

Ribosome heterogeneity results in leader sequence-mediated regulation of protein synthesis in *Francisella tularensis*

Hannah S. Trautmann,¹ Sierra S. Schmidt,¹ Steven T. Gregory,¹ Kathryn M. Ramsey^{1,2}

AUTHOR AFFILIATIONS See affiliation list on p. 18.

ABSTRACT Although ribosomes are generally examined in aggregate, ribosomes can be heterogenous in composition. Evidence is accumulating that changes in ribosome composition may result in altered function, such that ribosome heterogeneity may provide a mechanism to regulate protein synthesis. Ribosome heterogeneity in the human pathogen *Francisella tularensis* results from incorporation of one of three homologs of bS21, a small ribosomal subunit protein demonstrated to regulate protein synthesis in other bacteria. Loss of one homolog, bS21-2, results in genome-wide post-transcriptional changes in protein abundance. This suggests that bS21-2 can, either directly or indirectly, lead to preferential translation of particular mRNAs. Here, we examine the potential of bS21-2 to function in a leader sequence-dependent manner and to function indirectly, via Hfq. We found that the 5' untranslated region (UTR) of some bS21-2-responsive genes, including key virulence genes, is sufficient to alter translation in cells lacking bS21-2. We further identify features of a 5' UTR that allow responsiveness to bS21-2. These include an imperfect Shine-Dalgarno sequence and a particular six nucleotide sequence. Our results are consistent with a model in which a bS21 homolog increases the efficiency of translation initiation through interactions with specific leader sequences. With respect to bS21-2 indirectly regulating translation via the RNA-binding protein Hfq, we found that Hfq controls transcript abundance rather than protein synthesis, impacting virulence gene expression via a distinct mechanism. Together, we determined that ribosome composition in *F. tularensis* regulates translation in a leader sequence-dependent manner, a regulatory mechanism which may be used in other bacteria.

IMPORTANCE Ribosome heterogeneity is common in bacteria, and there is mounting evidence that ribosome composition plays a regulatory role in protein synthesis. However, mechanisms of ribosome-driven gene regulation are not well understood. In the human pathogen *Francisella tularensis*, which encodes multiple homologs for the ribosomal protein bS21, loss of one homolog impacts protein synthesis and virulence. Here, we explore the mechanism behind bS21-mediated changes in protein synthesis, finding that they can be linked to altered translation initiation and are dependent on specific sequences in the leaders of transcripts. Our data support a model in which ribosome composition regulates gene expression through translation, a strategy that may be conserved in diverse organisms with various sources of ribosome heterogeneity.

KEYWORDS *Francisella tularensis*, gene regulation, ribosomes, translation, virulence regulation, bS21, ribosomal proteins, T6SS

Ribosomes, the molecular machines that synthesize proteins, can be heterogeneous in composition (1). As bacterial ribosomes are composed of three ribosomal RNA molecules (rRNAs) and ~50 ribosomal proteins (r-proteins), heterogeneity can arise from differences in rRNA sequence among *rnm* operons, rRNA or r-protein modification, or

Editor Laurie E. Comstock, University of Chicago, Chicago, Illinois, USA

Address correspondence to Kathryn M. Ramsey, kramsey@uri.edu.

The authors declare no conflict of interest.

See the funding table on p. 18.

Received 23 April 2023

Accepted 7 July 2023

Copyright © 2023 American Society for Microbiology. All Rights Reserved.

r-protein content (2). The consequences of ribosome heterogeneity are incompletely understood, and much debate surrounds the hypothesis that distinct classes of ribosomes can have specialized functions by preferentially translating subsets of mRNA (3).

In *Escherichia coli*, the 30S subunit r-protein bS21 is not required for translation *in vitro* but is essential for viability (4, 5). It is one of the last proteins to be incorporated during 30S subunit assembly (6). While the precise function of bS21 in translation is not clear, early studies demonstrated its involvement in translation initiation (7–9). Association of bS21 with the 30S subunit is reversible, as it is easily exchanged among ribosomes (10) suggesting that the presence or absence of bS21 can act as a source of ribosomal heterogeneity.

More recently, multiple studies have suggested that bS21 might play a regulatory role in gene expression (11–14). High-resolution ribosome structures have shown bS21 to be located in the 30S subunit platform, near the mRNA exit channel, and in position to interact with the anti-Shine-Dalgarno (ASD) sequence of 16S rRNA in the 30S initiation complex (15) and with the Shine-Dalgarno/anti-Shine-Dalgarno (SD-ASD) helix in the 70S ribosome (16), suggesting possible mechanisms for regulation.

Recent work in the Bacteroidia species *Flavobacterium johnsoniae* clearly demonstrates that bS21 controls gene expression. In this organism, incorporation of the single bS21 homolog into the ribosome contributes to sequestration of the ASD (12). The mRNA encoding bS21, *rpsU*, is essentially the only *F. johnsoniae* mRNA with a strong SD sequence. Depletion of bS21 or removal of the region of bS21 necessary for ASD sequestration results in increased translation from the *rpsU* mRNA and mRNAs engineered to have a strong SD (17). These studies unambiguously demonstrate that ribosomes lacking bS21 have altered specificity for particular mRNAs in translation initiation, providing evidence that bS21 functions as a bona fide regulator of gene expression (17).

Francisella tularensis, a human pathogen that requires a type VI secretion system (T6SS) to cause disease, encodes three distinct homologs of bS21. We have shown that all three homologs can be incorporated into ribosomes (14). Accordingly, in *F. tularensis*, ribosome heterogeneity may be due to the identity of the bS21 homolog incorporated or, as in *F. johnsoniae*, the presence or absence of bS21 in the ribosome. Importantly, loss of one of the homologs, bS21-2, leads to reductions in most T6SS proteins that cannot be explained by changes in transcript abundance or protein stability. Loss of bS21-2 also results in defective intramacrophage replication in cells that can be complemented by restoration of bS21-2 but not by either of the other two homologs. These results indicate that bS21-2 specifically governs translation of virulence genes, including those that encode the T6SS (14).

While our observations support a model in which bS21 proteins in *F. tularensis* regulate gene expression at the level of translation, they are not likely to exert their effects in the same manner as the bS21 homologs in Bacteroidia. In *F. johnsoniae*, the C-terminal region of the single bS21 homolog is required to interact with and sequester the ASD; while this region is conserved among the Bacteroidia, it is not conserved among other bacterial lineages, including in Gammaproteobacteria like *F. tularensis* (12). Thus, in the current study, we aim to understand the mechanisms by which bS21-2 affects translation in *F. tularensis*.

Through reporter assays using translational fusions, we found that the 5′ untranslated region (5′ UTR) of some transcripts is sufficient to cause differences in protein production if bS21-2 is lost, indicating that these 5′ UTRs are responsive to bS21-2. By mutagenizing the bS21-2-responsive 5′ UTRs, we also found that transcripts with ideal SD sequences do not require bS21-2 for efficient translation. In an attempt to identify which component of the 5′ UTR is driving the responsiveness to bS21-2, we determined that two motifs enriched in bS21-2-controlled genes are not needed to cause changes in protein abundance and that the secondary structure of leader sequences does not play a clear role in responsiveness to bS21-2. Yet we have identified a short nucleotide

sequence in the 5' UTR of *mraY* that is critical for control by bS21-2. Finally, using qPCR and immunoblot analyses, we show that bS21-2 and the RNA-binding protein Hfq both govern T6SS protein abundance but do not act in a coordinated manner and function via distinct pathways. Our findings suggest that the r-protein bS21-2 governs protein abundance in *F. tularensis* by influencing translation from mRNA species with specific leader sequences.

RESULTS

bS21-2 promotes translation of specific genes in a 5' UTR-dependent manner

In *E. coli*, bS21 has been implicated in sequence-dependent translation initiation and, in the ribosome, is located adjacent to the 5' UTR of mRNAs during translation initiation (9, 15). Loss of bS21-2 in *F. tularensis* leads to changes in abundance for a subset of the proteome (14). This led us to hypothesize that bS21-2 may directly impact protein abundance by modulating translation initiation in a 5' UTR-dependent manner. In order to test this hypothesis, our goal was to assess the role of 5' UTR sequences in bS21-2-mediated translation of particular genes. To accomplish this, we developed a series of reporter constructs consisting of the experimentally determined or predicted 5' UTR with the first six codons of the gene of interest fused to a reporter gene (*lacZ* or *gfp*) (Fig. 1A). Reporter constructs were expressed from the *tul4* promoter, which is unaffected by the presence or absence of bS21-2 (14). This design allows for comparable transcription of reporter genes in both genotypes to compare relative translation initiation. The reporter constructs were introduced into either wild-type (WT) *F. tularensis* or *F. tularensis* lacking bS21-2 ($\Delta rpsU2$). Some experiments were completed using β -galactosidase reporters incorporated into the chromosome at the Tn7 site. Toxicity of plasmids that produce high levels of β -galactosidase in *E. coli* during plasmid production led us to use a green fluorescent protein (GFP)-based reporter system for some constructs. Reporter constructs using *gfp* were cloned into a multi-copy plasmid that is retained at essentially the same copy number in *F. tularensis* cells with and without bS21-2 (Fig. S1).

These reporter assays evaluated the relative efficiency of translation initiation of specific 5' UTRs in cells with or without bS21-2. We chose to assess the 5' UTRs corresponding to genes with significant changes in protein abundance in cells lacking bS21-2 (14). Consistent with the observed changes in protein abundance being due to changes in translation initiation, we found that the 5' UTRs of *pdpA*, *iglA*, *mraY*, FTL_0222, or FTL_1093 genes fused to *gfp* led to significantly less fluorescence in cells lacking bS21-2 compared to wild-type (Fig. 1B). In contrast, the 5' UTR of *tul4*, a gene not differentially expressed in cells lacking bS21-2, did not lead to a significant decrease in fluorescence in cells lacking bS21-2 (Fig. 1B). These data reveal that the 5' UTR of a gene is sufficient for bS21-2 to affect translation and is consistent with the idea that bS21 may be regulating translation initiation. We will refer to 5' UTRs that result in altered protein abundance in the presence of bS21-2 as "bS21-2-responsive." We also found that the 5' UTRs of some genes governed by bS21-2 in our proteomics analysis did not lead to reporter activity differences in cells lacking bS21-2, including FTL_0881 and FTL_0215 (Fig. S2). We do not have experimentally determined transcription start sites for these genes, so it is possible that the lack of regulation is due to inaccurate 5' UTR predictions. Other possibilities, including indirect regulation of these particular genes, are described in the Discussion. We confirmed that alterations in protein production in cells lacking bS21-2 could be complemented by ectopic expression of bS21-2 from a plasmid, indicating that the changes in translation are not due to polar effects of the *rpsU2* deletion (Fig. 1C).

An ideal Shine-Dalgarno sequence masks the positive effects of bS21-2 on translation

Based on structures of the *E. coli* ribosome during translation initiation, the bS21 residue R17 is close enough to directly contact the 16S rRNA nucleotide C1539, which is part of the anti-Shine-Dalgarno sequence (15) (Fig. 2A). R17 is conserved in all three *F. tularensis*

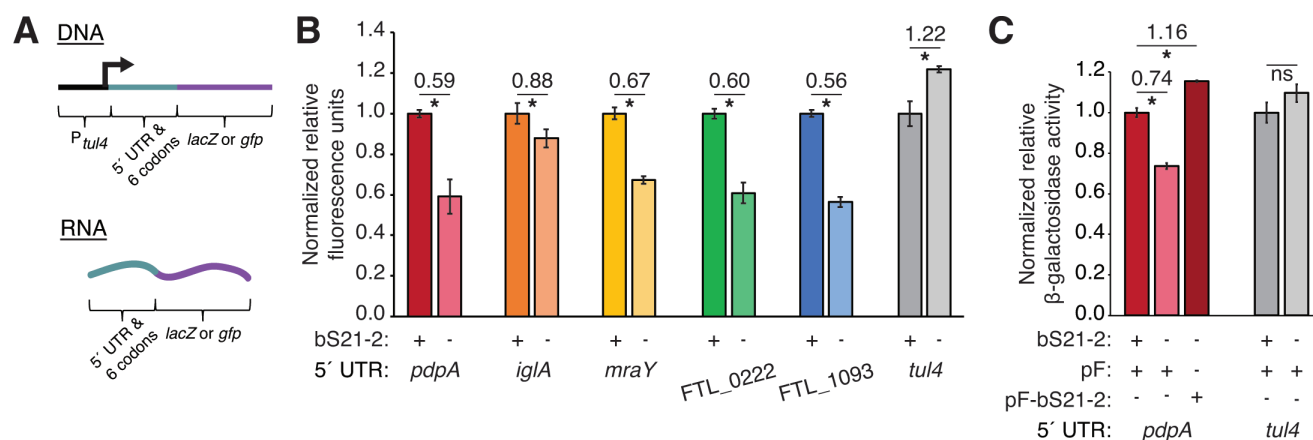


FIG 1 5' UTRs are sufficient to lead to bS21-2-mediated changes in translation. (A) Diagram of the translational reporter fusions used. Reporters used the *tul4* promoter to drive production of mRNA including the tested 5' UTR with the first six codons of the gene. Leader sequences are in-frame with either *lacZ* (integrated at the Tn7 site of the genome) or *gfp* (on a plasmid). (B) Relative fluorescence for indicated *gfp* translational fusion reporters in cells with (+; WT) or without (-; $\Delta rpsU2$) bS21-2. *tul4* serves as a control. 5' UTR sequences can be found in Table S1. (C) Relative β -galactosidase activity for indicated *lacZ* translational fusions in cells with (+; WT) or without (-; $\Delta rpsU2$) native bS21-2, or with ectopically expressed bS21-2 from a plasmid (pF-bS21-2). Strains without ectopically expressed bS21-2 contained an empty vector (pF). (B and C) Error bars represent 1 SD. Lines above bars indicate comparisons, values above line indicate ratio of reporter activity in cells lacking bS21-2 to wild-type cells. * $P < 0.05$. ns = not significant. Experiments were repeated at least twice in biological triplicate, and data from a representative experiment are shown.

bS21 homologs, and the rRNA-encoded ASD is identical in *F. tularensis* and *E. coli*. Thus, we hypothesized that bS21 homologs in *F. tularensis* may also contact the ASD and influence SD binding during translation initiation.

First, we assessed how SD strength correlates with responsiveness to bS21-2. We compared predicted SDs for genes whose proteins are positively affected ($n = 74$), negatively affected ($n = 84$), or unaffected ($n = 82$) by bS21-2 (14). We found that the genes positively affected by bS21-2 generally have weaker SD sequences, with only 39% having strong SD sequences (defined by four or more nucleotides [nt] complementary to the ASD), compared to 54% or 69% strong SDs in negatively affected or unaffected genes, respectively (Fig. 2B). These data are consistent with a model in which bS21-2 increases translation initiation predominately in the absence of strong SD-ASD interactions.

We specifically examined the influence of the SD sequences on the bS21-2-responsive 5' UTR of *pdpA*, a T6SS gene. In particular, we developed β -galactosidase translational reporters with altered SD sequences in the *pdpA* 5' UTR (Fig. 2C). The 5' UTRs with mutations that retained imperfect base-pairing between the ASD and SD (badSD, *tul4*SD) were still bS21-2-responsive. However, introducing an ideal SD, in two different positions (idealSD, ideal_movedSD), led to similar reporter gene expression in cells with and without bS21-2, indicating that these 5' UTRs are no longer bS21-2-responsive (Fig. 2C).

We replicated the impact of a perfect SD on the bS21-2-responsive 5' UTR of another gene, *mraY*. Modification of the imperfect *mraY* SD to an ideal SD resulted in no significant difference in GFP production in cells with or without bS21-2 (Fig. 3A). It is worth noting that in each of these cases, the addition of a perfect SD in the correct location (separated from the start codon by 4–9 nt) leads to increased total reporter production (Fig. S3 and S4). These data suggest that genes with perfect SD sequences do not require bS21-2 for efficient translation; in other words, an ideal SD may lead to such efficient translation that any response to bS21-2 becomes negligible. However, given that many *F. tularensis* genes have weak or non-perfect SDs and are not affected by bS21-2, there is an unidentified component of the 5' UTR that results in responsiveness to bS21-2.

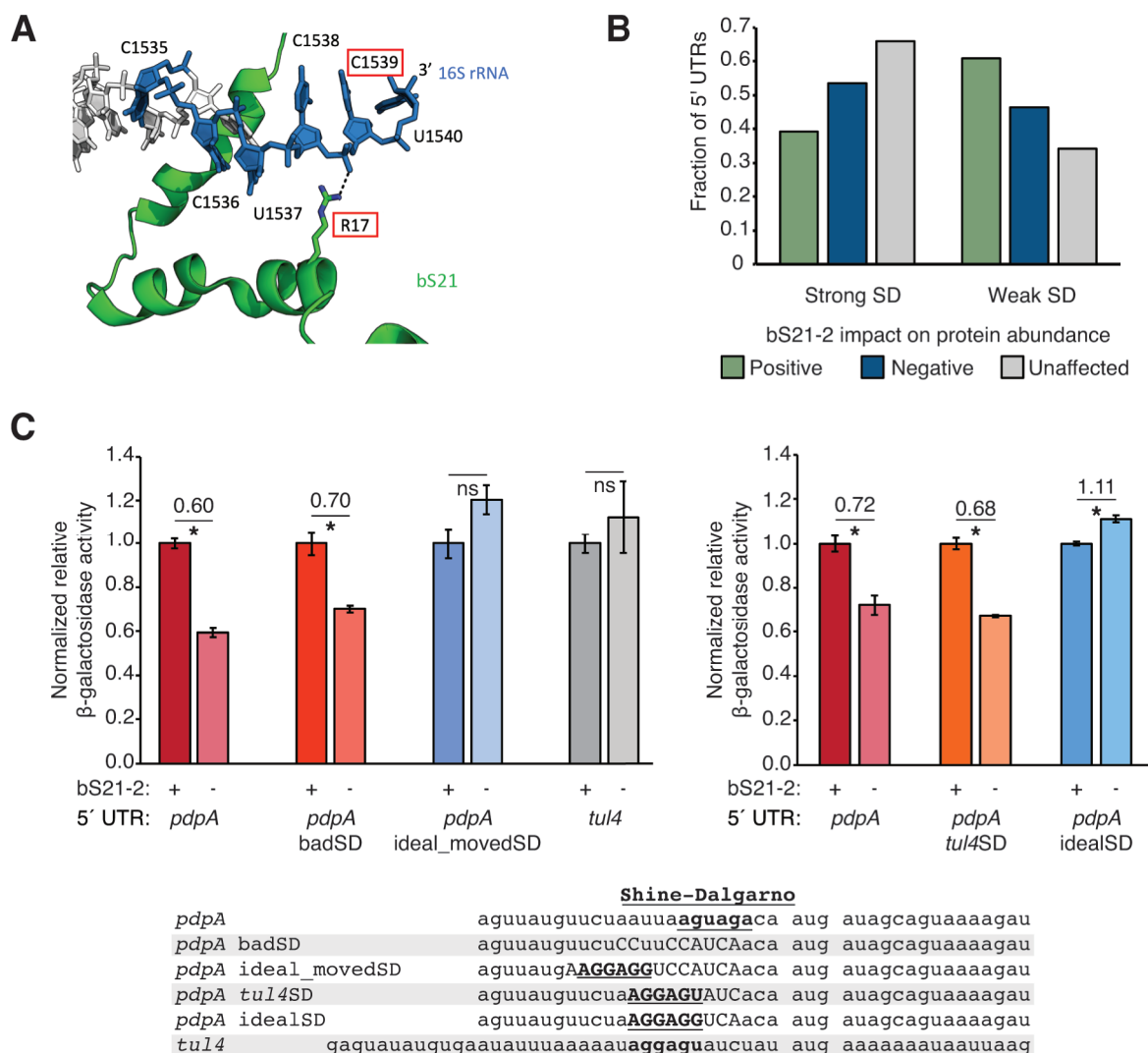


FIG 2 Genes with ideal Shine-Dalgarno sequences are not responsive to bS21-2. (A) bS21 interacts with the ASD sequence. In *E. coli*, amino acid R17 of the sole bS21 protein (green) directly interacts with C1539 of 16S rRNA, which is part of the ASD (blue). Measured distance is 2.7 Å (PDB 6o7k; 15). (B) The absence of strong SD-ASD interactions is correlated with bS21-2 influencing translation. Fraction of genes that are positively impacted ($n = 74$), negatively impacted ($n = 84$), or unaffected ($n = 82$) by bS21-2, categorized by strength of SD. “Strong” SD: four or more nucleotides complementary to ASD; “weak” SD: three or fewer complementary nucleotides. (C) Introduction of an ideal SD in the *pdpA* leader leads to loss of bS21-2 responsiveness. Top: relative β-galactosidase activity for indicated *lacZ* translational fusions in cells with (+; WT) or without (–; *ΔrpsU2*) bS21-2. Lines above bars indicate comparisons, values above line indicate ratio of reporter activity in cells lacking bS21-2 to wild-type cells. Error bars represent 1 SD. * $P < 0.05$ by t-test. ns = not significant. Experiments were repeated at least twice in biological triplicate, and data from a representative experiment are shown. Bottom: alignment of modifications to *pdpA* 5′ UTR. Capital letters: altered from WT; bold and underlined: predicted SD sequences; unnormalized β-galactosidase activity can be found in Fig. S3.

Sequence-specific motifs found in the 5′ UTR of genes governed by bS21-2 do not alter bS21-2 responsiveness

We reasoned that 5′ UTRs may be responsive to bS21-2 because they harbor a common sequence-specific element that mediates an altered interaction with bS21-2 or bS21-2-containing ribosomes. To identify such an element, we compiled 5′ UTR sequences including 100 nt upstream of the start codon and the first six codons of the gene for all proteins that were significantly less abundant in cells lacking bS21-2 compared to wild-type ($n = 74$; 100 nt was arbitrarily chosen because most *F. tularensis* transcription start sites have not been identified). As a control, we also compiled 5′ UTR sequences from 82 genes that were not impacted by bS21-2 presence. Using the motif-finding algorithm STREME, which identifies ungapped motifs enriched in large data sets (18), we identified

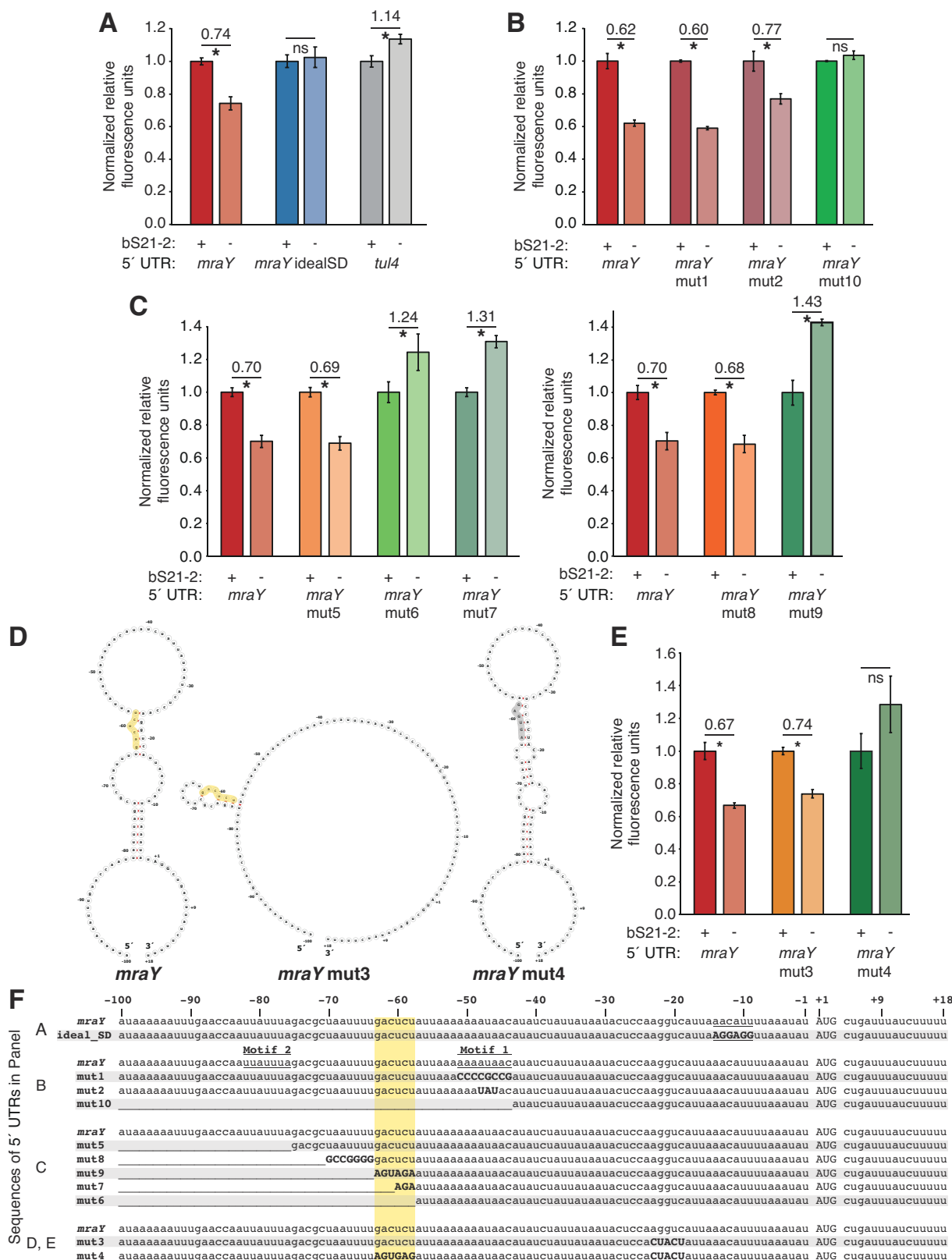


FIG 3 bS21-2-mediated translation of *mraY* depends on a weak SD and a specific six nucleotide sequence. Charts show relative fluorescence for indicated *gfp* translation fusion reporters in cells with (+; WT) or without (–; $\Delta rpsU2$) bS21-2. (A) Introduction of an ideal SD in the *mraY* leader leads to loss of bS21-2 responsiveness. (B) A motif common in bS21-2-responsive leader sequences (Motif 1, see F and Fig. S5) is not necessary for bS21-2-responsive translation in (Continued on next page)

FIG 3 (Continued) (A) A truncation that removes both motifs enriched in bS21-2-responsive leaders is no longer responsive (mut10). (C) The nucleotides between –58 and –63 in the *mraY* 5′ UTR, GACUCU, are essential for responsiveness to bS21-2. (D) Secondary structure predictions of wild-type and modified *mraY* 5′ UTRs. The six nt sequence necessary for responsiveness to bS21-2 is highlighted in yellow in *mraY* and *mraY* mut3 structures, and the location of mutated residues is highlighted in gray in *mraY* mut4. (E) Changing the *mraY* 5′ UTR secondary structure does not impact responsiveness to bS21-2. (A–C, E) Lines above bars indicate comparisons, values above line indicate ratio of reporter activity in cells lacking bS21-2 to wild-type cells. Error bars represent 1 SD. * $P < 0.05$ by *t*-test. ns = not significant. Experiments were repeated at least twice in biological triplicate, and data from a representative experiment are shown. Unnormalized fluorescence values can be found in Fig. S4. (F) Alignment of tested modifications to the *mraY* 5′ UTR. Sequences are grouped according to the panel in which they are found. Modifications from the wild-type leader sequence are capitalized and bold. Predicted and modified SD sequences are underlined in the sequences corresponding to panel A. Sequences corresponding to motifs enriched in bS21-2-responsive leaders are indicated. The key six nt sequence necessary for responsiveness to bS21-2 is found at nt –58 to –63 and is highlighted in yellow.

sequence motifs enriched in the 5′ UTRs of the 20 genes most positively governed by bS21-2 (14). The two motifs we identified were enriched compared to shuffled sequences and were not found to be enriched in the control sequences. These motifs, which we refer to as Motif 1 and Motif 2, are AU-rich (Motif 1 consensus AAAAUAAA and Motif 2 consensus UUAUUUA) and are found in 19 and 18 of the 20 sequences assessed, respectively (Fig. S5).

The *mraY* 5′ UTR contains sequences corresponding to both Motifs 1 and 2, so we generated targeted mutations to these sequences to assess their impact on responsiveness to bS21-2. Mutations 1 and 2 modified Motif 1 from AAAAUAAAC to CCCC GCCG, which altered the AU content of the entire motif, and AAAUAUACA, which altered the three most conserved nt in the motif. When assessed using the GFP reporter assay, neither of these modifications altered the responsiveness of the 5′ UTR to bS21-2 (Fig. 3B and F). To assess the contribution of Motif 2, we created mutation 5, a truncation of the 5′ end of the *mraY* 5′ UTR that removed 25 nt including Motif 2 (Fig. 3F). This mutant 5′ UTR also remained responsive to bS21-2 (Fig. 3C). We additionally created mutation 10, a truncation of the 5′ end of the *mraY* 5′ UTR which removes both Motif 1 and 2 (Fig. 3F). Notably, this mutation led to loss of responsiveness to bS21-2 (Fig. 3B). Together, these data indicate that while neither of the two STREME-predicted AU-rich motifs are necessary for the positive impact of bS21-2 on translation of the *mraY* 5′ UTR, some sequence between Motifs 1 and 2 is key for bS21-2 responsiveness.

A six nucleotide region of the *mraY* 5′ UTR is critical for bS21-2 responsiveness

While testing the importance of motifs enriched in bS21-2-responsive 5′ UTRs to responsiveness of the *mraY* 5′ UTR, our analysis revealed that the region key for bS21-2 responsiveness is between nucleotides 52 and 75 upstream from the initiation codon. To further clarify the sequence necessary for bS21-2 responsiveness in the *mraY* 5′ UTR, we made a series of truncations and modifications from the 5′ end of the leader sequence. Modifying the AU-rich region located 64–70 nt from the initiation codon (*mraY* mut8) did not impact bS21-2 responsiveness (Fig. 3C and F). But truncating the 5′ UTR to 57 nt (*mraY* mut6) led to loss of bS21-2 responsiveness, implicating the nucleotides 58–63 upstream of the initiation codon, GACUCU, in responsiveness to bS21-2 (Fig. 3C and F). We further assessed the importance of the GACUCU sequence using *mraY* mut7 (truncating the 5′ UTR to 60 nt and changing the first three nucleotides to AGA) and *mraY* mut9 (truncating the 5′ UTR to 63 nt and mutating nucleotides 58–63 to AGUGAG) and found neither was responsive to bS21-2 (Fig. 3C and F). These data allow us to conclude that the nucleotides 58–63 upstream of the *mraY* initiation codon, GACUCU, are critical for bS21-2-responsive translation of the *mraY* 5′ UTR. This is consistent with a model in which bS21-2-containing ribosomes interact directly or indirectly with a specific element of the leader sequence to facilitate efficient translation initiation on some transcripts.

Predicted secondary structures of 5' UTRs are not responsible for bS21-2 responsiveness

The secondary structure of mRNA molecules is an important determinant of translation initiation efficiency (19, 20). We considered the possibility that the key region necessary for bS21-2 responsiveness, the GACUCU sequence found 58–63 nt upstream of the *mraY* initiation codon, is important because it is part of a secondary structure recognized by bS21-2-containing ribosomes. In fact, in the predicted secondary structure of the *mraY* 5' UTR, the key GACUCU sequence base-pairs with the sequence 18–23 nt upstream of the initiation codon and 3 nt upstream from the predicted SD sequence (Fig. 3D). This base-pairing forms a stem-loop structure dependent on the key GACUCU sequence that could be found in close proximity to bS21 during translation initiation. To test the relevance of this potential structure, we mutated the sequence predicted to base-pair with the key region for bS21-2 responsiveness, at –23 to –18, thereby disrupting the structure (*mraY* mut3) (Fig. 3D and F). We also made complementary mutations to restore the structure (*mraY* mut4) (Fig. 3D and F) and assessed these 5' UTRs in a GFP reporter assay (Fig. 3E). We found that the disruption to the predicted secondary structure (*mraY* mut3) did not affect bS21-2 responsiveness. However, the complementary mutation, which altered the previously identified GACUCU sequence but restored the stem-loop structure, was no longer responsive to bS21-2 (Fig. 3E). These results reveal that the GACUCU sequence, rather than the structure it participates in forming, is key for conferring responsiveness of the *mraY* 5' UTR to bS21-2.

It is worth noting that loss of responsiveness to bS21-2 is not simply due to more or less protein synthesis. While many of the modifications that result in non-responsive leader sequences reduced overall reporter protein production (*mraY* mut4, 6, 7, 9), the *mraY* mut10 leader is similarly non-responsive but leads to increased reporter protein (Fig. S4).

In contrast to the *mraY* 5' UTR, the *pdpA* 5' UTR does not contain the sequence GACUCU, nor is it long enough to contain sequence at a conserved distance from the initiation codon (GACUCU is 58–63 nt upstream of the initiation codon; the *pdpA* 5' UTR is 24 nt long). We considered the possibility that, in contrast to the *mraY* 5' UTR, a structural element of the *pdpA* 5' UTR might confer responsiveness to bS21-2. We predicted the secondary structure of the *pdpA* 5' UTR using MXfold2, made targeted mutations to disrupt the predicted stem-loop structure (*pdpA* mut1), and generated β -galactosidase reporters at the Tn7 site in cells with (WT) and without ($\Delta rpsU2$) bS21-2 (Fig. S6). We also made complementary mutations to restore the original predicted secondary structure without maintaining the original sequence (*pdpA* mut2). In designing each mutation, we ensured that there was no significant disruption to the Shine-Dalgarno or start codon. Neither of these variants that altered the *pdpA* 5' UTR structure affected responsiveness to bS21-2, indicating that the secondary structure of this 5' UTR does not play a role in translation modulation by bS21-2 (Fig. S6). Thus, in our studies of two different bS21-2-responsive 5' UTRs, we did not find a specific predicted secondary structure that is necessary for bS21-2-responsive translation.

bS21-2 and Hfq influence T6SS protein abundance via different mechanisms

While we identified a key sequence for bS21-2 responsiveness in the *mraY* 5' UTR, neither this sequence nor a similar sequence is found in other 5' UTRs positively impacted by bS21-2 (Table S1), suggesting that other responsive 5' UTRs contain distinct sequences that similarly confer responsiveness or that there are multiple mechanisms by which bS21-2 exerts its effects. In considering the control of the T6SS proteins, we hypothesized that bS21-2 may be exerting its effects indirectly, altering the translation of a regulator that directly controls production of the T6SS proteins. One such candidate is the RNA-binding protein Hfq, which is known to control gene expression post-transcriptionally in many organisms and has been shown to impact the expression of T6SS proteins in *F. tularensis* (21, 22). Our proteomics analysis identified increased Hfq in cells lacking bS21-2 compared to wild-type (5.9-fold), although substantial variation in

biological replicates precluded the differences from reaching statistical significance (adj $P = 0.066$) (14). Additionally, many of the bS21-2-responsive 5' UTRs are AU-rich, raising the possibility that they may contain ARN motifs, known targets of Hfq. In fact, one of the motifs we found enriched in bS21-2-responsive 5' UTRs, Motif 1, could represent an ARN motif (Fig. S5). This led us to hypothesize that Hfq may play a role in bS21-2-mediated regulation of T6SS proteins.

To test further whether cells lacking bS21-2 have increased Hfq, we added DNA specifying a C-terminal vesicular stomatitis virus glycoprotein (VSV-G) epitope tag to *hfq* at its native locus on the chromosome, in cells with (WT) and without ($\Delta rpsU2$) bS21-2 and determined relative protein abundance by quantitative immunoblotting. We found a moderate increase in Hfq (about 30%, Fig. 4B) in cells lacking bS21-2 compared to wild-type; this is consistent with our prior proteomics findings, including that it is not a difference that would have reached our significance threshold.

Given the detectable increase in Hfq abundance in cells lacking bS21-2, we sought to determine if this is due to increased translation of the *hfq* mRNA. Using our reporter assay, we assessed the relative translation of mRNAs containing the *hfq* 5' UTR. We found that there is approximately two-fold more GFP produced from reporter fusions with the 5' UTR of *hfq* in cells lacking bS21-2, indicating this UTR leads to more efficient translation by ribosomes without bS21-2 (Fig. 4C). This suggests that the observed increase in Hfq in cells lacking bS21-2 is due to increased translation initiation.

If the moderate increase in Hfq leads to the observed reduction in T6SS proteins when bS21-2 is absent, Hfq must be acting as a repressor of the T6SS. However, inconsistent results have been reported with respect to the role of Hfq in regulating the T6SS genes in *F. tularensis*. A transcriptomic analysis of cells lacking Hfq found that one of the two Francisella Pathogenicity Island (FPI) operons encoding the T6SS, the *pdpA* operon, was upregulated in *hfq* mutant cells (diagram of FPI in Fig. 4A) (21). In another report, a proteomic analysis of *hfq* mutant cells found that proteins encoded by the other FPI operon, the *iglA* operon, are less abundant but observed no change in the abundance of proteins encoded by the *pdpA* operon (22). To clarify the role of Hfq in the regulation of T6SS protein abundance, we determined the abundance of several T6SS proteins encoded on both operons in cells with and without Hfq by quantitative immunoblotting (Fig. 4A and D) (14). PdpB, encoded in the *pdpA* operon, was more abundant in cells without Hfq compared to wild-type (>2 -fold; $P < 0.01$), while IgIB and IgIA, encoded on the *iglA* operon, were not impacted by the loss of Hfq (Fig. 4D). This is in contrast to cells lacking bS21-2, which contained reduced amounts of all three proteins (Fig. 4D) (14). These data suggest that Hfq regulates expression of proteins encoded by the *pdpA* operon, but not the *iglA* operon, consistent with the observations of Meibom et al. (21).

Hfq can exert its effects through a variety of mechanisms, some of which result in changes in translation initiation. To determine if the presence of *F. tularensis* Hfq, like bS21-2, impacts the translation of T6SS proteins, we analyzed the ability of cells lacking Hfq to translate mRNAs containing either the 5' UTRs of the T6SS protein gene *pdpA* or the control gene *tul4* fused to *gfp*. We found that translation of mRNAs with the *pdpA* or *tul4* 5' UTRs is not altered when Hfq is absent compared to wild-type, while translation of the mRNA with the *pdpA* 5' UTR decreases if bS21-2 is absent (Fig. 4E). These findings suggest that while bS21-2-associated changes in T6SS proteins can be attributed to changes in translation initiation, Hfq-associated changes in T6SS proteins cannot.

Since it has been proposed that Hfq represses *pdpA* operon transcript abundance (21) and Hfq does not impact translation of *pdpA* in a 5' UTR-dependent manner, we hypothesized that cells lacking Hfq might have increased PdpB due to increased *pdpB* (and *pdpA* operon) transcript abundance. To test this, we compared the abundance of mRNAs isolated from wild-type cells to those isolated from cells lacking bS21-2 ($\Delta rpsU2$) or Hfq (Δhfq) by qPCR and found that *pdpA* and *pdpB* transcripts have large, statistically significant increases when Hfq is not present (9.9-fold and 24-fold, respectively), but the relative impact on *iglA* transcript abundance is minor (2.5-fold increased) (Fig. 4F; Table

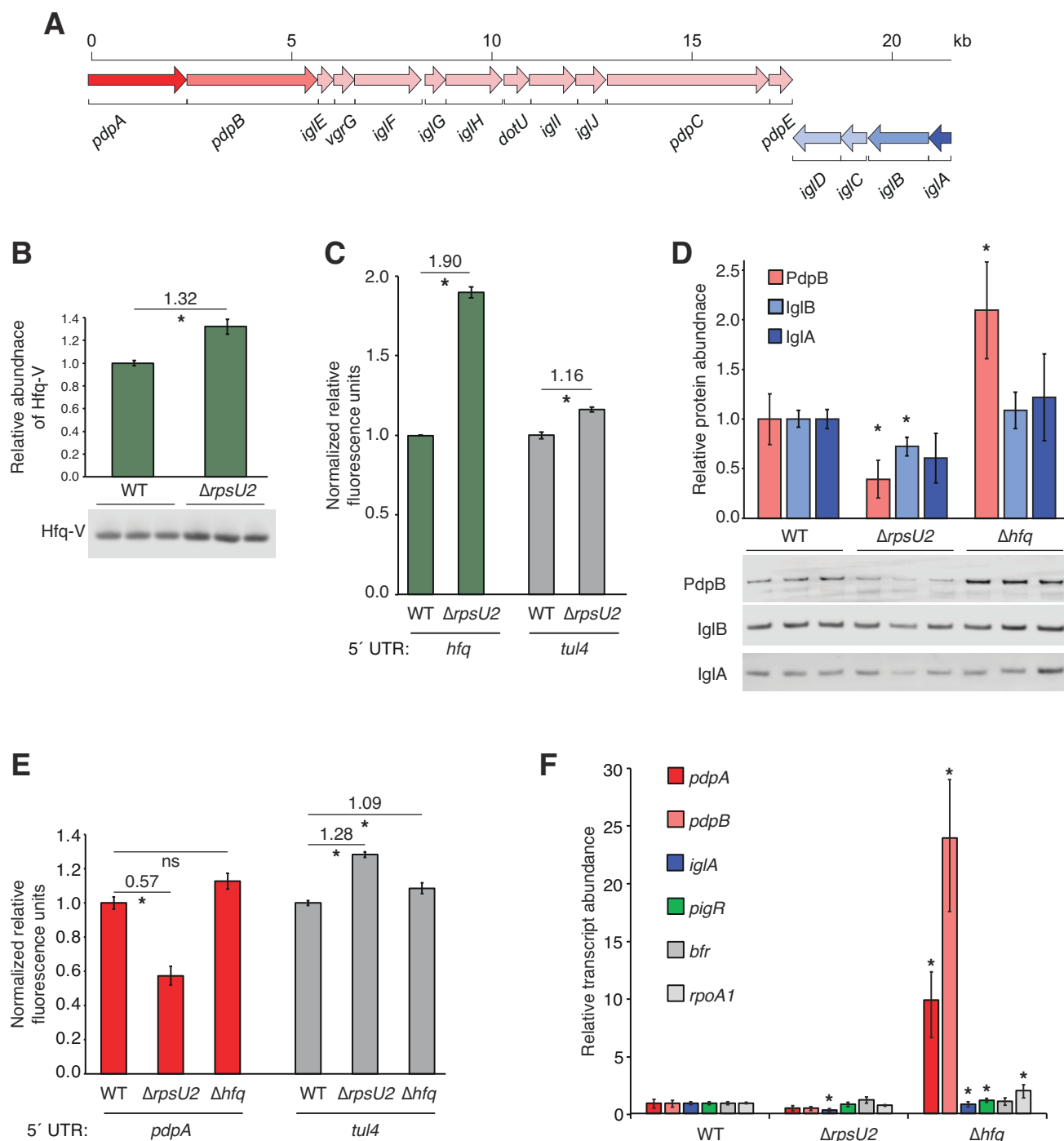


FIG 4 Hfq and bS21-2 influence production of T6SS proteins via different pathways. (A) Diagram of the Francisella Pathogenicity Island (FPI) genes, which encode T6SS proteins. (B) Cells lacking bS21-2 have more Hfq. Bottom: immunoblots probed with anti-VSV-G antibody. Whole cell lysates from bacteria containing Hfq-VSV-G and either with (WT) or without ($\Delta rpsU2$) bS21-2. Top: quantification of immunoblots. Band intensities for each protein were normalized to total protein per lane on the membrane. (C) Loss of bS21-2 leads to more *hfq* translation. Relative fluorescence for indicated *gfp* translational fusion reporters in cells with (+; WT) or without (–; $\Delta rpsU2$) bS21-2. (D) Only some of the T6SS proteins are influenced by loss of Hfq. Bottom: immunoblots probed with antibodies to indicated T6SS proteins in lysates of WT cells, cells lacking bS21-2 ($\Delta rpsU2$), or cells lacking Hfq (Δhfq). Top: quantification of immunoblots. Band intensities for each protein were normalized to total protein per lane on the membrane. (E) Hfq does not influence translation of the T6SS protein PdpA. Relative fluorescence for indicated *gfp* translational fusion reporters in WT cells, cells lacking bS21-2 ($\Delta rpsU2$), or cells lacking Hfq (Δhfq). (F) Hfq is a negative regulator of transcript abundance for some T6SS genes. Quantitative RT-PCR was used to determine the relative transcript abundance for indicated FPI-encoded genes in WT cells, cells lacking bS21-2 ($\Delta rpsU2$), or cells lacking Hfq (Δhfq), normalized to the *tul4* gene. The *rpoA1* and *bfr* genes are included as additional negative controls, as (Continued on next page)

FIG 4 (Continued)

their expression is not meaningfully influenced by bS21-2. (B–F) Lines above bars indicate comparisons, values above line indicate ratio of reporter activity in cells lacking bS21-2 to wild-type cells. Error bars represent 1 SD. Experiments were repeated at least twice in biological triplicate, and data from a representative experiment are shown. (A–D) * $P < 0.05$ after Bonferroni correction. (F) * $P < 0.005$ after Bonferroni correction, compared to WT.

S2); this is in contrast to the minor changes in *pdpA*, *pdpB*, and *iglA* transcripts in cells without bS21-2 (~2-fold or less).

Our findings indicate that Hfq exerts its effects on a subset of T6SS proteins by negatively regulating their transcript abundance (Fig. 4F). Conversely, bS21-2 influences the abundance of essentially all the T6SS proteins and does so by influencing their protein synthesis rather than their transcript abundance (Fig. 4E) (14). Given that Hfq and bS21-2 influence distinct groups of proteins at different points in gene expression, our work does not support a model in which Hfq mediates the effects of bS21-2 on the T6SS. Instead, our results suggest two distinct pathways of regulation for the genes encoding the T6SS: one in which bS21-2 improves the efficiency of translation initiation from both operons and one in which Hfq represses transcript abundance of only the *pdpA* operon.

DISCUSSION

In this work, we addressed two hypotheses regarding how bS21-2 exerts its effects on protein synthesis. The first is that ribosomes containing bS21-2 may influence translation initiation of specific transcripts in a leader sequence-dependent manner. Using reporter assays, we determined that specific 5' UTR sequences are sufficient to lead to altered translation in cells with or without bS21-2. In a comprehensive assessment of 5' UTR elements, we found that bS21-2-responsive 5' UTRs have imperfect SD sequences and, in a specific responsive 5' UTR, the presence of a particular six nucleotide sequence. Our second hypothesis was that the effects of bS21-2 on the T6SS proteins may be mediated by Hfq, a known regulator of T6SS proteins. However, since there have been conflicting reports regarding the impacts of Hfq on the T6SS, we also examined the effects of Hfq on T6SS protein and transcript abundance. Our work clearly demonstrates that Hfq is a negative regulator of genes in one of the two FPI operons encoding T6SS proteins and that this regulation influences transcript abundance rather than translation, consistent with and building upon a prior study (21). In contrast, the positive effects of bS21-2 on essentially all T6SS proteins can be attributed to differences in protein synthesis. Thus, we conclude that Hfq and bS21-2 function in independent pathways to regulate the T6SS proteins. Together, these results suggest that bS21-2 impacts protein synthesis by altering translation initiation on mRNAs with specific leader sequences.

This work demonstrates that bS21-2 exerts its effects on protein synthesis in a leader sequence-dependent manner and is validated in a subset of bS21-2-responsive 5' UTRs (Fig. 1B). While loss of bS21-2 results in altered abundance for about 160 proteins (14), we expect that changes in protein abundance for at least some of these may not be due to leader sequence-dependent effects but rather due to downstream or secondary effects. For example, bS21-2 may directly impact synthesis of proteins that influence the abundance of other proteins. Several proteases and peptidases were found to be differentially abundant in cells lacking bS21-2 (14). Thus, proteins like those encoded by FTL_0881 and FTL_0215, which do not have bS21-2-responsive leader sequences (Fig. S2), may have altered abundance in bS21-2 mutant cells due to changes in the abundance of proteases or protein processing enzymes.

In our search for an element necessary for responsiveness of leader sequences to bS21-2, we found that ideal SD sequences prevent responsiveness (Fig. 2C and 3A). This suggests that perfect SD-ASD complementarity leads to such efficient translation initiation that any contribution of bS21-2 to translation is minor and effectively masked. It is perhaps unsurprising that other regulators of translation, such as H-NS in *E. coli*, similarly function to regulate translation of mRNAs with imperfect SDs (23). We suggest this effect is due to the SD-ASD complementarity leading to increased stabilization of the 30S initiation complex, obscuring any positive effects of bS21-2, rather than due to overall changes in protein production. In particular, simply increasing translation is not

sufficient to lead to loss of responsiveness to bS21-2, as very similar amounts of protein are produced from reporters with the *pdpA tul4SD* and *pdpA idealSD* leaders, but only *pdpA idealSD* is not responsive to bS21-2 (Fig. 2C; Fig. S3). And similarly, reducing protein synthesis is not sufficient to prevent responsiveness to bS21-2. Moving the ideal SD sequence to a non-ideal location in translational fusions (*pdpA ideal_movedSD*) reduces the amount of reporter protein but still prevents responsiveness to bS21-2 (Fig. 2C; Fig. S3). In this case, we suggest that while the ideal SD prevents bS21-2 responsiveness, the reduced reporter abundance may be because the extended spacing between the SD and start codon influences the transition from initiation to elongation, affecting the rate of ribosome translocation immediately following the start codon (24).

While a simple common element across all bS21-2-responsive leader sequences remains elusive, we were able to hone in on the six nt sequence in the *mraY* 5' UTR that leads to bS21-2 responsiveness, GACUCU (Fig. 3). We speculate that the six nt GACUCU sequence may be a direct binding site for bS21-2 and that a direct interaction between this site and bS21-2 in the ribosome increases translation initiation of the *mraY* mRNA. However, this specific sequence is not found in other leader sequences positively impacted by bS21-2 (including the short, 24 nt *pdpA* 5' UTR; Table S1). Further work will be necessary to determine if this sequence is sufficient for bS21-2 responsiveness, to identify the commonalities among bS21-2-responsive 5' UTRs, and to determine how bS21-2 influences translation initiation on specific leader sequences.

While the key six nt sequence identified in the *mraY* 5' UTR is not universally required for responsiveness to bS21-2, it does reveal that specific leader sequences lead to increased translation by ribosomes with bS21-2. This is quite different from bS21 homologs in *Bacteroidia*, which also influence translation in a sequence-dependent manner, but exert their effects by altering accessibility to the anti-Shine-Dalgarno sequence (12, 17). This raises the possibility that other bS21 homologs, in *F. tularensis* and other organisms, may influence translation in a leader sequence-dependent manner.

In *F. tularensis*, loss of the RNA chaperone Hfq results in defective intramacrophage replication, which is essential for virulence. Yet how Hfq promotes *F. tularensis* intramacrophage replication remains poorly understood. Few small RNAs have been identified in *F. tularensis*, and none have been identified that are Hfq-dependent (25, 26). Our results demonstrate that Hfq represses *pdpA* operon transcript abundance but does not influence T6SS protein synthesis. The molecular mechanism by which Hfq exerts its effects on this operon, and if it involves a small RNA, remains unclear. Regardless, the change in production of T6SS components in cells lacking Hfq is consistent with the observed intramacrophage growth defect, as *F. tularensis* cells overproducing the T6SS are defective for intramacrophage survival (27).

MATERIALS AND METHODS

Bacterial strains and growth conditions

Unless otherwise noted, bacterial strains were grown as indicated. *Francisella tularensis* subsp. *holarctica* Live Vaccine Strain (LVS) cells were grown in Mueller-Hinton broth (BD Difco) supplemented with 0.025% iron pyrophosphate, 0.1% glucose, and 2% Isovitalex (sMHB), shaking aerobically or on cystine heart agar (BD Difco or prepared in house) plates with 1% hemoglobin (CHA-H) at 37°C. *Escherichia coli* XL1-Blue, DH5α (New England Biolabs), and DH5α λ-pir cells were grown in lysogeny broth (LB) shaking aerobically or on LB agar plates at 37°C. Kanamycin and nourseothricin were used at concentrations of 5 µg/mL (*F. tularensis*) or 50 µg/mL (*E. coli*); hygromycin B was used at concentrations of 200 µg/mL. *Saccharomyces cerevisiae* cells were grown in synthetic defined (SD) broth without uracil (-ura) shaking aerobically or on SD-ura agar plates at 30°C.

Vector construction

Tn7::lacZ plasmids

Mini-Tn7 plasmids for each β -galactosidase reporter were created from a plasmid derived from pMP749 (28). *E. coli lacZ* was amplified from pEX-pigR::lacZ (29) using a 5' primer specifying a NotI site and alanine linker (5'-GCGGCCGCT-3') and a 3' primer specifying a BamHI site. The amplified *lacZ* gene was cloned into NotI/BamHI-digested pMP749, resulting in pKR68 (Tn7-*lacZ*). Subsequently, two fragments were amplified from LVS genomic DNA (gDNA): (i) the *tul4* promoter with a 5' primer specifying a KpnI site and a 3' primer overlapping the second fragment; and (ii) either modified or wild-type UTRs from genes of interest, along with the first six codons of the corresponding gene, with a 3' primer specifying a NotI site and a 5' primer overlapping the first fragment. Overlap extension PCR was then conducted on the two fragments, and the PCR product was cloned into KpnI/NotI-digested pKR68 such that *lacZ* was in-frame with the first six codons of the gene of interest. The resulting plasmids are all indicated in Table S3. Modifications to wild-type UTRs were encoded on primers for PCR amplification.

Some reporter plasmids with the high-copy pUC *ori* produced enough β -galactosidase in *E. coli* to be toxic, so cloning required one of two alternate approaches. In one approach, the origin of the pMP749 plasmid was replaced by a low-copy R6K γ origin, amplified from pKL91 (30) using primers that encode an NspI site. The digested product was cloned into NspI-digested pMP749, resulting in pKR88 (Mini-Tn7_R6K γ), which was propagated in DH5 α λ -pir cells. Subsequently, the *tul4* promoter and 5' UTR were amplified from LVS gDNA using a 5' primer specifying a KpnI site and a 3' primer specifying a NotI site; *lacZ* was amplified from pKR68 using a 5' primer specifying a NotI site and alanine linker (5'-GCGGCCGCT-3') and a 3' primer specifying a BamHI site. The two fragments were cloned into BamHI/KpnI-digested pKR88 using a three-way ligation, resulting in pKR89 (Tn7_P*tul4*_tul4UTR-*lacZ*_R6K γ ; Table S3).

In a second approach, *lacZ* plasmids were cloned using *S. cerevisiae*. The 2 μ origin and *URA3* gene were isolated from pYES2 (Invitrogen) by digestion with PstI, then cloned into DraI-digested pKR68, disrupting the β -lactamase gene. The resulting plasmid, pKR128 pYES2 Tn7-*lacZ*, was used for subsequent cloning of 5' UTRs using alanine linkers and NotI sites, as described above and detailed in Table S3. pYES2-based plasmids were purified from overnight cultures of *S. cerevisiae* using the Zymoprep Yeast Plasmid Miniprep II kit.

pF-GFP plasmids

Multicopy GFP reporter plasmids were created from a previously described shuttle vector, pFNLT6 (31). A fragment containing the promoter, 5' UTR, and first six codons of *tul4* was digested from pKR89 with KpnI/NotI. sfGFP codon-optimized for expression in *E. tularensis* LVS was purchased as a gBlock (IDT) and digested with NotI/BamHI. Fragments were cloned into KpnI/BamHI-digested pF such that GFP was in-frame with the first six codons of *tul4*, resulting in pKR145 (pF-*tul4* UTR-GFP). The plasmid pKR146 (pF-*pdpA* UTR-GFP) was constructed similarly, after amplification from pKR74 of the *tul4* promoter and *pdpA* 5' UTR and first six codons and digestion of the PCR product with KpnI/NotI. Subsequent constructs were cloned into pKR145 to replace the *tul4* 5' UTR using the endogenously-encoded PacI site in the *tul4* promoter (Table S3 for details). For genes in which a transcription start site had not been annotated at the time of plasmid design, 100 nucleotides upstream of the start codon were included as the 5' UTR (Table S3 for details). Known transcription start sites for *tul4*, *iglA*, and *pdpA* were previously published by Ramsey et al. (32); the transcription start site for *hfq* was experimentally determined by Meibom et al. (21), and Chambers and Bender (33).

pF-nat complementation vector

The pF-*rpsU2*-V plasmid (pKR7) (14) was modified to replace the kanamycin resistance cassette with a nourseothricin (*nat*) resistance cassette from pF3-MgIA-V as previously described (34), yielding pKR15 pF-*nat-rpsU2*-V.

Allelic exchange plasmid

The plasmid pEX18kan was modified to create the in-frame deletion construct for deletion of *hfq* as previously described (35). Flanking regions of ~1,000 base pairs from both sides of the *hfq* gene were amplified by PCR. Primers amplifying the DNA adjacent to *hfq* included the first three or last three codons of the open reading frame and DNA specifying a NotI site, which also encodes an alanine linker (5'-GCGGCCGCT-3'). The two fragments were cloned into BamHI/KpnI-digested pEX18kan, yielding pKL111 pEX Δ *hfq*.

VSV-G-tagging integration vector

A single integration vector for VSV-G tagging of *hfq* was made by modifying pKL02 (32). The final 200 nucleotides of the 3' end of *hfq* were amplified using a 5' primer specifying a KpnI site and a 3' primer that lacked the native stop codon and included DNA specifying a NotI site. The fragment was cloned into KpnI/NotI-digested pKL02 such that the 3' end of *hfq* is in-frame with the codons specifying three alanine residues followed by the VSV-G epitope, resulting in pKR158 (pEX-*hfq*-V).

Strain construction

β -galactosidase reporter strains (Table 1) were constructed by site- and orientation-specific single chromosome integration using the Tn7 transposon as previously described (28). Helper plasmid pMP720 was electroporated into either wild-type (LVS) or bS21-2 mutant (Δ *rpsU2*) competent cells in 0.2 cm cuvettes with a 2.5-kV pulse, and hygromycin-resistant cells were selected by plating on CHA-H with hygromycin. Cells with the helper plasmid were electroporated with the appropriate mini-Tn7 plasmid and selected for on CHA-H with kanamycin. Colonies were screened for plasmid integration at the *attTn7* site using PCR. Candidate strains were confirmed by amplification of genomic DNA outside of the *attTn7* site and Sanger sequencing.

Reporter constructs encoded on pF plasmids were electroporated into LVS, LVS Δ *rpsU2*, or LVS Δ *hfq* cells as described above and selected for on CHA-H with kanamycin (Table 1). The complementation vectors or empty pF plasmids were electroporated into β -galactosidase reporter strains as described above and selected for on CHA-H with nourseothricin.

The *hfq* deletion strain was constructed by allelic exchange as previously described (14). Briefly, at least 1 μ g of allelic exchange plasmid pEX Δ *hfq* was electroporated into competent cells as above. Cells in which a single integration event occurred were selected for on CHA-H with kanamycin. Counterselection for the vector was accomplished by plating on CHA-H (BD Difco) containing 10% sucrose. Sucrose-resistant, kanamycin-sensitive colonies were screened for deletions using PCR. Candidate strains were confirmed by amplification of genomic DNA outside of the flanking regions on each side of the deletion and Sanger sequencing, validating LVS Δ *hfq*.

Cells with VSV-G-tagged Hfq were made as previously described (32). Briefly, at least 1 μ g of pKR158 pEX-*hfq*-V was electroporated into LVS and Δ *rpsU2* cells, and transformants were selected on CHA-H with kanamycin. Cells were confirmed to have a single integration by PCR amplification of DNA across the integration site and subsequent Sanger sequencing of the PCR product.

β -galactosidase assays

β -galactosidase assays using *F. tularensis* LVS or Δ *rpsU2* cells containing indicated reporter constructs were conducted as previously described (29). If significant yellow

TABLE 1 Strains used in this study

Strain number	Description	Genetic background	Plasmid
β-Galactosidase reporter strains			
KRLVS96	LVS Tn7::Ptul4-pdpA 5'UTR-lacZ <i>aphA</i>	LVS	pKR74
KRLVS97	LVS Δ rpsU2 Tn7::Ptul4-pdpA 5'UTR-lacZ <i>aphA</i>	Δ rpsU2	pKR74
KRLVS102	LVS Tn7::Ptul4-pdpA 5'UTR-mut1-lacZ <i>aphA</i>	LVS	pKR84
KRLVS106	LVS Δ rpsU2 Tn7::Ptul4-pdpA 5'UTR-mut1-lacZ <i>aphA</i>	Δ rpsU2	pKR84
KRLVS103	LVS Tn7::Ptul4-pdpA 5'UTR-mut2-lacZ <i>aphA</i>	LVS	pKR86
KRLVS108	LVS Δ rpsU2 Tn7::Ptul4-pdpA 5'UTR-mut2-lacZ <i>aphA</i>	Δ rpsU2	pKR86
KRLVS114	LVS Tn7::Ptul4-pdpA 5'UTR-badSD-lacZ <i>aphA</i>	LVS	pKR98
KRLVS117	LVS Δ rpsU2 Tn7::Ptul4-pdpA 5'UTR-badSD-lacZ <i>aphA</i>	Δ rpsU2	pKR98
KRLVS115	LVS Tn7::Ptul4-pdpA 5'UTR-ideal_movedSD-lacZ <i>aphA</i>	LVS	pKR99
KRLVS118	LVS Δ rpsU2 Tn7::Ptul4-pdpA 5'UTR-ideal_movedSD-lacZ <i>aphA</i>	Δ rpsU2	pKR99
KRLVS112	LVS Tn7::Ptul4-tul4 5'UTR-lacZ <i>aphA</i>	LVS	pKR89
KRLVS111	LVS Δ rpsU2 Tn7::Ptul4-tul4 5'UTR-lacZ <i>aphA</i>	Δ rpsU2	pKR89
KRLVS158	LVS Tn7::Ptul4-pdpA 5'UTR-idealSD-lacZ <i>aphA</i>	LVS	pKR129
KRLVS160	LVS Δ rpsU2 Tn7::Ptul4-pdpA 5'UTR-idealSD-lacZ <i>aphA</i>	Δ rpsU2	pKR129
KRLVS159	LVS Tn7::Ptul4-pdpA 5'UTR-tul4SD-lacZ <i>aphA</i>	LVS	pKR130
KRLVS161	LVS Δ rpsU2 Tn7::Ptul4-pdpA 5'UTR-tul4SD-lacZ <i>aphA</i>	Δ rpsU2	pKR130
KRLVS266	LVS Tn7::Ptul4-pdpA 5'UTR-lacZ <i>aphA</i> pF-nat	KRLVS96	pF-nat
KRLVS267	LVS Δ rpsU2 Tn7::Ptul4-pdpA 5'UTR-lacZ <i>aphA</i> pF-nat	KRLVS97	pF-nat
KRLVS268	LVS Δ rpsU2 Tn7::Ptul4-pdpA 5'UTR-lacZ <i>aphA</i> pF-nat-rpsU2-V	KRLVS97	pKR15
KRLVS269	LVS Tn7::Ptul4-tul4 5'UTR-lacZ <i>aphA</i> pF-nat	KRLVS112	pF-nat
KRLVS270	LVS Δ rpsU2 Tn7::Ptul4-tul4 5'UTR-lacZ <i>aphA</i> pF-nat	KRLVS111	pF-nat
pF-GFP reporter strains			
KRLVS180	LVS pF-tul4UTR-GFP	LVS	pKR145
KRLVS182	LVS Δ rpsU2 pF-tul4UTR-GFP	Δ rpsU2	pKR145
KRLVS234	LVS Δ hfq pF-tul4 UTR-GFP	Δ hfq	pKR145
KRLVS181	LVS pF-pdpAUTR-GFP	LVS	pKR146
KRLVS183	LVS Δ rpsU2 pF-pdpAUTR-GFP	Δ rpsU2	pKR146
KRLVS236	LVS Δ hfq pF-pdpA UTR-GFP	Δ hfq	pKR146
KRLVS188	LVS pF-mraYUTR-GFP	LVS	pKR151
KRLVS189	LVS Δ rpsU2 pF-mraYUTR-GFP	Δ rpsU2	pKR151
KRLVS190	LVS pF-FTL_0215UTR-GFP	LVS	pKR152
KRLVS191	LVS Δ rpsU2 pF-FTL_0215UTR-GFP	Δ rpsU2	pKR152

(Continued on next page)

TABLE 1 Strains used in this study (Continued)

Strain number	Description	Genetic background	Plasmid
KRLVS199	LVS pF- <i>mraY</i> UTR_mut1-GFP	LVS	pKR156
KRLVS200	LVS Δ <i>rpsU2</i> pF- <i>mraY</i> UTR_mut1-GFP	Δ <i>rpsU2</i>	pKR156
KRLVS201	LVS pF- <i>mraY</i> UTR_mut2-GFP	LVS	pKR157
KRLVS202	LVS Δ <i>rpsU2</i> pF- <i>mraY</i> UTR_mut2-GFP	Δ <i>rpsU2</i>	pKR157
KRLVS206	LVS pF- <i>iglA</i> UTR-GFP	LVS	pKR160
KRLVS207	LVS Δ <i>rpsU2</i> pF- <i>iglA</i> UTR-GFP	Δ <i>rpsU2</i>	pKR160
KRLVS208	LVS pF-FTL_0222UTR-GFP	LVS	pKR161
KRLVS209	LVS Δ <i>rpsU2</i> pF-FTL_0222UTR-GFP	Δ <i>rpsU2</i>	pKR161
KRLVS210	LVS pF-FTL_0881UTR-GFP	LVS	pKR162
KRLVS211	LVS Δ <i>rpsU2</i> pF-FTL_0881UTR-GFP	Δ <i>rpsU2</i>	pKR162
KRLVS212	LVS pF-FTL_1093UTR-GFP	LVS	pKR163
KRLVS213	LVS Δ <i>rpsU2</i> pF-FTL_1093UTR-GFP	Δ <i>rpsU2</i>	pKR163
KRLVS214	LVS pF- <i>mraY</i> UTR_mut5-GFP	LVS	pKR165
KRLVS215	LVS Δ <i>rpsU2</i> pF- <i>mraY</i> UTR_mut5-GFP	Δ <i>rpsU2</i>	pKR165
KRLVS222	LVS pF- <i>mraY</i> UTR_idealSD-GFP	LVS	pKR169
KRLVS223	LVS Δ <i>rpsU2</i> pF- <i>mraY</i> UTR_idealSD-GFP	Δ <i>rpsU2</i>	pKR169
KRLVS228	LVS pF- <i>hfq</i> UTR-GFP	LVS	pKR172
KRLVS229	LVS Δ <i>rpsU2</i> pF- <i>hfq</i> UTR-GFP	Δ <i>rpsU2</i>	pKR172
KRLVS243	LVS pF- <i>mraY</i> UTR_mut6-GFP	LVS	pKR175
KRLVS244	LVS Δ <i>rpsU2</i> pF- <i>mraY</i> UTR_mut6-GFP	Δ <i>rpsU2</i>	pKR175
KRLVS247	LVS pF- <i>mraY</i> UTR_mut7-GFP	LVS	pKR177
KRLVS248	LVS Δ <i>rpsU2</i> pF- <i>mraY</i> UTR_mut7-GFP	Δ <i>rpsU2</i>	pKR177
KRLVS252	LVS pF- <i>mraY</i> UTR_mut8-GFP	LVS	pKR179
KRLVS253	LVS Δ <i>rpsU2</i> pF- <i>mraY</i> UTR_mut8-GFP	Δ <i>rpsU2</i>	pKR179
KRLVS260	LVS pF- <i>mraY</i> UTR_mut9-GFP	LVS	pKR180
KRLVS261	LVS Δ <i>rpsU2</i> pF- <i>mraY</i> UTR_mut9-GFP	Δ <i>rpsU2</i>	pKR180
KRLVS262	LVS pF- <i>mraY</i> UTR_mut3-GFP	LVS	pKR182
KRLVS263	LVS Δ <i>rpsU2</i> pF- <i>mraY</i> UTR_mut3-GFP	Δ <i>rpsU2</i>	pKR182
KRLVS264	LVS pF- <i>mraY</i> UTR_mut4-GFP	LVS	pKR183
KRLVS265	LVS Δ <i>rpsU2</i> pF- <i>mraY</i> UTR_mut4-GFP	Δ <i>rpsU2</i>	pKR183
KRLVS197	LVS pF- <i>mraY</i> UTR_mut10-GFP	LVS	pKR155
KRLVS198	LVS Δ <i>rpsU2</i> pF- <i>mraY</i> UTR_mut10-GFP	Δ <i>rpsU2</i>	pKR155
Deletion strains			
KMLFT97	LVS Δ <i>hfq</i>	LVS	pKL111
VSV-G-tagged strains			
KRLVS194	LVS <i>hfq</i> -VSVG	LVS	pKR158
KRLVS195	LVS Δ <i>rpsU2</i> <i>hfq</i> -VSVG	Δ <i>rpsU2</i>	pKR158

color was not produced within 2 hours, reactions were stopped. Experiments were conducted at least twice in biological triplicate.

GFP assays

F. tularensis LVS, Δ *rpsU2*, or Δ *hfq* reporter constructs were grown in sMHB to mid-log phase in biological triplicate. Cells were pelleted and resuspended in PBS. OD₆₀₀ and fluorescence with excitation of 495 nm and emission of 535 nm were determined using ID3 plate reader (RI-INBRE CRCF), in technical triplicate. Fluorescence readings were normalized to OD₆₀₀ and fluorescence of LVS cells (lacking any GFP reporter) was subtracted from each reading to account for basal level fluorescence of the cells. Experiments were conducted at least twice in biological triplicate.

Plasmid copy number qPCR

Wild-type LVS or LVS $\Delta rpsU2$ cells with pF plasmids were grown to mid-log ($OD_{600} = 0.3\text{--}0.4$). Total DNA was extracted from 1 mL culture using the MasterPure Complete DNA Purification Kit (Lucigen). qRT-PCR was performed using PowerUp SYBR Green Master Mix (Applied Biosystems) and a Roche LightCycler 480 (RI-INBRE CRCF) essentially as described (35) with 0.05 ng of DNA. DNA abundances were calculated for an open reading frame (ORF3) on the plasmid, and relative abundance is reported compared to a chromosomally-encoded control gene, *tul4*. Experiments comparing wild-type and *rpsU2* mutant cells were performed three times in biological triplicate.

5' UTR secondary structure prediction

Secondary structures reported above were predicted using the MXfold2 web server (36).

5' UTR motif analyses

The 5' UTRs of genes that have decreases in protein, but not transcript, abundance in cells lacking bS21-2 compared to wild type (14) were analyzed. As there is insufficient annotation of transcriptional start sites in *F. tularensis* LVS, 100 nucleotides upstream of the start codon along with the first six codons were chosen for analysis.

STREME software (MEME suite) was used to analyze the 5' UTRs of the 20 genes with the largest fold decreases in protein in $\Delta rpsU2$ cells compared to LVS. These were compared to shuffled sequences to find two candidate motifs. As a control, the same parameters were used to compare the predicted 5' UTRs of 20 genes not differentially expressed in LVS and $\Delta rpsU2$.

Shine-Dalgarno predictions were made by highest similarity to the reverse complement of the anti-Shine-Dalgarno (5'-AGGAGG-3') within 20 nucleotides of the start codon.

Immunoblotting

Immunoblotting was completed as previously described (14). Briefly, cell lysates were separated by SDS-PAGE, transferred to PVDF, and analyzed for total protein with the Invitrogen No-Stain Protein labeling reagent for normalization. Membranes were probed with indicated monoclonal antibodies (BEI Resources, diluted 1:250 for PdpB, 1:1,000 for IgIB, and 1:1,000 for IgIA) or the VSV-G epitope tag (Sigma, diluted 1:2,222) in blocking buffer. Proteins were detected using IRDye 800 CW donkey anti-mouse IgG or donkey anti-rabbit IgG secondary antibodies (Li-Cor, diluted 1:10,000). Protein abundance was calculated as fluorescence of protein bands relative to total protein in each lane. Experiments were performed at least twice in biological triplicate.

RNA purification and qRT-PCR

Total RNA was purified according to the *RNAsnap* protocol (37). *F. tularensis* LVS was grown in biological triplicate to mid-log phase. Pelleted cells (10 mL) were resuspended in 100 μ L of fresh RES (95% formamide, 18 mM EDTA, 0.025% SDS, 1% BME), then incubated at 95°C for 7 minutes. Cell debris was pelleted by centrifugation, and the supernatant was preserved. Total nucleic acid was recovered with 0.3 M sodium acetate (pH 5.2) and 3 \times volumes 100% ethanol. Samples were stored at -80°C for 1 h, then nucleic acid was pelleted by centrifugation at 4°C for 30 minutes. The pellet was washed with 75% ethanol and resuspended in water. Purified nucleic acids were treated with RQ1 DNase (Promega) for 1 h at 37°C, and RNA was purified again with sodium acetate/ethanol precipitation.

cDNA was synthesized using Superscript III reverse transcriptase (Life Technologies) as previously described (35). qRT-PCR was performed with the PowerUP SYBR Green Master Mix (Applied Biosystems) and the Roche Lightcycler 480 (RI-INBRE CRCF). Transcript abundances of *pdpA*, *pdpB*, *iglA*, *pigR*, *rpoA1*, and *bfr* were normalized to a control gene, *tul4*. Experiments were conducted twice in biological triplicate.

ACKNOWLEDGMENTS

We thank the other members of the Ramsey laboratory for support and helpful discussions and Dr Simon Dove for his helpful comments on the manuscript and his support of K.M.R. as a post-doctoral fellow when she constructed the *F. tularensis* Δhfq strain. We also thank Janet Atoyan and the Rhode Island INBRE Molecular Informatics Core.

This work was funded by an NIGMS CARTD-COBRE Pilot Project Award (P20GM121344-KMR), an NIGMS/RI-INBRE Early Career Development Award (P20GM103430-KMR), and a Rhode Island Foundation Medical Research Grant (2798_20190602-KMR). This work was supported by the USDA National Institute of Food and Agriculture, Hatch Formula project accession number 1017848. This material is based upon work conducted at a Rhode Island NSF EPSCoR research facility, the Genomics and Sequencing Center, supported in part by the National Science Foundation EPSCoR Cooperative Agreements 0554548, EPS-1004057, and OIA-1655221. The research was made possible using equipment and services available through the Rhode Island Institutional Development Award (IDeA) Network of Biomedical Research Excellence from the National Institute of General Medical Sciences of the National Institutes of Health under grant number P20GM103430 through the Centralized Research Core facility and the Molecular Informatics Core (RRID:SCR_017685). Antibodies for the following *Francisella tularensis* proteins were obtained through BEI Resources, NIAID, NIH:PdpB, IgIA, and IgIB.

AUTHOR AFFILIATIONS

¹Department of Cell and Molecular Biology, University of Rhode Island, Kingston, Rhode Island, USA

²Department of Biomedical and Pharmaceutical Sciences, University of Rhode Island, Kingston, Rhode Island, USA

AUTHOR ORCIDs

Kathryn M. Ramsey  <http://orcid.org/0000-0002-9014-4416>

FUNDING

Funder	Grant(s)	Author(s)
HHS NIH National Institute of General Medical Sciences (NIGMS)	P20GM121344	Kathryn M. Ramsey
HHS NIH National Institute of General Medical Sciences (NIGMS)	P20GM103430	Kathryn M. Ramsey
Rhode Island Foundation	2798_20190602	Kathryn M. Ramsey
U.S. Department of Agriculture (USDA)	1017848	Kathryn M. Ramsey

AUTHOR CONTRIBUTIONS

Hannah S. Trautmann, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Writing – original draft, Writing – review and editing | Sierra S. Schmidt, Investigation, Validation | Steven T. Gregory, Formal analysis, Software, Visualization, Writing – review and editing | Kathryn M. Ramsey, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review and editing

ADDITIONAL FILES

The following material is available [online](#).

Supplemental Material

Fig S1 (JB00140-23-S0001.eps). Supplemental figure.

Fig S2 (JB00140-23-S0002.eps). Supplemental figure.

Fig S3 (JB00140-23-S0003.eps). Supplemental figure.

Fig S4 (JB00140-23-S0004.eps). Supplemental figure.

Fig S5 (JB00140-23-S0005.eps). Supplemental figure.

Fig S6 (JB00140-23-S0006.eps). Supplemental figure.

Legends for supplemental figures (JB00140-23-S00010.docx). Legends of Fig. S1 to S6.

TABLE S1 (JB00140-23-S0007.xlsx). Supplemental table.

TABLE S2 (JB00140-23-S0008.xlsx). Supplemental table.

TABLE S3 (JB00140-23-S0009.xlsx). Supplemental table.

REFERENCES

- Genuth NR, Barna M. 2018. The discovery of ribosome heterogeneity and its implications for gene regulation and organismal life. *Mol Cell* 71:364–374. <https://doi.org/10.1016/j.molcel.2018.07.018>
- Byrgazov K, Vesper O, Moll I. 2013. Ribosome heterogeneity: another level of complexity in bacterial translation regulation. *Curr Opin Microbiol* 16:133–139. <https://doi.org/10.1016/j.mib.2013.01.009>
- Ferretti MB, Karbstein K. 2019. Does functional specialization of ribosomes really exist? *RNA* 25:521–538. <https://doi.org/10.1261/rna.069823.118>
- Bubunenko M, Baker T, Court DL. 2007. Essentiality of ribosomal and transcription antitermination proteins analyzed by systematic gene replacement in *Escherichia coli*. *J Bacteriol* 189:2844–2853. <https://doi.org/10.1128/JB.01713-06>
- Shimojo M, Amikura K, Masuda K, Kanamori T, Ueda T, Shimizu Y. 2020. In vitro reconstitution of functional small ribosomal subunit assembly for comprehensive analysis of ribosomal elements in *E. coli*. *Commun Biol* 3:142. <https://doi.org/10.1038/s42003-020-0874-8>
- Mizushima S, Nomura M. 1970. Assembly mapping of 30s ribosomal proteins from *E. coli*. *Nature* 226:1214. <https://doi.org/10.1038/2261214a0>
- Held WA, Nomura M, Hershey JW. 1974. Ribosomal protein S21 is required for full activity in the initiation of protein synthesis. *Mol Gen Genet* 128:11–22. <https://doi.org/10.1007/BF00267291>
- Chang C, Craven GR. 1977. Identification of several proteins involved in the messenger RNA binding site of the 30 S ribosome by inactivation with 2-methoxy-5-nitropropone. *J Mol Biol* 117:401–418. [https://doi.org/10.1016/0022-2836\(77\)90135-8](https://doi.org/10.1016/0022-2836(77)90135-8)
- Van Duin J, Wijnands R. 1981. The function of ribosomal protein S21 in protein synthesis. *Eur J Biochem* 118:615–619. <https://doi.org/10.1111/j.1432-1033.1981.tb05563.x>
- Robertson WR, Dowsett SJ, Hardy SJS. 1977. Exchange of ribosomal proteins among the ribosomes of *Escherichia coli*. *Mol Gen Genet* 157:205–214. <https://doi.org/10.1007/BF00267399>
- Mizuno CM, Guyomar C, Roux S, Lavigne R, Rodriguez-Valera F, Sullivan MB, Gillet R, Forterre P, Krupovic M. 2019. Numerous cultivated and uncultivated viruses encode ribosomal proteins. *Nat Commun* 10:752. <https://doi.org/10.1038/s41467-019-08672-6>
- Jha V, Roy B, Jahagirdar D, McNutt ZA, Shatoff EA, Boleratz BL, Watkins DE, Bundschuh R, Basu K, Ortega J, Fredrick K. 2021. Structural basis of sequestration of the anti-Shine-Dalgarno sequence in the bacteroidetes ribosome. *Nucleic Acids Res* 49:547–567. <https://doi.org/10.1093/nar/gkaa1195>
- Chen L-X, Jaffe AL, Borges AL, Penev PI, Nelson TC, Warren LA, Banfield JF. 2022. Phage-encoded ribosomal protein S21 expression is linked to late-stage phage replication. *ISME COMMUN* 2. <https://doi.org/10.1038/s43705-022-00111-w>
- Trautmann HS, Ramsey KM. 2022. A ribosomal protein homolog governs gene expression and virulence in a bacterial pathogen. *J Bacteriol* 204:e0026822. <https://doi.org/10.1128/jb.00268-22>
- Kaledhonkar S, Fu Z, Caban K, Li W, Chen B, Sun M, Gonzalez RL, Frank J. 2019. Late steps in bacterial translation initiation visualized using time-resolved cryo-EM. *Nature* 570:400–404. <https://doi.org/10.1038/s41586-019-1249-5>
- Watson ZL, Ward FR, Méheust R, Ad O, Schepartz A, Banfield JF, Cate JH. 2020. Structure of the bacterial ribosome at 2 Å resolution. *Elife* 9:e60482. <https://doi.org/10.7554/eLife.60482>
- McNutt ZA, Roy B, Gemler BT, Shatoff EA, Moon K-M, Foster LJ, Bundschuh R, Fredrick K. 2023. Ribosomes lacking bS21 gain function to regulate protein synthesis in *Flavobacterium johnsoniae*. *Nucleic Acids Res* 51:1927–1942. <https://doi.org/10.1093/nar/gkad047>
- Bailey TL. 2021. STREME: accurate and versatile sequence motif discovery. *Bioinformatics* 37:2834–2840. <https://doi.org/10.1093/bioinformatics/btab203>
- Hall MN, Gabay J, Débarbouillé M, Schwartz M. 1982. A role for mRNA secondary structure in the control of translation initiation. *Nature* 295:616–618. <https://doi.org/10.1038/295616a0>
- de Smit MH, van Duin J. 1994. Control of translation by mRNA secondary structure in *Escherichia coli*. A quantitative analysis of literature data. *J Mol Biol* 244:144–150. <https://doi.org/10.1006/jmbi.1994.1714>
- Meibom KL, Forslund A-L, Kuoppa K, Alkhuder K, Dubail I, Dupuis M, Forsberg A, Charbit A. 2009. Hfq, a novel pleiotropic regulator of virulence-associated genes in *Francisella tularensis*. *Infect Immun* 77:1866–1880. <https://doi.org/10.1128/IAI.01496-08>
- Lenco J, Tambor V, Link M, Klimentova J, Dresler J, Peterek M, Charbit A, Stulik J. 2014. Changes in Proteome of the Δhfq strain derived from *Francisella tularensis* LVS correspond with its attenuated phenotype. *Proteomics* 14:2400–2409. <https://doi.org/10.1002/pmic.201400198>
- Park H-S, Ostberg Y, Johansson J, Wagner EGH, Uhlin BE. 2010. Novel role for a bacterial nucleoid protein in translation of mRNAs with suboptimal ribosome-binding sites. *Genes Dev* 24:1345–1350. <https://doi.org/10.1101/gad.576310>
- Wakabayashi H, Warnasooriya C, Ermolenko DN. 2020. Extending the spacing between the Shine-Dalgarno sequence and P-site codon reduces the rate of mRNA translocation. *J Mol Biol* 432:4612–4622. <https://doi.org/10.1016/j.jmb.2020.06.008>
- Postic G, Dubail I, Frapy E, Dupuis M, Dieppedale J, Charbit A, Meibom KL. 2012. Identification of a novel small RNA modulating *Francisella tularensis* pathogenicity. *PLoS One* 7:e41999. <https://doi.org/10.1371/journal.pone.0041999>
- Postic G, Frapy E, Dupuis M, Dubail I, Livny J, Charbit A, Meibom KL. 2010. Identification of small RNAs in *Francisella tularensis*. *BMC Genomics* 11:625. <https://doi.org/10.1186/1471-2164-11-625>
- Rohlfing AE, Ramsey KM, Dove SL. 2018. Polyphosphate kinase antagonizes virulence gene expression in *Francisella tularensis*. *J Bacteriol* 200:e00460-17. <https://doi.org/10.1128/JB.00460-17>
- LoVullo ED, Molins-Schneekloth CR, Schweizer HP, Pavelka MS. 2009. Single-copy chromosomal integration systems for *Francisella tularensis*. *Microbiology (Reading)* 155:1152–1163. <https://doi.org/10.1099/mic.0.022491-0>
- Charity JC, Blalock LT, Costante-Hamm MM, Kasper DL, Dove SL. 2009. Small molecule control of virulence gene expression in *Francisella tularensis*. *PLoS Pathog* 5:e1000641. <https://doi.org/10.1371/journal.ppat.1000641>
- Ramsey KM, Ledvina HE, Tresko TM, Wandzilak JM, Tower CA, Tallo T, Schramm CE, Peterson SB, Skerrett SJ, Mougous JD, Dove SL. 2020. Tn-Seq reveals hidden complexity in the utilization of host-derived

- glutathione in *Francisella tularensis*. PLoS Pathog 16:e1008566. <https://doi.org/10.1371/journal.ppat.1008566>
31. Maier TM, Havig A, Casey M, Nano FE, Frank DW, Zahrt TC. 2004. Construction and characterization of a highly efficient *Francisella* shuttle plasmid. Appl Environ Microbiol 70:7511–7519. <https://doi.org/10.1128/AEM.70.12.7511-7519.2004>
32. Ramsey KM, Osborne ML, Vvedenskaya IO, Su C, Nickels BE, Dove SL. 2015. Ubiquitous promoter-localization of essential virulence regulators in *Francisella tularensis*. PLoS Pathog 11:e1004793. <https://doi.org/10.1371/journal.ppat.1004793>
33. Chambers JR, Bender KS. 2011. The RNA chaperone Hfq is important for growth and stress tolerance in *Francisella novicida*. PLoS One 6:e19797. <https://doi.org/10.1371/journal.pone.0019797>
34. Rohlfing AE, Dove SL. 2014. Coordinate control of virulence gene expression in *Francisella tularensis* involves direct interaction between key regulators. J Bacteriol 196:3516–3526. <https://doi.org/10.1128/JB.01700-14>
35. Charity JC, Costante-Hamm MM, Balon EL, Boyd DH, Rubin EJ, Dove SL. 2007. Twin RNA polymerase-associated proteins control virulence gene expression in *Francisella tularensis*. PLoS Pathog 3:e84. <https://doi.org/10.1371/journal.ppat.0030084>
36. Sato K, Akiyama M, Sakakibara Y. 2021. RNA secondary structure prediction using deep learning with thermodynamic integration. Nat Commun 12:941. <https://doi.org/10.1038/s41467-021-21194-4>
37. Stead MB, Agrawal A, Bowden KE, Nasir R, Mohanty BK, Meagher RB, Kushner SR. 2012. RNAsnap: A rapid, quantitative and inexpensive, method for isolating total RNA from bacteria. Nucleic Acids Res 40:e156. <https://doi.org/10.1093/nar/gks680>