the pKK3535 plasmid is responsible for approximately 50% of total rRNA synthesis even in the presence of the seven chromosomal rRNA operons (Steen et al., 1986). As shown in Supplementary Figure 5, overexpression of the rrnB operon resulted in increased glgC-gfp expression, and was thus consistent with reduced CsrA activity under this condition. In our hands, however, the strains containing pKK3535 had impaired growth rates (see Supplementary Figure 6A) and thus we did not pursue additional experiments with this plasmid.

### The Concentration of RrA, B, F Increases Upon Translational Halt

The very modest effects of RrA, B, F on CsrA-controlled phenotypes suggest that if there is a physiological consequence of RrA, B, F interaction with CsrA then we might not have examined it under the proper growth conditions. In *E. coli*, rRNA expression is controlled by the second messenger ppGpp.
A decrease in ppGpp levels leads to increased rRNA expression (as observed upon treatment with antibiotics specifically blocking translation (Kurland and Maaloe, 1962; Muto et al., 1975)). To test whether the ppGpp effect also applies to rRNA, we induced translational arrest, either by chloramphenicol addition, or by expression of the toxin MazF, and investigated the level of RrA, B, F, by northern blot analysis at several time points after induction. Chloramphenicol blocks translation by binding to ribosomes and inhibiting the peptidyl transferase activity (Das et al., 1966) while MazF leads to translational arrest by disrupting ribosome biogenesis and cleaving mRNAs (Culviner and Laub, 2018), thereby removing the template for translation. In both cases of translational arrest, we detect a strong increase in RrA, B, F transcript levels 20–80 min after the treatment (Figure 9). As an independent indicator of transcription from the rRNA operons, we also investigated the level of tRNA<sub>Glu</sub>. This tRNA is exclusively expressed from genes located in four rRNA operons (rrnB, rrnC, rrnE, rrnG), so we expected its expression pattern to resemble that of RrA, B, F. Indeed, the level of tRNA<sub>Glu</sub> also increases during translational arrest, albeit not nearly to the same extent as RrA, B, F (Figure 9).

**DISCUSSION**

We identify three small genes that are co-transcribed from rRNA operons and processed to transcripts of defined lengths that bind the post-transcriptional regulator CsrA both in vitro and in vivo. As the transcription rate of rRNA is strictly correlated with the growth rate of *E. coli* (Potrykus et al., 2011), it is tempting to speculate that such short-lived sRNAs transcribed along with rRNA could function to align growth-rate-regulated transcription rates with other growth-phase-dependent processes in the cell, such as those regulated by CsrA. The relatively short half-life of RrA, B, F compared to other sRNAs makes biological sense if their abundance should reflect the transcriptional activity of the rRNA operons in real time.

We find a modest effect of overexpression of <i>rrB</i> on three of the CsrA-regulated phenotypes tested (Figure 8). The phenotypes observed upon overexpression of <i>rrB</i> agree with those expected from a mutant with modestly reduced CsrA activity. Such an apparent inhibition of CsrA by <i>rrB</i> overexpression might be due to binding of RrB to CsrA, resulting in regulation by titration in the classical way first described for CsrB/C. Direct binding between CsrA and RrB<sup>short</sup> is experimentally supported by the results of our RNA affinity purification and EMSA analyses as well as theoretically supported by the InvenireSRNA prediction and the presence of consensus CsrA-binding motifs in the RrB<sup>short</sup> structure. The fact that deletion of the <i>rrr</i> genes showed similar phenotypes to RrB overexpression is then counterintuitive (see Figure 8). Potentially, the chromosomally encoded form of RrB could fulfill a function for which the version of RrB expressed from our multicopy plasmid has a dominant negative effect (overriding the phenotype of the wildtype allele), in which case similar phenotypes of the chromosomal <i>rrr</i> genes and the RrB overexpression could be expected.

The curious effect or RrB overexpression was investigated further by introducing a plasmid expressing the entire <i>rrnB</i> operon (Supplementary Figure 6), which confirmed the results obtained by specific overexpression of RrB (Figure 8E). Since overexpression of <i>rrnB</i> by two different cloning tactics affected the <i>gltC-gfp</i> fusion similarly, we conclude that deletion and overexpression of RrB both result in phenotypes consistent with reduced CsrA activity, but cannot currently provide a mechanistic explanation for this curious observation.

While we have demonstrated an interaction between the RrA, B, F RNAs and CsrA, deletion and overexpression of <i>rrr</i> show only modest changes of the CsrA-related phenotypes tested in this study. We want to note that CsrA is the major hub in a regulatory network with many inputs (Romeo and Babitzke, 2018). For that reason, one might not expect that the absence, or presence in excess, of RrA, B, F would lead to prominent phenotypes, but merely fine tuning of the activity of CsrA. Further, EMSA analyses (Figure 7) showed that CsrB totally competed RrB<sup>short</sup> off CsrA at a ten-fold lower concentration than RrB<sup>short</sup> itself (Figure 7B). This shows that CsrB is a higher-affinity CsrA binding partner than RrA, B, F and that we therefore might expect small effects, if any, of changes to the RrA, B, F levels in situations where CsrA-regulation by CsrB/C is at play.

Our northern analysis shows that RrA, B, F are most abundant when the rRNA operons are expressed, namely during exponential growth (Figure 2C), and especially upon treatments that lower ppGpp production by arresting translation (Figure 9). During translational halt, RrA, B, F levels increased more than the tRNA<sub>Glu</sub>, which is also expressed solely from rRNA operons. This difference in the extent of induction could either mean that there is differential expression of the genes in the rRNA operons upon translational arrest, or, more likely, translational arrest may result in increased stability of the RrA, B, F sRNAs relative to tRNA<sub>Glu</sub> by an unknown mechanism. Our data do not allow us to distinguish between these possibilities. Nevertheless, the experiments show that conditions exist in which the abundance of the RrA, B, F RNAs increase substantially. If they play a greater role in cell physiology than the fine-tuning of CsrA-regulated phenotypes shown in Figure 8, their impact should maybe be sought under such conditions, or at least under conditions where CsrB/C concentrations are low but the transcription rate of rRNA is high.

We find that <i>rrr</i> is well conserved in various different genera of bacteria including numerous pathogenic species. It is well known that growth-phase-dependent regulation is important for bacterial virulence (Dalebroux and Swanson, 2012; Kitamoto et al., 2016). Thus, if RrA, B, F is used as an indicator of growth rate in the cell, it could also potentially have a regulatory effect on the pathogenesis of these virulent stains.

A hallmark of all the members of the TLR family found in *E. coli* is the presence of a short (18–19 bp) repeated sequence identical to the 3′-end of the mature tRNA or rRNA from the same operon. For the tRNA or rRNA this 3′-end sequence is important for processing (reviewed in Mackie, 2013). The 3′-end of the 173-nt form of RrB was mapped 3 nt downstream of the repeated sequence that is identical to the final 18 bp of mature 5S RNA. As this repeated sequence and structure is known to be important for the RNase E cleavage immediately
downstream of 5S, we suggest that the repeat in the RrB 3'-end is also recognized by RNase E, and that this explains the function of the repeated sequence.

RrB<sub>short</sub> was enriched on Hfq in the co-purification experiment (Figure 2), and showed specific binding to Hfq by EMSA analysis (Supplementary Figure 5), which could suggest a role for Hfq in folding or processing of RrB. Hfq could affect folding and the kinetics of the processing but is not essential for processing since RrB<sub>short</sub> was also observed in a Δhfq mutant (Supplementary Figure 7). Alternatively, the enrichment could reflect that RrB<sub>short</sub> functions as a base-pairing sRNA to regulate one or more target RNAs, chaperoned by Hfq. GadY (Parker et al., 2017) and McaS (Jorgensen et al., 2013) are two recent examples of dual function sRNAs that can both be involved in endothelial inflammation and atherosclerosis (Son et al., 2017), and the murine-specific miR-712, transcribed as a part of the rrB gene, are among the most highly used transcription terminator sequences on plasmids for cloning and protein expression (Denèfle et al., 1987; Andrews et al., 1996; Rogers et al., 1997). Our demonstration that the Rra, B, F molecules can interact with at least two pleiotropic regulatory proteins in *E. coli* (Hfq and CsrA), warrants careful consideration of whether expression of such putative sRNAs from a high copy number plasmid is desirable for a given application.

Several studies on eukaryotic organisms have identified potential miRNAs transcribed as part of rRNA primary transcripts (Wei et al., 2013; Chak et al., 2015; Asha and Soniya, 2017), and the murine-specific miR-712, transcribed as a part of a spacer element in the pre-rRNA transcript, has been shown to be involved in endothelial inflammation and atherosclerosis (Son et al., 2013). To our knowledge, the Rra, B, F sRNAs described here represent the first report of sRNAs transcribed as part of rRNA primary transcripts in a prokaryote. In conclusion, three out of seven rRNA operons in *E. coli* contains an arrangement of dual terminators with small defined RNAs encoded in between, and these RNAs are conserved both in structure and sequence among species belonging to different genera of Enterobacteriales. We have found that these RNAs interact with both Hfq and CsrA but has been unable to demonstrate clear phenotypes of 

**FIGURE 9** Levels of Rra, B, F after addition of chloramphenicol or induction of *mazF* expression. The strains MAS1081 (A) and TSS261 (MAS1081 + pMazF) (B) were grown in MOPS minimal media supplemented with 0.2% glucose or 0.4% glycerol, respectively. Treatment was induced by addition of either chloramphenicol (20 µg/ml) (A) or arabinose (0.1%) (B). RNA was harvested at the indicated times after treatment. Equal volumes of RNA (6–15 µg/lane) were separated on 6% denaturing polyacrylamide gels and northern blotted. The blots were probed for tRNA<sub>Glu</sub> and Rra, B, F as indicated on the figure. Quantification of the radioactivity signal in each lane is given relative to the average value of the steady-state levels and is shown below each lane. nd: not determined due to high background.
both deletion and overexpression mutants. We presume that the conservation among species indicate a function for the RNAs but that this function is still not fully understood.

**DATA AVAILABILITY STATEMENT**

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: http://www.ebi.ac.uk/pride/archive/projects/PXD013749.

**AUTHOR CONTRIBUTIONS**

TS, MK, SS, and MS wrote the manuscript, conceived and designed the research. TS, MK, BK, EH, SS, and MS designed the experiments. MK (Hfq) and TS (CsrA) performed the experiments. All authors contributed to the article and approved the submitted version.

**REFERENCES**

Aiba, H., Adhya, S., and de Crombrugghe, B. (1981). Evidence for two functional gal promoters in intact *Escherichia coli* cells. *J. Biol. Chem.* 256, 11905–11910.

Altier, C., Suyemoto, M., Ruiz, A. I., Burnham, K. D., and Maurer, R. (2002). A deeply conserved, noncanonical miRNA hosted by ribosomal DNA. *EMBO J.* 21, 921–929.

Andrews, B., Adari, H., Hannig, G., Lahue, E., Gosselin, M., Martin, S., et al. (1996). A tightly regulated high level expression vector that utilizes a thermosensitive lac repressor: production of the human T cell receptor Vβ5.3 in *Escherichia coli*. *Gene* 182, 101–109. doi: 10.1016/0378-1119(96)00523-9

Asha, S., and Soniya, E. V. (2017). The sRNAome mining revealed existence of Baglio, S. R., Rooijers, K., Koppers-Lalic, D., Verweij, F. J., Perez Lanzón, M., Zini, Berk, A. J., and Sharp, P. A. (1977). Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. *Cell* 12, 721–732. doi: 10.1016/0092-8674(77)90027-0

Bardwell, V. J., and Wickens, M. (1990). Purification of RNA and RNA-protein complexes by an R17 coat protein affinity method. *Nucleic Acids Res.* 18, 6587–6594.

Becquet, D., Kovaleva, E., and Schatz, P. J. (2008). A minimal peptide substrate in a pseudoknot assisted by an RNA chaperone. *J. Mol. Biol.* 266, 11753–11760. doi: 10.1016/S0022-2836(08)01393-8

Bordew, V., and Felden, B. (2014). Curli synthesis and biofilm formation in enteric bacteria are controlled by a dynamic small RNA module made up of a pseudoknot assisted by an RNA chaperone. *Nucleic Acids Res.* 42, 4682–4696. doi: 10.1093/nar/gku908

Beisel, C. L., and Storz, G. (2011). The base-pairing RNA Spot 42 participates in a multioutput feedforward loop to help enact catabolite repression in *Escherichia coli*. *J. Mol. Biol.* 414, 1599–1610. doi: 10.1016/j.jmb.2011.02.0982.x

Beisel, C. L., and Storz, G. (2011). Control of RNA synthesis in *Escherichia coli* at increased RRN gene dosage. Role of guanosine tetraphosphate and ribosome feedback. *J. Biol. Chem.* 286, 11753–11760. doi: 10.1074/jbc.M111.2925818

Bilusic, I., Popitsch, N., Rescheneder, P., Schroeder, R., and Lybecker, M. (2014). Revisiting the coding potential of the *E. coli* genome through Hfq co-immunoprecipitation. *RNA Biol.* 11, 641–654. doi: 10.4161/rna.29299

Borode, V., and Felden, B. (2014). Curli synthesis and biofilm formation in enteric bacteria are controlled by a dynamic small RNA module made up of a pseudoknot assisted by an RNA chaperone. *Nucleic Acids Res.* 42, 4682–4696. doi: 10.1093/nar/gku908

Bos, M., and Kersken, H. (1991). A novel RNA product of the tyrT operon of *Escherichia coli*. *Nucleic Acids Res.* 19, 5863–5870. doi: 10.1093/nar/19.21.5863

Bos, M., and Kersken, H. (1994). Organization and functions of genes in the upstream region of tyrT of *Escherichia coli*: phenotypes of mutants with partial deletion of a new gene (tgs). *J. Bacteriol.* 176, 221–231. doi: 10.1128/jb.176.1.221-231.1994

Brantl, S., and Jahn, N. (2015). sRNAs in bacterial type I and type III toxin-antitoxin systems. *FEMS Microbiol. Rev.* 39, 413–427.

Brosius, J. (1984). Toxicity of an overproduced foreign gene product in *Escherichia coli* and its use in plasmid vectors for the selection of transcription terminators. *Gene* 27, 161–172. doi: 10.1016/0378-1119(84)90137-9

Brosius, J., Cate, R. L., and Perlmutter, A. P. (1982). Precise location of two promoters for the beta-lactamase gene of pBR322. *S1 mapping of ribonuclease acid isolated from Escherichia coli* synthesized in vitro. *J. Biol. Chem.* 257, 9205–9210.

Brosius, J., Null, T. J., Sleet, D. D., and Noller, H. F. (1981). Gene organization and primary structure of a ribosomal RNA operon from *Escherichia coli*. *J. Mol. Biol.* 148, 107–127. doi: 10.1016/0022-2836(81)90058-8

Chak, L. L., Mohammad, J., Lai, E. C., Tucker-Kellogg, G., and Okamura, K. (2015). A deeply conserved, noncanonical miRNA hosted by ribosomal DNA. *RNA* 21, 375–384. doi: 10.1261/rna.049098.114

Chao, Y., Papenfort, K., Reinhardt, R., Sharma, C. M., and Vogel, J. (2012). An atlas of Hfq-bound transcripts reveals 3’ UTRs as a genomic reservoir of regulatory small RNAs. *EMBO J.* 31, 4005–4019. doi: 10.1038/emboj.2012.229

Ciesiolka, J., Michalowski, D., Wrzesinski, J., Krajewski, J., and Krzyzosiak, W. J. (1998). Patterns of cleavages induced by lead ions in defined RNA secondary structure motifs 1 Edited by L. Tinoco. *J. Mol. Biol.* 275, 211–220. doi: 10.1006/jmbi.1997.1462

Corcoran, C. P., Rieder, R., Podkaminski, D., Hofmann, B., and Vogel, J. (2012). “Use of aptamer tagging to identify in vivo protein binding partners of small resolution using two-color fluorescent DNA microarrays. *Proc. Natl. Acad. Sci.* U.S.A. 99, 9607–9702. doi: 10.1073/pnas.1123181199

FUNDING

This work was supported by the Lundbeck Foundation [R108-A10583 to MS], the Danish council for Independent Research | Natural Sciences [1323-00343B to SS] and the Danish National Research Foundation [DNRF120 to SS; MS].

**ACKNOWLEDGMENTS**

The authors thank Marit Warrer, Bertil Gummesson, and Pilar Menéndez Gil for excellent technical assistance. The authors thank Jörg Vogel for help and support to TS during his visit in Würzburg. The authors also thank Anders Boysem for the kind supply of purified Hfq.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.625585/full#supplementary-material
Sawitzke, J. A., Thomason, L. C., Costantino, N., Bubunenko, M., Datta, S., and Rudd, K. E. (1999). Novel intergenic repeats of
Said, N., Rieder, R., Hurwitz, R., Deckert, J., Urlaub, H., and Vogel, J. (2009). Romeo, T., Gong, M., Liu, M. Y., and Brun-Zinkernagel, A.-M. (1993).
Romeo, T., and Babitzke, P. (2018). “Global regulation by CsrA and its RNA
Robertson, H. D. (1990).
Raden, M., Ali, S. M., Alkhnbashi, O. S., Busch, A., Costa, F., Davis, J. A., et al.
Potts, A. H., Vakulskas, C. A., Pannuri, A., Yakhnin, H., Babitzke, P., and Romeo, T.
Rasmussen, A. A., Johansen, J., Nielsen, J. S., Overgaard, M., Kallipolitis, B., and
Stenum et al. sRNA Encoded in rRNA Operons
Shenk, T. E., Rhodes, C., Rigby, P. W., and Berg, P. (1975). Biochemical method
der 1996. Frontiers in Microbiology | www.frontiersin.org

May 2021 | Volume 12 | Article 625585

Stenum et al.

sRNA Encoded in rRNA Operons