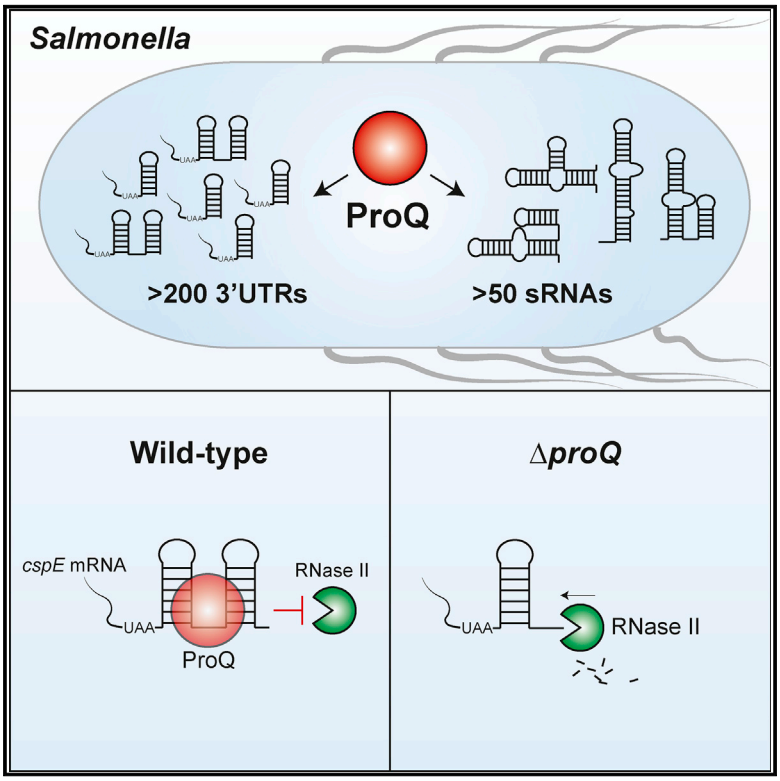
Resource

Molecular Cell

Global Maps of ProQ Binding *In Vivo* Reveal Target Recognition via RNA Structure and Stability Control at mRNA 3′ Ends

Graphical Abstract



Highlights

- CLIP-seq maps hundreds of ProQ binding sites in *Salmonella* and *E. coli*
- ProQ binding is driven by RNA structure
- Small RNAs and mRNA 3′ UTRs are enriched among ProQ ligands
- ProQ prevents RNA decay by counteracting exoribonuclease activity

Authors

Erik Holmqvist, Lei Li, Thorsten Bischler, Lars Barquist, Jörg Vogel

Correspondence

joerg.vogel@uni-wuerzburg.de

In Brief

Using CLIP-seq, Holmqvist et al. map transcriptome-wide interactions of the emerging global RNA-binding protein ProQ in *Salmonella* and *E. coli*. Their data suggest ProQ to target sRNAs and mRNA 3′ UTRs primarily through a structural code and to stabilize some mRNAs by counteracting 3′ exoribonuclease activity.

Holmqvist et al., 2018, Molecular Cell *70*, 971–982
June 7, 2018
https://doi.org/10.1016/j.molcel.2018.04.017

CellPress

Molecular Cell
Resource

Global Maps of ProQ Binding *In Vivo* Reveal Target Recognition via RNA Structure and Stability Control at mRNA 3′ Ends

Erik Holmqvist,[1,2] Lei Li,[1] Thorsten Bischler,[1] Lars Barquist,[1,3] and Jörg Vogel[1,3,4,*]
[1]Institute for Molecular Infection Biology, University of Würzburg, 97080 Würzburg, Germany
[2]Department of Cell and Molecular Biology, Biomedical Center, Uppsala University, 75124 Uppsala, Sweden
[3]Helmholtz Institute for RNA-based Infection Research (HIRI), 97080 Würzburg, Germany
[4]Lead Contact
*Correspondence: joerg.vogel@uni-wuerzburg.de
https://doi.org/10.1016/j.molcel.2018.04.017

SUMMARY

The conserved RNA-binding protein ProQ has emerged as the centerpiece of a previously unknown third large network of post-transcriptional control in enterobacteria. Here, we have used *in vivo* UV cross-linking and RNA sequencing (CLIP-seq) to map hundreds of ProQ binding sites in *Salmonella enterica* and *Escherichia coli*. Our analysis of these binding sites, many of which are conserved, suggests that ProQ recognizes its cellular targets through RNA structural motifs found in small RNAs (sRNAs) and at the 3′ end of mRNAs. Using the *cspE* mRNA as a model for 3′ end targeting, we reveal a function for ProQ in protecting mRNA against exoribonucleolytic activity. Taken together, our results underpin the notion that ProQ governs a post-transcriptional network distinct from those of the well-characterized sRNA-binding proteins, CsrA and Hfq, and suggest a previously unrecognized, sRNA-independent role of ProQ in stabilizing mRNAs.

INTRODUCTION

Globally acting RNA-binding proteins (RBPs) that function in conjunction with small noncoding RNAs (sRNAs) are key components of post-transcriptional control networks in diverse organisms (Gorski et al., 2017; Wagner and Romby, 2015). In bacteria, the only two known examples of such global RBPs until recently were CsrA and Hfq. CsrA-like proteins generally function as post-transcriptional repressors by binding to AUGGA sequences in mRNA 5′ UTRs; their activity is typically controlled by antagonistic sRNAs that sequester them by the presentation of high-affinity binding sites (Vakulskas et al., 2015). Hfq also mediates post-transcriptional regulation in mRNA 5′ regions, though primarily by facilitating the activities of base-pairing sRNAs; similarly to eukaryotic microRNAs, these Hfq-associated sRNAs typically target multiple mRNAs (De Lay et al., 2013; Updegrove et al., 2016; Vogel and Luisi, 2011). In *Escherichia coli* and *Salmonella*, two enteric bacteria in which the targetomes of CsrA and Hfq have been extensively mapped, these proteins bind at least one-third of all expressed transcripts and broadly impact bacterial physiology and virulence (Chao et al., 2012; Feliciano et al., 2016; Holmqvist et al., 2016; Hör and Vogel, 2017; Potts et al., 2017; Sowa et al., 2017; Zhang et al., 2003).

The focus on CsrA and Hfq notwithstanding, evidence has accumulated that other global RBPs must be present in bacteria. For example, extensive mapping efforts concluded that more than half of the known *Salmonella* sRNAs may not be associated with CsrA or Hfq *in vivo* (Chao et al., 2012; Holmqvist et al., 2016). Moreover, many other bacteria use sRNA-mediated post-transcriptional control but do not contain functional homologs of these proteins (Durand et al., 2015). Proving this conjecture correct, the protein ProQ has recently been discovered as the third sRNA-related RBP with a prominent function in post-transcriptional control in bacteria (Attaiech et al., 2016; Olejniczak and Storz, 2017; Smirnov et al., 2016). Based on co-immunoprecipitation (coIP) data, ProQ may form complexes with dozens of sRNAs and hundreds of mRNAs in *Salmonella* (Smirnov et al., 2016). One of these sRNAs, RaiZ, was subsequently shown to repress translation of the *trans*-encoded *hupA* mRNA in a ProQ-dependent manner (Smirnov et al., 2017). A similar protein, RocC, was found to be essential for sRNA-mediated mRNA repression of competence genes in *Legionella pneumophila* (Attaiech et al., 2016).

ProQ is an ~25-kDa protein that was originally identified as being important for osmolyte accumulation in *E. coli* by increasing cellular levels of the proline transporter ProP (Kunte et al., 1999; Milner and Wood, 1989). Its RNA-binding activity was hypothesized based upon sequence similarity with FinO, a plasmid-encoded RBP required for *cis*-antisense RNA-mediated control of conjugation (Smith et al., 2004). ProQ/FinO-domain proteins are now known to be widespread in α-, β-, and γ-proteobacteria, encoded by chromosomes and mobile elements, such as phages and plasmids (Attaiech et al., 2017; Glover et al., 2015; Olejniczak and Storz, 2017; Smirnov et al., 2016). Pioneering biochemical work by the Wood and Glover groups showed that ProQ not only has high affinity to an RNA substrate derived from FinP (the antisense RNA target of FinO) but also possesses RNA strand exchange and RNA annealing activities (Chaulk et al., 2011). According to the available biochemical and

structural data, the N-terminal FinO domain of ProQ is the primary determinant of its RNA binding capacity, though both the linker and the C-terminal domain may also contribute (Chaulk et al., 2011; Gonzalez et al., 2017). Given the growing interest in ProQ, a high-resolution map of the protein's interactions with transcripts *in vivo* is needed to inform mechanistic studies, define major binding determinants, and guide the synthesis of RNA ligands for co-crystallization efforts.

UV crosslinking and immunoprecipitation followed by RNA sequencing (CLIP-seq) has become a powerful tool to globally profile RBP binding sites in cellular transcripts. In comparison to standard coIP analysis, CLIP-seq offers RNA-protein recovery under denaturing conditions and also often reveals binding sites through mutations at the crosslinked nucleotide (reviewed in König et al., 2012). CLIP-seq has recently been used with great success to provide RNA sequence recognition motifs for CsrA and Hfq in several enteric bacteria (Holmqvist et al., 2016; Potts et al., 2017; Tree et al., 2014). In the present study, we applied CLIP-seq to map in the *Salmonella* and *E. coli* transcriptomes hundreds of ProQ binding sites, most of which are located in highly structured sRNAs and mRNA 3′ UTRs. Our results suggest that, unlike CsrA and Hfq, which select RNA clients through distinct short primary sequences, ProQ recognizes its cellular targets primarily on the level of RNA structure. Whereas post-transcriptional control of bacterial mRNAs typically takes place near their 5′ ends, we reveal predominant interactions of ProQ with mRNA 3′ regions and report on a previously unknown function of this RBP in protecting mRNAs from 3′ exonucleolytic decay.

RESULTS

Identification of ProQ Binding Sites by *In Vivo* UV Crosslinking

To map cellular ProQ-RNA interactions at high resolution, we applied our recently established protocol for bacterial CLIP-seq (Holmqvist et al., 2016) to *Salmonella* that express ProQ with a C-terminal triple FLAG tag (Smirnov et al., 2016). To facilitate comparison with the previous CLIP-seq data for CsrA and Hfq (Holmqvist et al., 2016), the bacteria were analyzed in the same growth phase (early stationary phase) as before. Strong radioactive signals after labeling of co-immunoprecipitated and ProQ-associated RNA were dependent on prior UV treatment (which induces covalent RNA-protein bridges), indicating that unspecifically bound transcripts were successfully depleted (Figure 1A). Purified RNA from three independent experiments was converted into cDNA and subjected to Illumina sequencing. For background correction (Friedersdorf and Keene, 2014), we sequenced cDNA libraries prepared from both crosslinked and non-crosslinked cultures. The triplicates showed clear read enrichments in crosslinked samples as compared to matched background controls (Figure S1A).

ProQ binding sites were predicted with a customized peak calling algorithm (Holmqvist et al., 2016). Based on three biological replicates, we identified 467 ProQ peaks with statistically significant enrichment ($q \leq 0.01$) in the crosslinked samples (Table S1). These peaks came from 499 annotated transcripts expressed from all parts of the *Salmonella* genome (Figure 1B; Table S1); 47 of them mapped to 5′ UTRs, 247 to coding sequences (CDSs), 233 to 3′ UTRs, 70 to sRNAs, 3 to tRNAs, and 26 to two of the three *Salmonella* SL1344 plasmids, whereas 27 peaks mapped to orphan regions in the chromosome (Figure 1C). Within each RNA class, 3′ UTRs and sRNAs had the highest percentage of peaks (Figure S1B). The number of peaks mapping to features exceeds the actual number peaks (626 versus 467, respectively) because many peaks map to overlapping (e.g., 3′ UTRs and sRNAs) or adjacent (e.g., CDSs and 3′ UTRs) annotations (Table S1).

The general distribution of ProQ peaks along all *Salmonella* mRNAs was approached by meta-gene analysis with start or stop codons as reference points. This analysis showed a strong peak enrichment around stop codons (Figure 1D). The peak enrichment close to stop codons could reflect either a propensity of ProQ to bind within 3′ UTRs near the stop codon or a mechanistic link between ProQ binding and a stop-codon-dependent process, such as translation termination. To discriminate between these two possibilities, meta-gene analysis was applied to all *Salmonella* open reading frames (ORFs) after their separation into two bins: ORFs that are followed by a 3′ UTR versus ORFs that are followed by another ORF in an operon. Strikingly, almost all peaks located close to stop codons fell in the first category (terminal 3′ UTRs), suggesting that ProQ is attracted to an mRNA 3′ end feature rather than a terminating ribosome (Figure 1D).

According to a recent biochemical classification by Grad-seq, most of the potential target sRNAs of ProQ may not associate with either CsrA or Hfq (Smirnov et al., 2016). Cross-comparison of 138 sRNAs captured by CLIP-seq in both the present ProQ and the previous CsrA and Hfq analyses (Holmqvist et al., 2016) corroborates the view of three distinct sRNA classes; Hfq and ProQ sRNAs, which dominate by their individual numbers, clearly segregate concerning crosslinking to either of these two RBPs (Figures 1E and S2). Nonetheless, some sRNAs crosslinked to two or even all three RBPs. Overall, our analysis indicates that ProQ recognizes specific features embedded in a group of sRNAs, most of which are not bound by Hfq and CsrA.

Absence of a Common Sequence Motif from ProQ Binding Sites

RBPs bind to RNA motifs consisting of specific sequences, structures, or a combination of both, embedded in their ligands (Lunde et al., 2007). Crosslink-specific mutations in CLIP-seq reads indicate RNA-protein contacts at single-nucleotide resolution, which can be exploited to compute RBP binding motifs (Zhang and Darnell, 2011). In our data, 29% (127/441) of the ProQ peaks mapping to the *Salmonella* chromosome contained cDNA mutations (see STAR Methods), most of which were C→T transitions (Figure 2A). However, contrasting with our previous success in defining Hfq and CsrA recognition motifs from mutation-proximal regions (Holmqvist et al., 2016) using the MEME algorithm (Bailey et al., 2015), we failed to identify a common sequence motif for ProQ (Figure S3A). By building a GraphProt model on the full set of ProQ peaks, a sequence motif could be generated for the 50 top scoring sequences under the model (Figure S3B). The strong guanine-cytosine (GC) enrichment in this motif suggested that ProQ binding sites are structured. Indeed, upon assessing

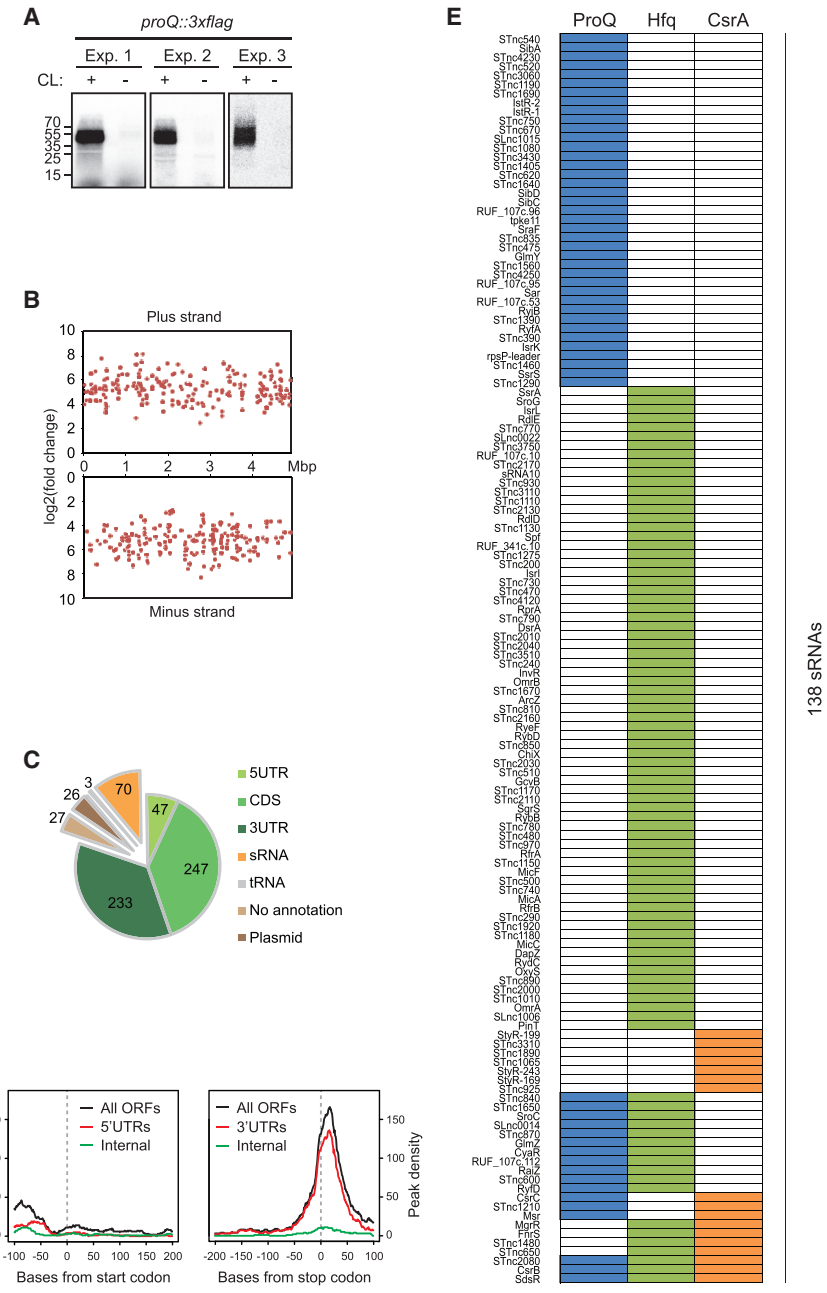


Figure 1. CLIP-Seq Reveals Hundreds of ProQ Binding Sites on the *Salmonella* Transcriptome

(A) Autoradiographs of radioactively labeled RNA fragments covalently bound to ProQ after *in vivo* UV crosslinking, immunoprecipitation, gel electrophoresis, and membrane transfer. Each caption is from an independent experiment. CL, UV crosslinking; Exp, experiment.

(B) Fold change versus genomic position of significantly enriched ($q \leq 0.01$) ProQ peaks.

(C) Distribution of significant ProQ peaks among indicated RNA classes.

(D) Meta-gene analysis of ProQ peaks using mRNA start or stop codons as reference points. For 5′- and 3′-UTR-specific distributions, start and stop codons of the first and last cistrons, respectively, in polycistronic mRNAs were used.

(E) All *Salmonella* sRNAs detected by CLIP-seq of ProQ, Hfq, and/or CsrA (this study and Holmqvist et al., 2016). A colored box indicates crosslinking of the indicated protein.

the propensity for forming RNA structure at ProQ binding sites (comparing the predicted folding energy of sequences surrounding crosslink sites to sequence-scrambled counterparts), sequences immediately 3′ of crosslink sites displayed significantly more structure than those of the background model (Figure 2B). Application of the co-variance-based algorithm CMfinder (Yao et al., 2006) to mutation-proximal sequences predicted a motif consisting of a hairpin structure with the highest confidence for base pairing at the base of the stem (Figure 2C). Similar to MEME, CMfinder did not detect a clear sequence motif, suggesting that ProQ tends to recognize its cellular targets by RNA structure rather than primary sequence.

Figure 2D shows ProQ peaks and crosslinking sites mapped onto the predicted secondary structures of several top-enriched sRNAs. Echoing the above peak characteristics, the ProQ crosslink sites are located upstream or within strong secondary structures. For example, the RaiZ sRNA base-pairs with and regulates translation of the *hupA* mRNA in *trans* in a ProQ-dependent manner (Smirnov et al., 2017). A single CLIP-seq peak indicates that ProQ contacts two stem loops in the 3′ part of RaiZ, in good agreement with *in vitro* foot printing of ProQ binding on this sRNA (Smirnov et al., 2017). SibC is a *cis*-encoded antitoxin of a type 1 toxin-antitoxin system (Fozo et al., 2008), whereas RyfD and STnc1690 are of unknown function. Two of these sRNAs, RaiZ and SibC, are known to require ProQ for their intracellular stability (Smirnov et al., 2016, 2017). Comparison between the locations of ProQ peaks and available RNase E data for *Salmonella* (Chao et al., 2017) indicates that ProQ binding overlaps with several RNase E cleavage sites in each of these sRNAs. Globally, 162 of the 441 ProQ peaks mapped to the *Salmonella* chromosome overlap with RNase E cleavage sites, a significant ($p \leq 2.2e{-}16$) enrichment compared to peaks of the same size and number randomly distributed on the *Salmonella* transcriptome (see STAR Methods). It is reasonable to speculate that ProQ blocks access of RNase E to some of these sRNAs.

With the exception of RaiZ, the sRNAs in Figure 2D are predicted to form stable secondary structures. In contrast, most Hfq-dependent sRNAs exhibit extensive single-stranded regions

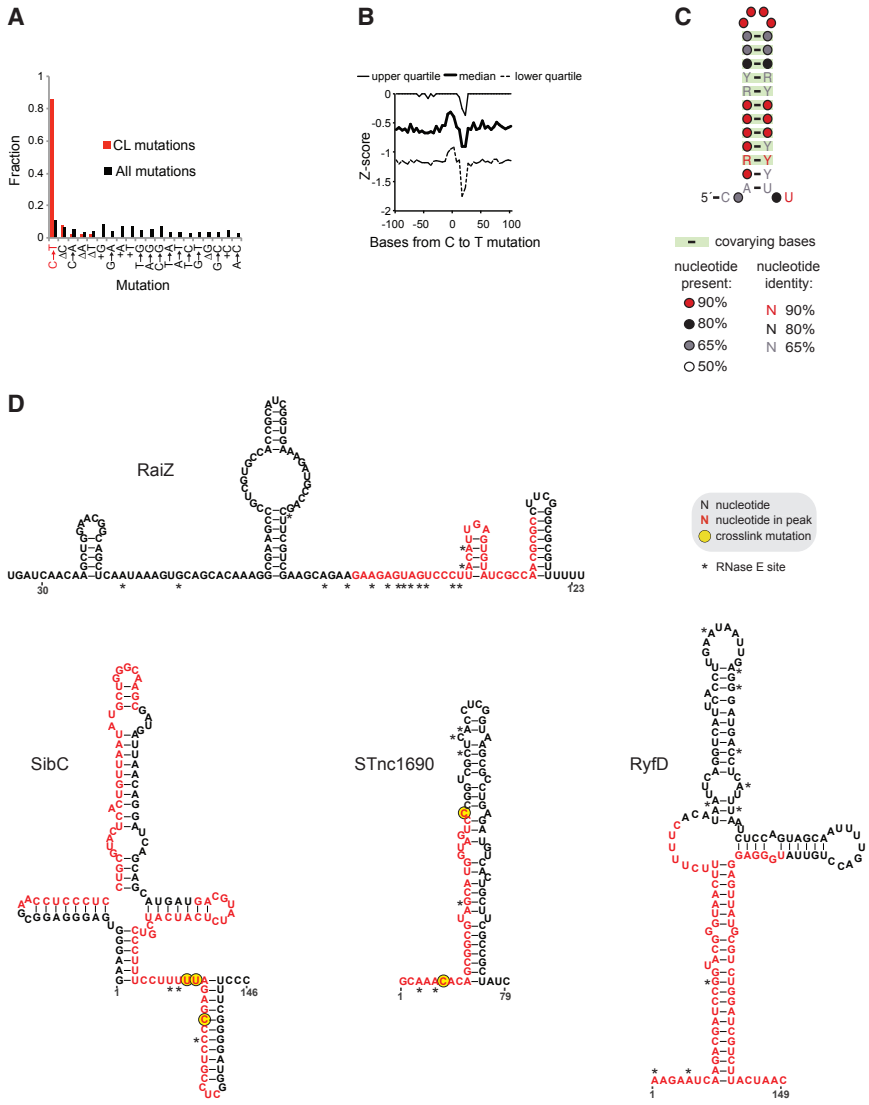


Figure 2. ProQ Binds to Highly Structured RNA

(A) Relative abundance of crosslink-specific read mutations.

(B) *In silico* RNA folding energy of sequences flanking crosslink-specific C to T read mutations. *Z* scores were calculated by comparing energy values for actual sequences to those of 1,000 shuffled sequences with the same dinucleotide content.

(C) The CMfinder algorithm was applied to all sequences around crosslink-specific C to T mutations. The highest scoring motif is shown.

(D) Predicted secondary structures of the indicated sRNAs. Nucleotides corresponding to a ProQ peak and positions of crosslink-induced mutations are colored in red and yellow, respectively.

to enable binding to both Hfq and mRNA targets. In particular, Hfq-bound sRNAs carry Rho-independent terminators ending in long single-stranded poly(U) tails, to which Hfq binds with high affinity (Ishikawa et al., 2012; Sauer and Weichenrieder, 2011). ProQ peaks also map close to Rho-independent terminators (Figure S3C), but only a small number of sRNAs crosslink to both ProQ and Hfq (Figure 1E), suggesting the existence of two distinct types of terminators recognized by either of these two RBPs. Indeed, a global analysis gives strengthened support for sRNA terminators bound by ProQ having short and those bound by Hfq carrying long single-stranded U tails (Figures S3D–S3F). Thus, terminator characteristics may be one feature that drives sRNA association with different RBPs and may at least partly explain the dichotomy in sRNA association with Hfq or ProQ.

Comparative CLIP-Seq Shows Similar ProQ Binding in *Salmonella* and *E. coli*

Evolutionary conservation of macromolecular interactions is a strong indicator of biological function. For a comparison to the ProQ-RNA interactions in *Salmonella*, we used CLIP-seq to identify ProQ sites in *E. coli* W3110, again with a chromosomal *proQ::3xflag* allele (Table S2). Mirroring the *Salmonella* results (Figure 1D), meta-gene analysis of *E. coli* mRNAs showed a strong enrichment of ProQ peaks in mRNA 3′ regions (Figure 3A).

To investigate a potential sequence conservation of ProQ sites, we compared ProQ peaks mapping to genomic regions present in both these species. The global similarity of peak positioning was then determined by, for each *Salmonella* peak, calculating the distance to the closest *E. coli* peak. Strikingly, many peaks fell within a 10-nt distance, indicating a high degree of conservation of ProQ binding positions between *E. coli* and *Salmonella* (Figure 3B).

RNA sequence per se seems to be a poor indicator of ProQ binding (Figures 2C and S1B), suggesting that binding instead might rely upon conserved RNA structure. To test this, we searched for shared ProQ sites with no obvious sequence similarity in the two species. A striking example is the *fabB* mRNA, encoding an enzyme involved in fatty acid biosynthesis. Both species showed a ProQ peak in the *fabB* 3′ UTR (Figure 3C) despite little if any shared nucleotide sequence (Figure 3D, top). Nonetheless, the peak regions are predicted to adopt a similar structure, with a highly similar pattern of ProQ coverage between *E. coli* and *Salmonella* (Figure 3D, bottom). This further strengthens the notion that ProQ binding is largely driven by structure rather than sequence. Taken together, this global comparative CLIP-seq analysis indicates conserved functions of ProQ at the level of individual transcripts.

ProQ Positively Regulates mRNA Abundance and Stability

The pervasive binding to mRNA 3′ regions (Figures 1C and 1D) suggested that ProQ may influence the expression of these

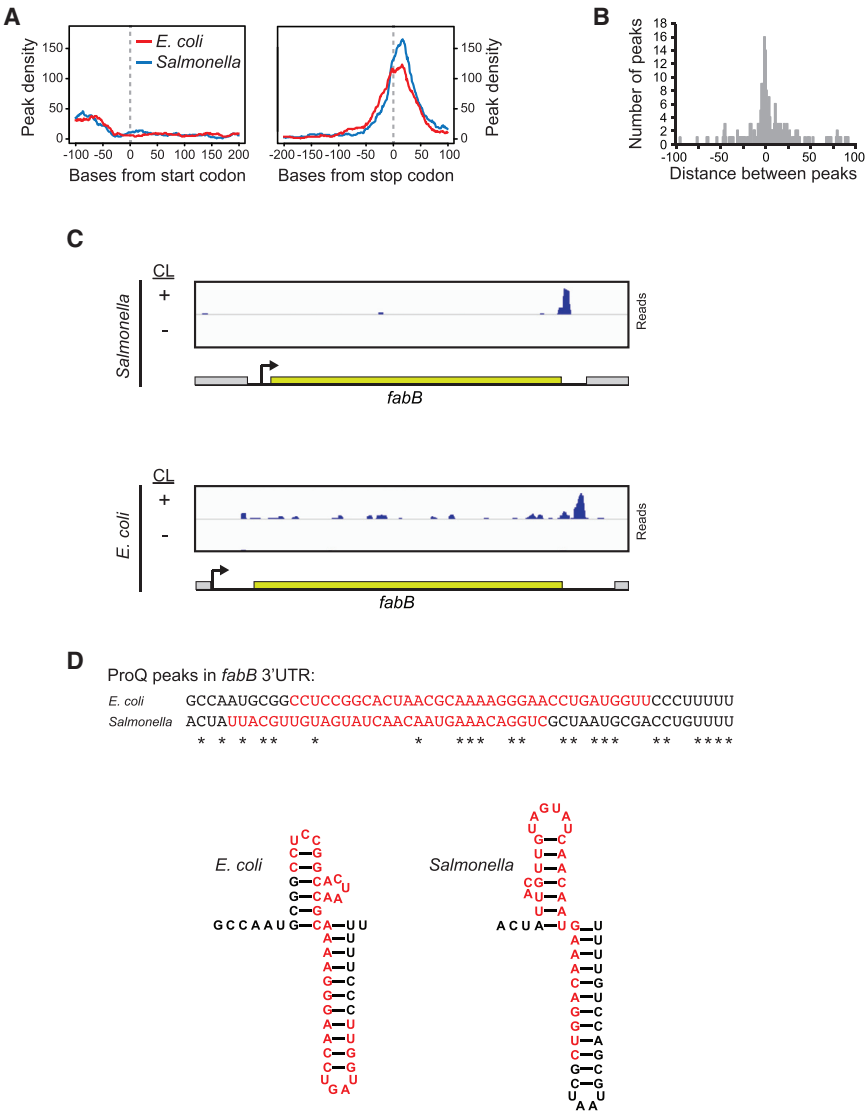


Figure 3. ProQ Binding Is Conserved between *Salmonella* and *E. coli*

(A) Meta-gene analysis of ProQ peak distributions on all *Salmonella* or *E. coli* mRNAs.

(B) Distance between *E. coli* and *Salmonella* peaks mapped to a SuperGenome based on the genomes of *Salmonella enterica serovar* Typhimurium strain SL1344 and *E. coli* K-12 strain W3110. Shown are peaks that mapped within 100 nt from each other.

(C) CLIP-seq read distributions at the *fabB* locus in *E.coli* and *Salmonella*.

(D) Sequence alignment of the ProQ peak region in the *fabB* 3′ UTR (top). Predicted secondary structures of the *fabB* 3′ UTRs from *E. coli* and *Salmonella* (bottom) are shown. Nucleotides corresponding to a ProQ peak are colored in red.

transcripts on the post-transcriptional level, similarly to its previously described effects on many sRNAs (Smirnov et al., 2016). To measure a possible effect of ProQ on its mRNA ligands, the steady-state levels of four well-expressed mRNAs (*cspE*, *ompD*, *ompF*, and *fliC*), all of which had one or several 3′-located ProQ peaks, were analyzed by northern blot (Figure 4A). *Salmonella* wild-type and Δ*proQ* strains harboring a control plasmid, as well as the Δ*proQ* strain harboring a plasmid that slightly overexpresses ProQ (*proQ*+), were grown to early stationary phase. Each of these mRNAs showed reduced levels in the absence of ProQ (Δ*proQ*), which were fully restored to wild-type levels in the *proQ*+ complementation strain. Thus, ProQ binding might be a positive determinant for mRNA abundance.

For further analysis, we chose the *cspE* mRNA as model because it is short (300 nt) and abundantly recovered by coIP with ProQ (Figure S4; Smirnov et al., 2016). Reduced levels in the Δ*proQ* strain could be due to a reduction in transcription rate or RNA stability. However, effects on *cspE* transcription in the Δ*proQ* strain are unlikely because lower mRNA levels were also seen when *cspE* was transcribed from a heterologous P_{BAD} promoter (Figure 4B). To test whether ProQ promotes *cspE* mRNA stability, a rifampicin run-out experiment was conducted. Indeed, whereas the *cspE* mRNA decays with a half-life of ~4 min in wild-type *Salmonella*, its half-life drops to 1 min in the Δ*proQ* strain; *proQ* complementation restores it to 4 min, as in the wild-type strain (Figures 4C and 4D).

To test whether ProQ-dependent mRNA stabilization reflects a more general phenomenon, we monitored the stability of several other mRNAs (*cspC*, *cspD*, *ompD*, and *yfiD*) with ProQ binding sites in their 3′ regions (Table S1) on northern blots. Three out of four of these mRNAs were destabilized in the absence of ProQ, whereas complementation of ProQ from a plasmid restored or even increased stability (Figure S5). Thus, binding of ProQ to transcribed 3′ UTRs may be a general mechanism to stabilize mRNAs.

The 3′ End of *cspE* Is Sufficient for ProQ Binding and mRNA Stabilization

ProQ-dependent stabilization of *cspE* mRNA combined with a ProQ site in the 3′ UTR suggested that this protein controls *cspE* stability by acting at the 3′ end. The *cspE* 3′ UTR is predicted to fold into two stem loops, the second of which is a Rho-independent terminator (Figure 5A). The CLIP-seq results located ProQ binding across parts of both stem loops, including the single-stranded linker (Figure 5A). This site is highly conserved, as judged by almost identical crosslink-specific enrichments obtained in *Salmonella* and *E. coli* (Figure 5B).

To dissect how different parts of the *cspE* mRNA contribute to ProQ-dependent regulation, stepwise 5′ truncated mRNAs were expressed from a heterologous promoter. Northern blot analysis

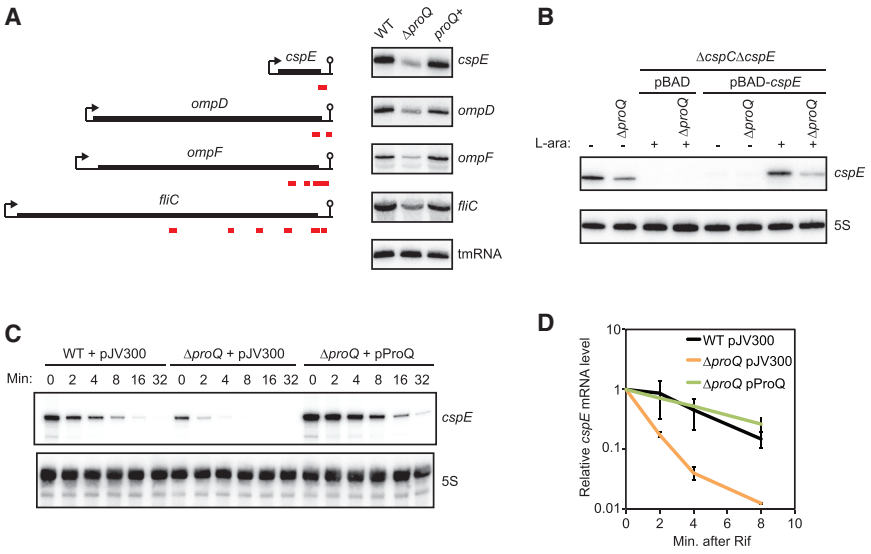


Figure 4. ProQ Stabilizes Its mRNA Targets

(A) Northern blot analysis of the steady-state levels of indicated mRNAs with respect to the absence or presence of ProQ in *Salmonella* grown to early stationary phase (OD_{600} of 2.0). *proQ*+ refers to a strain expressing ProQ from a plasmid in a *proQ* deletion background. Transfer-messenger RNA (tmRNA) served as loading control.

(B) Northern blot analysis of endogenous and plasmid-expressed *cspE* mRNA. Wild-type (WT) and Δ*proQ* strains with indicated plasmids were grown to early stationary phase (OD_{600} of 2.0) and induced for pBAD expression for 30 min. 5S served as loading control.

(C) Northern blot analysis of the *cspE* mRNA before and after addition of rifampicin to inhibit transcription. The indicated strains were grown to early stationary phase (OD_{600} of 2.0) before addition of rifampicin. 5S served as loading control.

(D) Quantification of *cspE* mRNA abundance measured by northern blot after addition of rifampicin. Signals for *cspE* mRNA were normalized to 5S signals. Means and error bars are based on three independent experiments.

showed that all *cspE* mRNA variants, the shortest one comprising only the last 77 nt (*cspE*+234), required ProQ for stability (Figure 5C). Moreover, because most of these truncated *cspE* transcripts lack signals for translation initiation, the observed ProQ-dependent stabilization must be independent of translation. To verify ProQ binding to the 3′ region of *cspE* mRNA, we performed *in vitro* binding assays with the *cspE*+234 RNA and purified ProQ. Titration of ProQ shifted radiolabeled *cspE*+234 in a concentration-dependent manner (Figure 5D), consistent with a recent study (Gonzalez et al., 2017). The analysis of peak characteristics suggested that ProQ preferentially binds structured RNA (Figures 2B and 2C). To test this experimentally, we conducted binding assays between ProQ and a *cspE*+234 variant ("*cspE* mut"; Figure 5D) with three point mutations that destabilize the terminator stem. Whereas ProQ binding was clearly affected by the structure-destabilizing mutations, re-stabilizing the terminator stem by introducing compensatory mutations ("*cspE* comp") fully restored binding (Figure 5D). The interaction between ProQ and *cspE*+234 was specific, because an excess of unlabeled *cspE*+234, but not an unspecific competitor RNA, effectively competed with labeled *cspE*+234 for ProQ binding (Figure 5E).

ProQ Antagonizes Exoribonucleolytic Activity at the 3′ UTR of *cspE* mRNA

The ProQ-dependent stabilization of *cspE* mRNA suggested that ProQ—directly or indirectly—counteracted a ribonucleolytic activity. Because the ProQ binding site is located at the 3′ end of the *cspE* mRNA (Figure 5A), and stabilization is independent of upstream sequence (Figure 5C), we reasoned that ProQ would interfere with a 3′-dependent exoribonuclease (Hui et al., 2014).

Salmonella possess three main 3′–5′ exoribonucleases: RNase R; PNPase; and RNase II, encoded by *rnr*, *pnp*, and *rnb*, respectively (Andrade et al., 2009). To test their involvement in ProQ-dependent regulation, the *cspE* mRNA was probed in *Salmonella* strains deleted for each of these exoribonucleases, alone or in combination with the Δ*proQ* allele. Whereas the absence of PNPase and RNase R did not affect *cspE* mRNA levels, inactivation of RNase II rendered the *cspE* mRNA insensitive to the *proQ* deletion (Figure 6A). Furthermore, the Δ*rnb* Δ*proQ* strain displayed the same half-life of *cspE* mRNA, as did wild-type *Salmonella* (Figure 6B), which argues that RNase II is the main 3′ decay enzyme ProQ protects against.

DISCUSSION

In this paper, we have used CLIP-seq to capture RNA ligands associated with the emerging global RBP ProQ in *E. coli* and *Salmonella*. This approach detected hundreds of RNA target sites in intact cells, providing the first direct evidence for ProQ interactions across the transcriptome. Our results support the existence of a third global regulon of sRNA-mediated gene regulation in bacteria distinct from the well-characterized CsrA and Hfq regulons. Based on these results, we speculate that major mechanisms of ProQ-mediated regulation may involve mRNA 3′ ends, which we demonstrated using the *cspE* mRNA as a model. Our investigation of the traits of ProQ binding sites raises the exciting possibility of a global regulatory network that relies upon an RNA structure code rather than the recognition of primary sequence. These findings have additional important implications for the broader family of FinO domain-containing RBPs, which are widely distributed among the proteobacteria.

Comparison with CsrA and Hfq

The 467 ProQ binding sites defined here are quantitatively comparable to the 640 and 467 binding sites of Hfq and CsrA, respectively (Holmqvist et al., 2016), identified under the same conditions in the same *Salmonella* strain. As did CsrA and Hfq (Holmqvist et al., 2016), ProQ crosslinked primarily to sRNAs and mRNAs (Figure 1), despite the much higher cellular copy numbers of rRNAs or tRNAs. In other words, ProQ is a globally

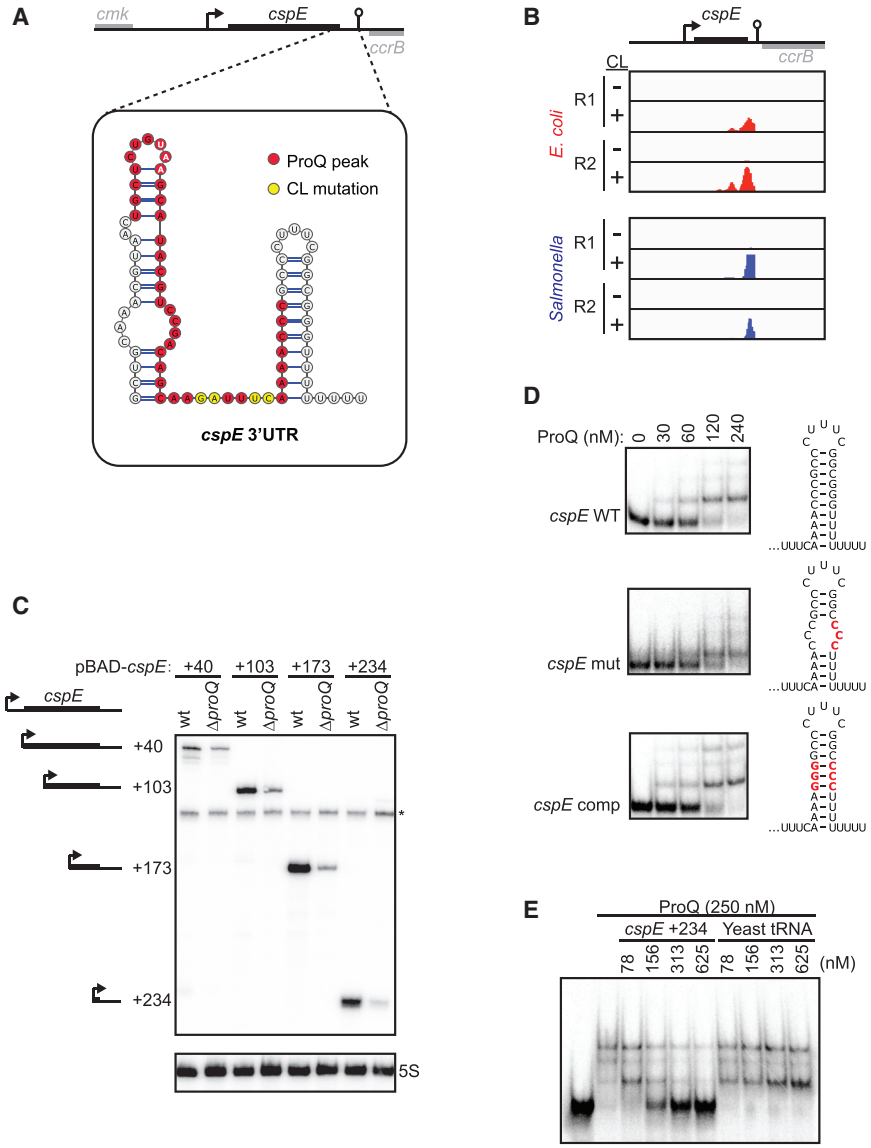


Figure 5. ProQ Binds to the 3′ UTR of *cspE* to Protect against mRNA Degradation

(A) Predicted secondary structure of the *cspE* 3′ UTR in *Salmonella*. The structure prediction was generated with RNAfold (Lorenz et al., 2011) and visualized using VARNA (Darty et al., 2009). The ProQ peak sequence and crosslink-specific mutations are indicated. The *cspE* stop codon is highlighted in white letters.

(B) Read coverage at the *cspE* locus from ProQ CLIP-seq experiments in *Salmonella* and *E. coli*, respectively.

(C) Northern blot analysis of *cspE* mRNA variants truncated from the 5′ end. *Salmonella* harboring the indicated plasmids were grown to early stationary phase (OD_{600} of 2.0) and induced for pBAD expression for 30 min. 5S served as loading control. The signal indicated with an asterisk is of unknown origin but likely reflects readthrough from the flippase recognition target (FRT) scar replacing the *cspE* gene in the Δ*cspE*Δ*cspC* strain background.

(D) Migration of *in vitro* transcribed and radioactively labeled *cspE*+234 RNA and variants thereof in a non-denaturing gel after 5 min of incubation with purified ProQ at increasing concentrations.

(E) Migration of *in vitro* transcribed and radioactively labeled *cspE*+234 RNA in a non-denaturing gel after incubation with ProQ and increasing concentrations of non-labeled *cspE*+234 or yeast tRNA.

acting RBP with a defined target suite that does not indiscriminately bind to just any abundant cellular RNA.

Comparison of the sRNA ligands of ProQ with those of CsrA and Hfq corroborates the concept of a distinct class of ProQ-associated sRNAs (Figure 1E). Associations with two or even all three RBPs were observed, but whether these overlaps imply a functional dependence of a crosslinked sRNA on each of the RBPs in question is yet unclear. That is, even though CLIP-seq is sensitive, it is only semiquantitative and may also detect interactions that are of no functional consequence. For example, the RaiZ sRNA was originally identified in Hfq coIP studies (Chao et al., 2012), but later studies showed RaiZ to depend solely on ProQ for both its intracellular stability and its ability to repress the translation of its main target, the *hupA* mRNA (Smirnov et al., 2017). We also note that CsrB and SdsR, both of which have been crosslinked to all three RBPs (Figure 1E), belong to the most abundant sRNAs in *E. coli* and *Salmonella*. Nonetheless, it is exciting to speculate that at least some of these multi-protein crosslinks reflect functional sorting of sRNAs into specialized complexes. As an example, in *E. coli*, the McaS sRNA employs two different mechanisms and acts through both Hfq and CsrA to regulate the formation of bacterial biofilms (Jørgensen et al., 2013).

A few multiple associations notwithstanding, the CLIP-seq data reinforce the idea of a distinct class of ProQ-associated sRNAs. What are their main characteristics? In contrast to Hfq-dependent sRNAs, which generally target mRNAs in *trans*, ProQ crosslinks to many *cis*-encoded sRNAs, including antisense RNAs of type 1 toxin-antitoxin systems, such as *sib-ibs* and *istR-tisB* (Fozo et al., 2008; Vogel et al., 2004). There is so far only one characterized sRNA—RaiZ—which similarly to Hfq-associated sRNAs base pairs with the 5′ UTR of a *trans*-encoded mRNA (Smirnov et al., 2017). It is intriguing that fewer 5′ UTR crosslinks were observed with ProQ than with Hfq (compare Figure 1C with Figure 2C in Holmqvist et al., 2016). This indicates that multiple mRNA targeting via 5′ UTRs by ProQ-associated sRNAs may be less common than in the Hfq regulon. As caveat of this prediction, we observed no ProQ crosslinks in the 5′ UTR of *hupA* mRNA (Table S1), despite it being a validated *trans*-encoded target of RaiZ.

With respect to the sRNAs, an interesting pattern emerges at their 3′ ends. Comparing their terminators, we find that

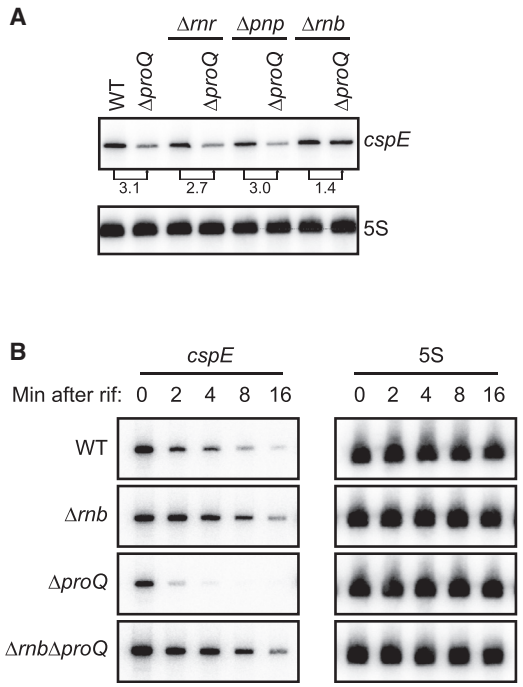


Figure 6. RNase II Degrades the *cspE* mRNA in the Absence of ProQ

(A) Northern blot analysis of *cspE* mRNA in strains with indicated genotypes. Each strain was grown to early stationary phase (OD_{600} of 2.0). 5S served as loading control.

(B) Northern blot analysis of the *cspE* mRNA before and after addition of rifampicin to inhibit transcription. The indicated strains were grown to early stationary phase (OD_{600} of 2.0) before addition of rifampicin. 5S served as loading control.

ProQ-bound sRNAs tend to have few single-stranded terminal uridines, in contrast to longer 3′ U tails in Hfq-bound sRNAs (Figure S3C). A long poly(U) tail is important for Hfq binding, both *in vitro* and *in vivo* (Otaka et al., 2011; Sauer and Weichenrieder, 2011), which may explain why Hfq fails to bind most ProQ-specific sRNAs. However, it is less clear why ProQ then does not interact with the majority of Hfq-bound sRNAs, even though ProQ is much more abundant than Hfq (Smirnov et al., 2016). Either a long poly(U) tail is a negative determinant for ProQ binding or the higher affinity for terminators with longer U tails helps Hfq to outcompete ProQ at transcript 3′ ends, as discussed recently (Olejniczak and Storz, 2017). In this regard, the now available CLIP-seq data for both Hfq and ProQ should suggest promising candidate RNAs for mechanistic studies of the RBP competition at the 3′ ends of sRNAs. Further, as the vast majority of ProQ-binding sRNAs remain of unknown function, it may be worth considering that some of them act as sponges of ProQ (Bossi and Figueroa-Bossi, 2016).

Control at the mRNA 3′ End

In contrast to eukaryotes, in which most post-transcriptional control occurs within the 3′ UTRs of transcripts, 5′ UTRs are considered the business ends, where bacterial global RBPs exert control of mRNA stability and translation (Van Assche et al., 2015). This is where CsrA acts primarily—by steric inhibition of ribosome binding (Romeo et al., 2013)—and where most Hfq-associated RNAs or even Hfq alone repress or activate mRNAs (Chen and Gottesman, 2017; Melamed et al., 2016; Waters et al., 2017). By contrast, a key finding from our CLIP-seq data is that ProQ binds many mRNAs at the 3′ end.

Using the *cspE* mRNA as a model, we have presented genetic evidence that ProQ can stabilize mRNAs by counteracting decay that is dependent on the 3′→5′ exoribonuclease RNase II (Figures 5 and 6). In addition to the *cspE* mRNA, several other mRNAs with ProQ crosslinks in their 3′ UTR were unstable in the Δ*proQ* deletion strain, suggesting a more general role for ProQ in 3′-end-dependent protection from RNA degradation. Whereas, at least for ProQ, the attention may shift to the 3′ end of transcripts, we note that Rho-independent terminators located in 3′ UTRs are prominent targets of Hfq as well (Holmqvist et al., 2016). Furthermore, Hfq has been implicated in 3′-end-dependent processes. Hfq-bound sRNAs become highly unstable in the absence of Hfq but regain stability in the absence of PNPase (Andrade et al., 2012), suggesting that Hfq can protect against exoribonuclease activity through 3′ end binding. At some mRNA 3′ ends, Hfq stimulates polyadenylation (Hajnsdorf and Régnier, 2000; Le Derout et al., 2003). Thus, both Hfq and ProQ act at RNA 3′ ends, albeit not on the same RNA species (Figure S2). Considering other recent findings of regulatory sequences in transcribed 3′ UTRs (López-Garrido et al., 2014; Ruiz de los Mozos et al., 2013), the 3′ end of bacterial mRNAs warrants more investigations into its importance for post-transcriptional control.

Mechanism of RNA Stabilization

What are the possible mechanisms by which ProQ counteracts RNase activity? To stabilize the *cspE* mRNA, ProQ could directly interfere with RNase II activity. However, evidence that ProQ and RNase II interact *in vivo* is so far lacking, as judged from coIP experiments (Figure S6). In a more likely scenario, ProQ may sterically obstruct access of RNase II to the mRNA. RNase II efficiently degrades single-stranded RNAs that carry 3′ poly(A) tails (Coburn and Mackie, 1996); in *E. coli*, the *cspE* mRNA both undergoes polyadenylation and is a known substrate for RNase II (Mohanty and Kushner, 2006). However, polyadenylation by poly(A) polymerase I does not significantly affect the stability of *cspE* mRNA, neither in the presence nor in the absence of ProQ (Figure S7). Because stable RNA stem loops impede the processive action of RNase II (Spickler and Mackie, 2000), the *cspE* terminator is likely to block RNase II activity. Similar to the enhanced effect of PNPase activity at structured RNA by helicase RhlB (Py et al., 1996), full RNase-II-dependent decay of *cspE* mRNA may require unwinding of the 3′ terminator hairpin by other factors. To address this, it will be important to identify these putative factors and test their contributions to RNase-II-mediated decay *in vitro*, together with that of ProQ.

Whereas the *cspE* case is straightforward, the loss of ProQ causes variable effects among the tested mRNAs with 3′ located ProQ sites (Figure S5). For example, the *yobF-cspC* mRNA is much less affected by *proQ* deletion than is the *cspE* mRNA, even though ProQ binds both 3′ UTRs (Table S2). Possibly, the effects of ProQ on mRNA stability vary dependent on which decay pathway dominates. Thus, ProQ may only contribute

substantially when an exoribonucleolytic activity drives the overall decay rate. Conversely, when endonucleolytic cleavage near the 5′ end dominates decay kinetics, the inhibition ProQ exerts on exonucleolytic activity would hardly affect the decay rate. In agreement with such a scenario, *cspE* mRNA has only one predicted RNase E cleavage site, whereas six are known in the *yobF-cspC* mRNA (Chao et al., 2017).

In addition to protecting from 3′ exoribonuclease activity, ProQ may also stabilize transcripts against endonucleolytic, RNase E-mediated decay. Cross-comparison of the *Salmonella* ProQ CLIP-seq with global RNase E data (Chao et al., 2017) shows that 35% (162/467) of all ProQ peaks overlap with RNase E cleavage sites. One example is RaiZ, in which the ProQ binding site covers several RNase E cleavage sites (Figure 2D). Clearly, a global analysis of RNA decay is needed to assess the generality of ProQ's role in RNA stabilization.

A Structure Recognition Code for FinO-Domain Proteins?

Analysis of binding site characteristics suggested that ProQ interacts with structured RNA, without any preference for a specific sequence (Figures 2 and S3). In accordance with this, structure-destabilizing mutations in the *cspE* terminator stem resulted in reduced ProQ binding (Figure 5D). Comparative CLIP-seq revealed that many ProQ binding sites are conserved between *E. coli* and *Salmonella*, which may rely on structural rather than sequence conservation (Figure 3). This pattern suggests a structural code whereby ProQ selects its sRNA and mRNA targets from thousands of different cellular transcripts. We predict that the structural code is read by the FinO domain of ProQ. The FinO protein binds to the 3′ flank of the terminator structure of the FinP antisense RNA (Arthur et al., 2011; Jerome and Frost, 1999), which is strongly reminiscent of the motif predicted for ProQ binding sites (Figure 2C). Moreover, the FinO-like protein RocC in *Legionella* depends on the terminator stem loop of the RocR sRNA for binding (Attaiech et al., 2016).

Despite their preferences for similar RNA motifs, FinO domain proteins display puzzling differences in specificity. In *Salmonella*, ProQ, expressed from the chromosome, and FinO, encoded on plasmid pSLT, are present in the same cell. Yet ProQ binds hundreds of different transcripts, whereas FinO appears to interact only with FinP and its target, *traJ* mRNA. *Legionella* RocC is also much more selective than ProQ; its main ligands are the sRNA RocR and a handful of RocR-regulated mRNAs (Attaiech et al., 2016). As the homology between ProQ, FinO, and RocC is restricted to the FinO domain (Attaiech et al., 2017), perhaps the specificity is not defined solely by the FinO domain itself but rather depends on the different flanking accessory domains of these proteins. This is an attractive possibility, because RNA binding activity has been shown to involve not only the FinO domain but also the N-terminal extension of FinO, as well as the linker and C-terminal domain of ProQ (Ghetu et al., 2002; Gonzalez et al., 2017). The now available information about hundreds of ProQ binding sites will enable the selection of representative RNA ligands for *in vitro* structure-function studies to understand what determines the varied degrees of target specificity of FinO-like proteins. Moreover, different types of *in vivo* target site mapping, for example, formaldehyde crosslinking-based fCLIP-seq, which maps interactions between dsRNA-binding proteins and structured RNAs (Kim et al., 2017), may be used to refine the binding sites of these proteins *in vivo*.

Conclusions

Our present work adds ProQ binding sites to a growing list of crosslinking-based target site maps for global RBPs, including Hfq and CsrA, in *Salmonella* and *E. coli* (Holmqvist et al., 2016; Potts et al., 2017; Tree et al., 2014; this work). This list is predicted to be incomplete, as coIP-based studies have suggested hundreds of putative RNA targets for candidate global RBPs, such as the cold-shock proteins (Michaux et al., 2017). Understanding how these proteins work together to orchestrate gene expression remains a major challenge, but it is becoming clear that RBP functions in bacteria are more global than thought until recently.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- METHOD DETAILS
  - Media and Growth Conditions
  - Bacterial Strains and Plasmids
  - UV Crosslinking, Immunoprecipitation, and RNA Purification
  - cDNA Library Preparation
  - Sequencing
  - Processing of Sequence Reads and Mapping
  - Read Count Normalization
  - Peak Calling
  - Analysis of Crosslink-Specific Mutations
  - Meta-gene Analysis
  - Analysis of Sequence and Structure Motifs
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Comparison of Salmonella and *E. coli* Peaks
  - Comparison of ProQ Binding Sites and RNase E Cleavage Sites
  - Comparison of Hfq and ProQ Terminators
  - Northern Blot
  - Western Blot
  - Electromobility Shift Assays
  - coIP
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and five tables and can be found with this article online at https://doi.org/10.1016/j.molcel.2018.04.017.

ACKNOWLEDGMENTS

We thank Stan Gorski, Kai Papenfort, Cynthia Sharma, Alexander Westermann, and Gerhart Wagner for reading the manuscript; Victoria McParland and Tatjana Achmedov for excellent technical assistance; and Yanjie Chao, Alexandre Smirnov, and Sandra Viegas for sharing bacterial strains. This

work was supported by DFG grant Vo875-14/1 (Vogel). Erik Holmqvist received support from EMBO (long-term fellowship) and the Wenner-Gren Foundations.

AUTHOR CONTRIBUTIONS

J.V. and E.H. designed the experiments. E.H. performed the experiments. L.L., T.B., and L.B. carried out bioinformatic analyses. E.H. and J.V. wrote the manuscript, which was discussed, modified, and improved by all authors. J.V. supervised the project.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: December 12, 2017
Revised: April 5, 2018
Accepted: April 19, 2018
Published: May 24, 2018

REFERENCES

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. *11*, R106.

Andrade, J.M., Pobre, V., Silva, I.J., Domingues, S., and Arraiano, C.M. (2009). The role of 3′-5′ exoribonucleases in RNA degradation. Prog. Mol. Biol. Transl. Sci. *85*, 187–229.

Andrade, J.M., Pobre, V., Matos, A.M., and Arraiano, C.M. (2012). The crucial role of PNPase in the degradation of small RNAs that are not associated with Hfq. RNA *18*, 844–855.

Arthur, D.C., Edwards, R.A., Tsutakawa, S., Tainer, J.A., Frost, L.S., and Glover, J.N. (2011). Mapping interactions between the RNA chaperone FinO and its RNA targets. Nucleic Acids Res. *39*, 4450–4463.

Attaiech, L., Boughammoura, A., Brochier-Armanet, C., Allatif, O., Peillard-Fiorente, F., Edwards, R.A., Omar, A.R., MacMillan, A.M., Glover, J.N., and Charpentier, X. (2016). Silencing of natural transformation by an RNA chaperone and a multitarget small RNA. Proc. Natl. Acad. Sci. USA *113*, 8813–8818.

Attaiech, L., Glover, J.N.M., and Charpentier, X. (2017). RNA chaperones step out of Hfq's shadow. Trends Microbiol. *25*, 247–249.

Bailey, T.L., Johnson, J., Grant, C.E., and Noble, W.S. (2015). The MEME suite. Nucleic Acids Res. *43* (W1), W39–W49.

Bossi, L., and Figueroa-Bossi, N. (2016). Competing endogenous RNAs: a target-centric view of small RNA regulation in bacteria. Nat. Rev. Microbiol. *14*, 775–784.

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.

Chao, Y., Li, L., Girodat, D., Förstner, K.U., Said, N., Corcoran, C., Śmiga, M., Papenfort, K., Reinhardt, R., Wieden, H.J., et al. (2017). In vivo cleavage map illuminates the central role of RNase E in coding and non-coding RNA pathways. Mol. Cell *65*, 39–51.

Chaulk, S.G., Smith Frieday, M.N., Arthur, D.C., Culham, D.E., Edwards, R.A., Soo, P., Frost, L.S., Keates, R.A., Glover, J.N., and Wood, J.M. (2011). ProQ is an RNA chaperone that controls ProP levels in Escherichia coli. Biochemistry *50*, 3095–3106.

Chen, J., and Gottesman, S. (2017). Hfq links translation repression to stress-induced mutagenesis in *E. coli*. Genes Dev. *31*, 1382–1395.

Chen, B., Yun, J., Kim, M.S., Mendell, J.T., and Xie, Y. (2014). PIPE-CLIP: a comprehensive online tool for CLIP-seq data analysis. Genome Biol. *15*, R18.

Chinni, S.V., Raabe, C.A., Zakaria, R., Randau, G., Hoe, C.H., Zemann, A., Brosius, J., Tang, T.H., and Rozhdestvensky, T.S. (2010). Experimental identification and characterization of 97 novel npcRNA candidates in Salmonella enterica serovar Typhi. Nucleic Acids Res. *38*, 5893–5908.

Coburn, G.A., and Mackie, G.A. (1996). Overexpression, purification, and properties of Escherichia coli ribonuclease II. J. Biol. Chem. *271*, 1048–1053.

Darling, A.C., Mau, B., Blattner, F.R., and Perna, N.T. (2004). Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res. *14*, 1394–1403.

Darty, K., Denise, A., and Ponty, Y. (2009). VARNA: interactive drawing and editing of the RNA secondary structure. Bioinformatics *25*, 1974–1975.

De Lay, N., Schu, D.J., and Gottesman, S. (2013). Bacterial small RNA-based negative regulation: Hfq and its accomplices. J. Biol. Chem. *288*, 7996–8003.

Dugar, G., Herbig, A., Förstner, K.U., Heidrich, N., Reinhardt, R., Nieselt, K., and Sharma, C.M. (2013). High-resolution transcriptome maps reveal strain-specific regulatory features of multiple Campylobacter jejuni isolates. PLoS Genet. *9*, e1003495.

Durand, S., Tomasini, A., Braun, F., Condon, C., and Romby, P. (2015). sRNA and mRNA turnover in Gram-positive bacteria. FEMS Microbiol. Rev. *39*, 316–330.

Feliciano, J.R., Grilo, A.M., Guerreiro, S.I., Sousa, S.A., and Leitão, J.H. (2016). Hfq: a multifaceted RNA chaperone involved in virulence. Future Microbiol. *11*, 137–151.

Fender, A., Elf, J., Hampel, K., Zimmermann, B., and Wagner, E.G.H. (2010). RNAs actively cycle on the Sm-like protein Hfq. Genes Dev. *24*, 2621–2626.

Förstner, K.U., Vogel, J., and Sharma, C.M. (2014). READemption-a tool for the computational analysis of deep-sequencing-based transcriptome data. Bioinformatics *30*, 3421–3423.

Fozo, E.M., Kawano, M., Fontaine, F., Kaya, Y., Mendieta, K.S., Jones, K.L., Ocampo, A., Rudd, K.E., and Storz, G. (2008). Repression of small toxic protein synthesis by the Sib and OhsC small RNAs. Mol. Microbiol. *70*, 1076–1093.

Friedersdorf, M.B., and Keene, J.D. (2014). Advancing the functional utility of PAR-CLIP by quantifying background binding to mRNAs and lncRNAs. Genome Biol. *15*, R2.

Gardner, P.P., Barquist, L., Bateman, A., Nawrocki, E.P., and Weinberg, Z. (2011). RNIE: genome-wide prediction of bacterial intrinsic terminators. Nucleic Acids Res. *39*, 5845–5852.

Ghetu, A.F., Arthur, D.C., Kerppola, T.K., and Glover, J.N. (2002). Probing FinO-FinP RNA interactions by site-directed protein-RNA crosslinking and gelFRET. RNA *8*, 816–823.

Glover, J.N., Chaulk, S.G., Edwards, R.A., Arthur, D., Lu, J., and Frost, L.S. (2015). The FinO family of bacterial RNA chaperones. Plasmid *78*, 79–87.

Gonzalez, G.M., Hardwick, S.W., Maslen, S.L., Skehel, J.M., Holmqvist, E., Vogel, J., Bateman, A., Luisi, B.F., and Broadhurst, R.W. (2017). Structure of the *Escherichia coli* ProQ RNA-binding protein. RNA *23*, 696–711.

Gorski, S.A., Vogel, J., and Doudna, J.A. (2017). RNA-based recognition and targeting: sowing the seeds of specificity. Nat. Rev. Mol. Cell Biol. *18*, 215–228.

Hajnsdorf, E., and Régnier, P. (2000). Host factor Hfq of Escherichia coli stimulates elongation of poly(A) tails by poly(A) polymerase I. Proc. Natl. Acad. Sci. USA *97*, 1501–1505.

Hoffmann, S., Otto, C., Doose, G., Tanzer, A., Langenberger, D., Christ, S., Kunz, M., Holdt, L.M., Teupser, D., Hackermüller, J., and Stadler, P.F. (2014). A multi-split mapping algorithm for circular RNA, splicing, trans-splicing and fusion detection. Genome Biol. *15*, R34.

Holmqvist, E., Wright, P.R., Li, L., Bischler, T., Barquist, L., Reinhardt, R., Backofen, R., and Vogel, J. (2016). Global RNA recognition patterns of post-transcriptional regulators Hfq and CsrA revealed by UV crosslinking in vivo. EMBO J. *35*, 991–1011.

Hör, J., and Vogel, J. (2017). Global snapshots of bacterial RNA networks. EMBO J. *36*, 245–247.

Hui, M.P., Foley, P.L., and Belasco, J.G. (2014). Messenger RNA degradation in bacterial cells. Annu. Rev. Genet. *48*, 537–559.

Ishikawa, H., Otaka, H., Maki, K., Morita, T., and Aiba, H. (2012). The functional Hfq-binding module of bacterial sRNAs consists of a double or single hairpin

preceded by a U-rich sequence and followed by a 3′ poly(U) tail. RNA *18*, 1062–1074.

Jerome, L.J., and Frost, L.S. (1999). In vitro analysis of the interaction between the FinO protein and FinP antisense RNA of F-like conjugative plasmids. J. Biol. Chem. *274*, 10356–10362.

Jørgensen, M.G., Thomason, M.K., Havelund, J., Valentin-Hansen, P., and Storz, G. (2013). Dual function of the McaS small RNA in controlling biofilm formation. Genes Dev. *27*, 1132–1145.

Kim, B., Jeong, K., and Kim, V.N. (2017). Genome-wide mapping of DROSHA cleavage sites on primary microRNAs and noncanonical substrates. Mol. Cell *66*, 258–269.e5.

Kingsford, C.L., Ayanbule, K., and Salzberg, S.L. (2007). Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. Genome Biol. *8*, R22.

König, J., Zarnack, K., Luscombe, N.M., and Ule, J. (2012). Protein-RNA interactions: new genomic technologies and perspectives. Nat. Rev. Genet. *13*, 77–83.

Kröger, C., Colgan, A., Srikumar, S., Händler, K., Sivasankaran, S.K., Hammarlöf, D.L., Canals, R., Grissom, J.E., Conway, T., Hokamp, K., and Hinton, J.C. (2013). An infection-relevant transcriptomic compendium for Salmonella enterica Serovar Typhimurium. Cell Host Microbe *14*, 683–695.

Kunte, H.J., Crane, R.A., Culham, D.E., Richmond, D., and Wood, J.M. (1999). Protein ProQ influences osmotic activation of compatible solute transporter ProP in Escherichia coli K-12. J. Bacteriol. *181*, 1537–1543.

Langenberger, D., Bermudez-Santana, C., Hertel, J., Hoffmann, S., Khaitovich, P., and Stadler, P.F. (2009). Evidence for human microRNA-offset RNAs in small RNA sequencing data. Bioinformatics *25*, 2298–2301.

Le Derout, J., Folichon, M., Briani, F., Dehò, G., Régnier, P., and Hajnsdorf, E. (2003). Hfq affects the length and the frequency of short oligo(A) tails at the 3′ end of Escherichia coli rpsO mRNAs. Nucleic Acids Res. *31*, 4017–4023.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Group (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

López-Garrido, J., Puerta-Fernández, E., and Casadesús, J. (2014). A eukaryotic-like 3′ untranslated region in Salmonella enterica hilD mRNA. Nucleic Acids Res. *42*, 5894–5906.

Lorenz, R., Bernhart, S.H., Höner Zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P.F., and Hofacker, I.L. (2011). ViennaRNA package 2.0. Algorithms Mol. Biol. *6*, 26.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Lunde, B.M., Moore, C., and Varani, G. (2007). RNA-binding proteins: modular design for efficient function. Nat. Rev. Mol. Cell Biol. *8*, 479–490.

Maticzka, D., Lange, S.J., Costa, F., and Backofen, R. (2014). GraphProt: modeling binding preferences of RNA-binding proteins. Genome Biol. *15*, R17.

Matin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal *17*, 10–12.

Melamed, S., Peer, A., Faigenbaum-Romm, R., Gatt, Y.E., Reiss, N., Bar, A., Altuvia, Y., Argaman, L., and Margalit, H. (2016). Global mapping of small RNA-target interactions in bacteria. Mol. Cell *63*, 884–897.

Michaux, C., Holmqvist, E., Vasicek, E., Sharan, M., Barquist, L., Westermann, A.J., Gunn, J.S., and Vogel, J. (2017). RNA target profiles direct the discovery of virulence functions for the cold-shock proteins CspC and CspE. Proc. Natl. Acad. Sci. USA *114*, 6824–6829.

Milner, J.L., and Wood, J.M. (1989). Insertion proQ220:Tn5 alters regulation of proline porter II, a transporter of proline and glycine betaine in Escherichia coli. J. Bacteriol. *171*, 947–951.

Mohanty, B.K., and Kushner, S.R. (2006). The majority of Escherichia coli mRNAs undergo post-transcriptional modification in exponentially growing cells. Nucleic Acids Res. *34*, 5695–5704.

Olejniczak, M., and Storz, G. (2017). ProQ/FinO-domain proteins: another ubiquitous family of RNA matchmakers? Mol. Microbiol. *104*, 905–915.

Otaka, H., Ishikawa, H., Morita, T., and Aiba, H. (2011). PolyU tail of rho-independent terminator of bacterial small RNAs is essential for Hfq action. Proc. Natl. Acad. Sci. USA *108*, 13059–13064.

Papenfort, K., Pfeiffer, V., Mika, F., Lucchini, S., Hinton, J.C., and Vogel, J. (2006). SigmaE-dependent small RNAs of Salmonella respond to membrane stress by accelerating global omp mRNA decay. Mol. Microbiol. *62*, 1674–1688.

Perkins, T.T., Kingsley, R.A., Fookes, M.C., Gardner, P.P., James, K.D., Yu, L., Assefa, S.A., He, M., Croucher, N.J., Pickard, D.J., et al. (2009). A strand-specific RNA-seq analysis of the transcriptome of the typhoid bacillus Salmonella typhi. PLoS Genet. *5*, e1000569.

Potts, A.H., Vakulskas, C.A., Pannuri, A., Yakhnin, H., Babitzke, P., and Romeo, T. (2017). Global role of the bacterial post-transcriptional regulator CsrA revealed by integrated transcriptomics. Nat. Commun. *8*, 1596.

Py, B., Higgins, C.F., Krisch, H.M., and Carpousis, A.J. (1996). A DEAD-box RNA helicase in the Escherichia coli RNA degradosome. Nature *381*, 169–172.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

Romeo, T., Vakulskas, C.A., and Babitzke, P. (2013). Post-transcriptional regulation on a global scale: form and function of Csr/Rsm systems. Environ. Microbiol. *15*, 313–324.

Ruiz de los Mozos, I., Vergara-Irigaray, M., Segura, V., Villanueva, M., Bitarte, N., Saramago, M., Domingues, S., Arraiano, C.M., Fechter, P., Romby, P., et al. (2013). Base pairing interaction between 5′- and 3′-UTRs controls icaR mRNA translation in Staphylococcus aureus. PLoS Genet. *9*, e1004001.

Saramago, M., Domingues, S., Viegas, S.C., and Arraiano, C.M. (2014). Biofilm formation and antibiotic resistance in Salmonella Typhimurium are affected by different ribonucleases. J. Microbiol. Biotechnol. *24*, 8–12.

Sauer, E., and Weichenrieder, O. (2011). Structural basis for RNA 3′-end recognition by Hfq. Proc. Natl. Acad. Sci. USA *108*, 13065–13070.

Sittka, A., Pfeiffer, V., Tedin, K., and Vogel, J. (2007). The RNA chaperone Hfq is essential for the virulence of Salmonella typhimurium. Mol. Microbiol. *63*, 193–217.

Smirnov, A., Förstner, K.U., Holmqvist, E., Otto, A., Günster, R., Becher, D., Reinhardt, R., and Vogel, J. (2016). Grad-seq guides the discovery of ProQ as a major small RNA-binding protein. Proc. Natl. Acad. Sci. USA *113*, 11591–11596.

Smirnov, A., Wang, C., Drewry, L.L., and Vogel, J. (2017). Molecular mechanism of mRNA repression in *trans* by a ProQ-dependent small RNA. EMBO J. *36*, 1029–1045.

Smith, M.N., Crane, R.A., Keates, R.A., and Wood, J.M. (2004). Overexpression, purification, and characterization of ProQ, a posttranslational regulator for osmoregulatory transporter ProP of Escherichia coli. Biochemistry *43*, 12979–12989.

Sowa, S.W., Gelderman, G., Leistra, A.N., Buvanendiran, A., Lipp, S., Pitaktong, A., Vakulskas, C.A., Romeo, T., Baldea, M., and Contreras, L.M. (2017). Integrative FourD omics approach profiles the target network of the carbon storage regulatory system. Nucleic Acids Res. *45*, 1673–1686.

Spickler, C., and Mackie, G.A. (2000). Action of RNase II and polynucleotide phosphorylase against RNAs containing stem-loops of defined structure. J. Bacteriol. *182*, 2422–2427.

Stocker, B.A., Hoiseth, S.K., and Smith, B.P. (1983). Aromatic-dependent "Salmonella sp." as live vaccine in mice and calves. Dev. Biol. Stand. *53*, 47–54.

Tree, J.J., Granneman, S., McAteer, S.P., Tollervey, D., and Gally, D.L. (2014). Identification of bacteriophage-encoded anti-sRNAs in pathogenic Escherichia coli. Mol. Cell *55*, 199–213.

Updegrove, T.B., Zhang, A., and Storz, G. (2016). Hfq: the flexible RNA matchmaker. Curr. Opin. Microbiol. *30*, 133–138.

Vakulskas, C.A., Potts, A.H., Babitzke, P., Ahmer, B.M., and Romeo, T. (2015). Regulation of bacterial virulence by Csr (Rsm) systems. Microbiol. Mol. Biol. Rev. *79*, 193–224.

Van Assche, E., Van Puyvelde, S., Vanderleyden, J., and Steenackers, H.P. (2015). RNA-binding proteins involved in post-transcriptional regulation in bacteria. Front. Microbiol. *6*, 141.

Viegas, S.C., Pfeiffer, V., Sittka, A., Silva, I.J., Vogel, J., and Arraiano, C.M. (2007). Characterization of the role of ribonucleases in Salmonella small RNA decay. Nucleic Acids Res. *35*, 7651–7664.

Vogel, J., and Luisi, B.F. (2011). Hfq and its constellation of RNA. Nat. Rev. Microbiol. *9*, 578–589.

Vogel, J., Argaman, L., Wagner, E.G., and Altuvia, S. (2004). The small RNA IstR inhibits synthesis of an SOS-induced toxic peptide. Curr. Biol. *14*, 2271–2276.

Wagner, E.G.H., and Romby, P. (2015). Small RNAs in bacteria and archaea: who they are, what they do, and how they do it. Adv. Genet. *90*, 133–208.

Waters, S.A., McAteer, S.P., Kudla, G., Pang, I., Deshpande, N.P., Amos, T.G., Leong, K.W., Wilkins, M.R., Strugnell, R., Gally, D.L., et al. (2017). Small RNA interactome of pathogenic E. coli revealed through crosslinking of RNase E. EMBO J. *36*, 374–387.

Weinberg, Z., and Breaker, R.R. (2011). R2R–software to speed the depiction of aesthetic consensus RNA secondary structures. BMC Bioinformatics *12*, 3.

Xu, H., Luo, X., Qian, J., Pang, X., Song, J., Qian, G., Chen, J., and Chen, S. (2012). FastUniq: a fast de novo duplicates removal tool for paired short reads. PLoS ONE *7*, e52249.

Yao, Z., Weinberg, Z., and Ruzzo, W.L. (2006). CMfinder–a covariance model based RNA motif finding algorithm. Bioinformatics *22*, 445–452.

Zhang, C., and Darnell, R.B. (2011). Mapping in vivo protein-RNA interactions at single-nucleotide resolution from HITS-CLIP data. Nat. Biotechnol. *29*, 607–614.

Zhang, A., Wassarman, K.M., Rosenow, C., Tjaden, B.C., Storz, G., and Gottesman, S. (2003). Global analysis of small RNA and mRNA targets of Hfq. Mol. Microbiol. *50*, 1111–1124.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-mouse-HRP	Thermo Fisher Scientific	Cat# 31430
Anti-FLAG	Sigma-Aldrich	Cat#F3165
Bacterial and Virus Strains		
JVS-1574: *Salmonella typhimurium* SL1344, Str^R *hisG rpsL xyl*	Laboratory strain collection	N/A
JVS-10314: SL1344 *proQ-3xflag*	Smirnov et al., 2016	N/A
JVS-10315: SL1344 Δ*proQ::KanR*	Smirnov et al., 2016	N/A
JVS-10317: SL1344 Δ*proQ*	Smirnov et al., 2016	N/A
JVS-5084: SL1344 Δ*cspC* Δ*cspE*	This study	N/A
JVS-11295: SL1344 Δ*cspC* Δ*cspE* Δ*proQ*::KmR	This study	N/A
JVS-871: SL1344 Δ*rnr*	This study	N/A
JVS-872: SL1344 Δ*pnp*	This study	N/A
JVS-873: SL1344 Δ*rnb*	This study	N/A
EHS-1286: SL1344 Δ*rnr* Δ*proQ*::KanR	This study	N/A
EHS-1287: SL1344 Δ*pnp* Δ*proQ*::KanR	This study	N/A
EHS-1288: SL1344 Δ*rnb* Δ*proQ*::KanR	This study	N/A
EHS-1471: SL1344 *rnb-3xflag*	This study	N/A
JVS-10520: *Escherichia coli* W3110 proQ-3xflag::KanR	This study	N/A
JVS-869: SL1344 Δ*pcnB*	This study	N/A
EHS-1289: SL1344 Δ*pcnB* Δ*proQ*::KanR	This study	N/A
Chemicals, Peptides, and Recombinant Proteins		
Anti-FLAG magnetic beads	Sigma-Aldrich	Cat# M8823
DNase I	Thermo Fisher Scientific	Cat# 4716728001
Benzonase nuclease	Sigma-Aldrich	Cat# E8263
Calf Intestinal Alkaline Phosphatase	New England Biolabs	Cat# M0290
T4 Polynucleotide Kinase	Thermo Fisher Scientific	Cat# EK0031
SUPERase In	Thermo Fisher Scientific	Cat# AM2696
Proteinase K	Thermo Fisher Scientific	Cat# EO0491
GlycoBlue Coprecipitant	Thermo Fisher Scientific	Cat# AM9515
SYBRGold	Thermo Fisher Scientific	Cat# S11494
3xFLAG peptide	Sigma-Aldrich	Cat# F4799
2xLongAmp Taq PCR Master Mix	New England Biolabs	Cat# M0287
Critical Commercial Assays		
NEBNext Multiplex Small RNA Library Prep Set	New England Biolabs	Cat# E7300
MEGAscript T7 transcription kit	Thermo Fisher Scientific	Cat# AM1334
Deposited Data		
Raw and analyzed sequencing data	This study	GEO: GSE106633
Original imaging data	Mendeley Data	https://data.mendeley.com/datasets/pcvymmzps8/1
Oligonucleotides		
See Table S3 for all oligonucleotides used in this study.	This paper	N/A
Software and Algorithms		
Cutadapt (v. 1.5/1.7.1)	Matin, 2011	https://cutadapt.readthedocs.io/en/stable/#
cmpfastq	NIHR Biomedical Research Centre for Mental Health	http://compbio.brc.iop.kcl.ac.uk/software/cmpfastq.php

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
FastUniq	Xu et al., 2012	https://sourceforge.net/projects/fastuniq/
READemption (v.0.3.7)	Förstner et al., 2014	https://pythonhosted.org/READemption/
segemehl (v.0.2.0)	Hoffmann et al., 2014	http://www.bioinf.uni-leipzig.de/Software/segemehl/
PEAKachu		https://github.com/tbischler/PEAKachu
DESeq2	Love et al., 2014	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
samtools (v.0.1.19)	Li et al., 2009	http://www.htslib.org/doc/samtools.html
PIPE-CLIP	Chen et al., 2014	http://pipeclip.qbrc.org/
RNAfold	Lorenz et al., 2011	http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi
MEME (v.4.9.1)	Bailey et al., 2015	http://meme-suite.org
CMfinder (v.0.2)	Yao et al., 2006	http://bio.cs.washington.edu/yzizhen/CMfinder/
GraphProt	Maticzka et al., 2014	http://www.bioinf.uni-freiburg.de/Software/GraphProt/
TSSpredator	Dugar et al., 2013	http://it.inf.uni-tuebingen.de/?page_id=190
bedtools	Quinlan and Hall, 2010	http://bedtools.readthedocs.io/en/latest/#
TransTermHP	Kingsford et al., 2007	http://transterm.ccb.jhu.edu

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Jörg Vogel (joerg.vogel@uni-wuerzburg.de).

METHOD DETAILS

Media and Growth Conditions

Salmonella Typhimurium strain SL1344 (Stocker et al., 1983) or *Escherichia coli* strain W3110 were grown in liquid LB medium or on solid LB agar medium at 30 or 37°C. When appropriate, liquid or solid media were supplemented with 30 μg/mL chloramphenicol, 100 μg/mL ampicillin, 50 μg/mL Kanamycin, 500 μg/ml rifampicin, or 0.002% L-arabinose.

Bacterial Strains and Plasmids

All bacterial strains used in this study are listed in Table S4. All plasmids used in this study are listed in Table S5. Construction of plasmids pJV300, pZE12-ProQ, and pKP8-35 have been previously described (Papenfort et al., 2006; Sittka et al., 2007; Smirnov et al., 2016).

To construct strain JVS-5084, PCR products (primers JVO-4496/JVO-4497 and JVO-4500/JVO-4501) amplified from plasmids pKD3 or pKD4, respectively, were recombined into *Salmonella* using the Lambda Red recombination system. The mutations were sequentially transferred to a WT *Salmonella* strain with phage P22 transduction followed by removal of the antibiotic resistance cassettes using plasmid pCP20.

To construct strain EHS-1471, a PCR product (JVO-14435/JVA-14436) amplified from pSUB11 was recombined into *Salmonella* using the Lambda Red recombination system. The mutations were sequentially transferred to a WT *Salmonella* strain with phage P22 transduction followed by removal of the antibiotic resistance cassettes using plasmid pCP20.

Strains JVS-871, JVS-872 and JVS-873 were constructed by pCP20-mediated removal of antibiotic cassettes from the parental strains CMA-701, CMA-539 and CMA-700, respectively (Saramago et al., 2014; Viegas et al., 2007). Bacterial cultures were grown at 37°C with shaking at 220 rpm in LB broth supplemented with 100 μg/ml ampicillin, 50 μg/ml kanamycin, 20 μg/ml chloramphenicol and 500 μg/ml rifampicin, where appropriate.

Plasmid pEH691 was constructed by ligating a PCR product (primers JVO-4308/JVO-4309) encoding the *cspE* transcript to a PCR product (primers JVO-900/JVO-901) of plasmid pBAD33. Both fragments were cut with XbaI prior to ligation.

Plasmids pEH694, pEH695, pEH696 and pEH697 were constructed by ligating PCR products (forward primer JVO-4856 and reverse primers JVO-13866, JVO-13867, JVO-13868, and JVO-13869, respectively) amplified from pEH691.

Plasmid pEH313 was constructed by first ligating two linear DNA fragments generated by annealing of oligonucleotides JVO-9724/JVO-9725 and JVO-9726/JVO-972, respectively. The resulting fragment comprising a 2xStrepII-3xFLAG tandem affinity tag sequence was ligated into the NotI and HindIII sites of plasmid pBAD24, resulting in plasmid pEH299. pEH313 was constructed by cloning PCR products (primers JVO-9867/JVO-9868) into the NcoI site in pEH299.

UV Crosslinking, Immunoprecipitation, and RNA Purification

For each biological replicate, 400 mL bacterial culture was grown until an OD_{600} of 2.0. Half of the culture was directly placed in a 22x22 cm plastic tray and irradiated with UV-C light at 800 mJ. Cells were pelleted by centrifugation at 4°C, resuspended in lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 0.05% Tween, pH 8.0, lysozyme (5 mg/ml), DNase I (0,05 U/μl)) and mixed with 1 mL glass beads (0.1 mm radius). Cells were lysed on a Retsch MM400 instrument at 30 Hz for 10 min in the presence of 0.1 mm glass beads followed by two rounds of centrifugation for 15 min at 13000 rpm and 4°C. The cleared lysates were mixed with one volume of Urea-NP-T byffer (50 mM NaH_2PO_4, 300 mM NaCl, 0.05% Tween, pH 8.0, 8 M Urea), incubated 5 min at 65°C with shaking at 900 rpm and diluted 10x in ice-cold NP-T buffer.

Anti-FLAG magnetic beads (Sigma) were washed three times in NP-T buffer (60 μL 50% bead suspension per 200 mL bacterial culture), added to the lysate, and the mixture was rotated for 1 hr at 4°C. Beads were collected by centrifugation at 800xg, resuspended in 1 mL NP-T buffer, transferred to new tubes, and washed 2x with High-Salt buffer (50 mM NaH_2PO_4, 1 M NaCl, 0.05% Tween, pH 8.0) and 2x with NP-T buffer. Beads were resuspended in 200 μL NP-T buffer with1 mM $MgCl_2$ and 0.02 μL benzonase nuclease (Sigma), incubated 10 min at 37°C with shaking at 800 rpm, followed by a 2 min incubation on ice. After one wash with High-Salt buffer, and two washes with CIP buffer (100 mM NaCl, 50 mM Tris-HCl pH 7.4, 10 mM $MgCl_2$), the beads were resuspended in 200 μL CIP buffer with 10 units of Calf Intestinal Alkaline Phosphatase (NEB) and incubated 30 min at 37°C with shaking at 800 rpm. After one wash with High-Salt buffer and two washes with PNK buffer (50 mM Tris-HCl pH 7.4, 10 mM $MgCl_2$, 0.1 mM spermidine), the beads were resuspended in 200 μL PNK buffer with 20 units of T4 Polynucleotide Kinase and 20 μCi γ-^{32}P-ATP and incubated 30 min at 37°C.

After three washes with NP-T buffer, the beads were resuspended in 30 μL Protein Loading buffer (0.3 M Tris-HCl pH 6.8, 0.05% Bromophenol blue, 10% Glycerol, 7% DTT) and incubated 5 min at 95°C. The elution was repeated one time. The magnetic beads were collected on a magnetic separator and the supernatant was loaded and separated on a 15% SDS-polyacrylamide gel followed by transfer to a nitrocellulose membrane. The protein size marker was highlighted with a radioactively labeled marker pen, and the membrane was exposed to a phosphor screen for 30 min.

The autoradiogram was used as a template to cut out the regions of the membrane with radioactive signals from labeled RNA-protein complexes and equivalent regions for control samples. Each membrane piece was further cut into smaller pieces and incubated 30 min at 37°C with shaking at 1000 rpm in 400 μL PK solution (50 mM Tris-Hcl pH 7.4, 75 mM NaCl, 6 mM EDTA, 1% SDS, 10 units of SUPERaseIN (LifeTechnologies) and 0,4 mg Proteinase K (ThermoScientific)) whereafter 100 μL 9 M urea was added and the incubation was continued for an additional 30 min. 450 μL of the PK solution/urea was mixed with 450 μL Phenol:Chloroform:Isoamyl alcohol in a PhaseLock tube and incubated 5 min at 30°C with shaking at 1000 rpm followed by centrifugation for 12 min at 13000 rpm and 4°C. The aqueous phase was precipitated over night with 3 volumes of ice-cold ethanol, 1/10 volume of 3 M NaOAc pH 5.2, and 1 μL of GlycoBlue (Life Technologies) in LoBind tubes (Eppendorf). The precipitate was pelleted by centrifugation (30 min, 13000 rpm, 4°C), washed with 80% ethanol, centrifuged again (15 min, 13000 rpm, 4°C), dried 2 min at room temperature and resuspended in 10 μL sterile water.

cDNA Library Preparation

To enable sequencing on Illumina instruments, libraries were prepared using the NEBNext Multiplex Small RNA Library Prep Set for Illumina (#E7300, New England Biolabs) according to the manufacturer's instructions. 2.5 μL purified RNA (or sterile water as negative control) was mixed with 0.5 μL 3′ SR Adaptor (diluted 1:10) and 0.5 μL Nuclease Free Water, incubated 2 min at 70°C and chilled on ice. After addition of 5 μL 3′ Ligation Reaction Buffer and 1.5 μL 3′ Ligation Enzyme Mix the samples were incubated 60 min at 25°C. 0.25 μL of the SR RT Primer and 2.5 μL Nuclease Free Water was added followed by incubation 5 min at 75°C, 15 min at 37°C, and 15 min at 25°C. For ligation of the 5′ adaptor, the sample was mixed with 0.5 μL 5′ SR adaptor (denatured, diluted 1:10), 0.5 μL 10x Ligation Reaction Buffer, and 1.24 μL Ligation Enzyme Mix and incubated 60 min at 25°C. cDNA synthesis was carried out by the addition of 4 μL First Strand Synthesis Reaction buffer, 0.5 μL Murine RNase Inhibitor, and 0.5 μL ProtoScript Reverse Transcriptase and incubation at 50°C for 60 min. The reverse transcription activity was inhibited by a 15 min incubation at 70°C.

The cDNA was amplified by PCR by mixing 10 μL cDNA sample with 25 μL 2xLongAmp Taq PCR Master Mix, 1.25 μL SR Primer and 17.5 μL Nuclease Free Water in a thermal cycler with the following program: 30 s at 94°C, 20 rounds of (15 s at 94°C, 30 s at 62°C, 15 s at 70°C). The PCR reactions were purified on columns (QIAGEN), eluted in 10 μL sterile water, and loaded on 6% polyacrylamide gels with 7 M Urea together with a 50 bp DNA size marker (ThermoScientific). Gels were stained with SYBRGold (Life Technologies) and fragments between 140-250 bp were excised from the gels. Elution of DNA fragments were performed in 500 μL DNA elution buffer (NEB) at 16°C over night at 1000 rpm followed by EtOH precipitation. Pellets were resuspended in 10 μL sterile water. 2 μL gel purified DNA was mixed with 25 μL 2xLongAmp Taq PCR Master Mix, 2 μL each of primer JVO-11007 and JVO-11008 (10 μM), 19 μL sterile water and amplified using the following program: 30 s at 94°C, 6 rounds of (15 s at 94°C, 30 s at 60°C, 15 s at 65°C). PCR reactions were purified on columns (QIAGEN) and eluted in 15 μL sterile water.

Sequencing

High-throughput sequencing was performed at vertis Biotechnologie AG, Freising, Germany. cDNA libraries were pooled on an Illumina NextSeq 500 mid-output flow cell and sequenced in paired-end mode (2x75 cycles). Raw sequencing reads in FASTQ format and coverage files normalized by DEseq2 size factors (see below) are available via Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo) under accession number GEO: GSE106633.

Processing of Sequence Reads and Mapping

To assure high sequence quality, read 1 (R1) and read 2 (R2) files containing the Illumina paired-end reads in FASTQ format were trimmed independently from each other with a Phred score cutoff of 20 by the program fastq_quality_trimmer from FASTX toolkit version 0.0.13. In the same step, after quality trimming, the NEBNext R1 and R2 3′-adapters (R1: AGATCGGAAGAGCACACGTCT GAACTCCAGTCAC, R2: GATCGTCGGACTGTAGAACTCTGAACGTGTAGATCTCGGTGGTCGCCGTATCATT) were trimmed using Cutadapt version 1.5/1.7.1 (Matin, 2011) and reads without any remaining bases were discarded. Afterward, reads without a mate in the complementary read file were excluded using cmpfastq (http://compbio.brc.iop.kcl.ac.uk/software/cmpfastq.php). In order to remove putative PCR duplicates, paired-end reads were collapsed using FastUniq (Xu et al., 2012). Subsequently, a size filtering step was applied in which read pairs with at least one read shorter than 12 nt or longer than 25 nt were eliminated. The collections of remaining reads were mapped to the *Salmonella* Typhimurium SL1344 chromosome (NCBI Acc.-No: NC_016810.1) and plasmid (NCBI Acc.-No: NC_017718.1, NC_017719.1, NC_017720.1) or *E. coli* K12 (NC_000913.3) reference sequences using the RNA-seq pipeline READemption version 0.3.7 (Förstner et al., 2014) and segemehl version 0.2.0 (Hoffmann et al., 2014) with an accuracy cutoff of 80%. From the results, only reads mapping uniquely to one genomic position were considered for all subsequent analysis. For all analyses related to annotated genomic features such as CDSs, tRNAs, and rRNAs, gene annotations from NCBI were used.

We defined *ad hoc* transcriptional units (TUs) based on NCBI CDS annotations, transcription start site (TSS) annotations from (Kröger et al., 2013) and Rho-independent terminator predictions by RNIE (Gardner et al., 2011). Briefly, TUs were defined as starting on annotated primary TSSes and ending either with a predicted Rho-independent terminator or in the presence of an intergenic gap greater than 500 nt on the coding strand. In the absence of an upstream TSS, an arbitrary 100 nt 5′UTR was added upstream of the first CDS in the TU, and similarly in the absence of a terminator, an arbitrary 100 nt 3′UTR was added. In the event of a predicted primary TSS within an intergenic gap of less than 500 nt on the coding strand, the TU was ended 100 nt downstream of the preceding CDS, or at the end of the preceding CDS if the predicted primary TSS was less than 100 nt downstream. We defined 5′UTRs as the regions from the start of each predicted TU to the position upstream of the first CDS in the TU and 3′UTRs as the regions from one nt downstream of the last CDS in the TU to the end of the TU. sRNA annotations are based on (Chinni et al., 2010; Kröger et al., 2013; Perkins et al., 2009), and KU Förstner and J Vogel (unpublished data).

Read Count Normalization

As the CLIP-seq procedure involves specific capture of ProQ-bound transcripts, scaling normalization procedures that assume a substantial core of transcripts with unchanged abundances may fail. We performed an exploratory analysis of read counts summarized per position to isolate a core of positions present in pairs of crosslinked libraries and non-crosslinked controls, which could then be used for normalization. First, low-count positions were filtered by removing all positions with read counts less than 6 standard deviations from 0 for both libraries. The difference in read counts between the two libraries was then calculated. Plotted as a histogram, this displayed a clear bimodal distribution: one mode corresponding to positions with high counts in both crosslinked and non-crosslinked libraries, and a second mode corresponding to those with high counts in the crosslinked library and low counts in the non-crosslinked control. Positions in the first mode across all replicate crosslinked and control libraries were then used to compute normalization factor using the DESeq normalization procedure (Anders and Huber, 2010). These size factors were subsequently used for peak calling via the tool PEAKachu (see below).

Peak Calling

Peak calling was performed using the adaptive approach implemented in the tool PEAKachu (https://github.com/tbischler/PEAKachu, manuscript in preparation) in a similar way as described previously (Holmqvist et al., 2016). The tool was run in paired-end (-P) and paired-replicates (-r) mode using BAM files for the respective pairs of crosslinked and control libraries as input. The maximum fragment size (-M) was set to 25 and annotations in GFF format (see above) were used to map overlapping features to called peaks. For normalization, ‘‘manual’’ mode was selected together with previously calculated size factors (see above). Peak calling via the adaptive approach is performed in two consecutive steps. In the first step, initial peaks are defined via heuristic decomposition of read clusters computed by the blockbuster algorithm (Langenberger et al., 2009) based on pooled read alignments from all crosslinked libraries. In this step, PEAKachu applies a set of parameters for which default values were used. In the second step, PEAKachu runs DESeq2 (Love et al., 2014) to test each peak for significant enrichment of normalized read counts in the crosslinked compared to the control libraries. Initial peaks were tested for significance using the following parameter values: mad-multiplier (-m) 1.0, fold change (-f) 1.0 and adjusted p value (-Q) 0.01. Normalized coverage plots representing the numbers of mapped reads per nt were generated by PEAKachu for each replicon and strand to facilitate data visualization in a genome browser.

Analysis of Crosslink-Specific Mutations

To detect crosslinking-induced mutation sites from the CLIP-Seq data, we followed our previous approach (Holmqvist et al., 2016). Briefly, mutated sites were required to be present in both paired reads and subsequently the first read in the mutated pair was extracted using samtools (v 0.1.19). The enrichment of mutations in each library was detected using PIPE-CLIP (Chen et al., 2014), using a false discovery rate (FDR) cut-off of 0.1, and requiring at least 10 reads mapped to the mutation site. The crosslinked samples were further compared with non-crosslinked samples to identify the crosslink-specific mutations.

Meta-gene Analysis

The number of ProQ peaks was calculated by cumulatively counting sites relative to annotated genomic features, e.g., the first start codon and last stop codon of respective gene operons.

Analysis of Sequence and Structure Motifs

To evaluate the structural profiles around cross-link mutations, a sliding window of 20 nucleotides was stepped 5 nucleotides at a time across the sequence of interest. The minimum free energy (MFE) of these windows was calculated using the RNAfold algorithm from the ViennaRNA Package (Lorenz et al., 2011). The MFEs for each window were then compared to the 1000 shuffled sequences preserving dinucleotide composition, and *Z*-scores with upper and lower quantiles were reported.

To identify motifs in the vicinity of crosslink-induced mutations, MEME (version 4.9.1) (Bailey et al., 2015) was used on the sequences extended by 50 nt downstream with allowing the motif to have one nucleotide shift while the other parameters were kept at their default values. To further identify candidate structural motifs, CMfinder (version 0.2) (Yao et al., 2006) was applied on extended sequences using default parameters. The motifs were further visualized using R2R (Weinberg and Breaker, 2011).

To identify potentially structured sequence motifs, we ran GraphProt (Maticzka et al., 2014). As this method is discriminative, we constructed a negative control dataset consisting of random sites on transcripts expressed at OD 2.0 (Kröger et al., 2013), excluding ProQ-bound transcripts, selected to have the same distribution across annotated regions (5′ UTRs, CDS, 3′UTRs, sRNAs) as genuine ProQ binding sites. We selected twice as many random negative sequences as the positive sequences, as selecting more than this would require either reusing sequence regions or losing the balance in selected regions between true positives and our negative control set. In keeping with the recommendations in the GraphProt paper, we extended peaks by 15 nucleotides on either side to provide a ''viewpoint'' within which motif-finding occurs. The sequences were additionally extended on either side to a total length of 300 bases to provide flanking sequence for more accurate folding.

We first optimized the GraphProt parameters using the ls command, then used the built-in 10-fold cross validation to test whether an informative model could in principle be built using GraphProt. This procedure gave an average precision of ~68%, and an AUROC of ~.78, suggesting that a structure-informed GraphProt model could discriminate at least a subset of ProQ peaks from unbound sequence. We trained a GraphProt model using the full dataset, and used the top scoring 50 sequences under this model to generate a motif (Figure S3B).

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis was carried out in R. Welch's two sample t tests and Wilcoxon rank-sum tests were applied to the analysis of 3′ tail length in Figures S3E and S3F to avoid assumptions of equal variances or normality, respectively.

Comparison of Salmonella and *E. coli* Peaks

In order to compare specific ProQ binding site locations between *Salmonella* and *E. coli* we applied the SuperGenome approach implemented in the tool TSSpredator (Dugar et al., 2013) to map the two genomes to a common coordinate system. At first, Mauve (Darling et al., 2004) was applied to generate a whole genome alignment from the *Salmonella* and *E. coli* chromosome reference sequences. This alignment was used as input for TSSpredator version 1.06 to generate position-wise mappings between coordinates of SuperGenome and single reference sequences. Afterward, normalized coverage and called peaks for *Salmonella* and *E. coli* were mapped to the SuperGenome to allow comparison of the two datasets. Note that each nucleotide position in the SuperGenome can either be present in only one or both reference sequences. This can lead to a splitting of peak regions if a peak in one genome is interrupted by an indel relative to the other genome.

Comparison of ProQ Binding Sites and RNase E Cleavage Sites

To address the association of ProQ binding sites with RNase E cleavage sites, we retrieved the 22,033 RNase E cleavage sites from recent TIER-seq data (Chao et al., 2017), and compared them with ProQ binding peaks using bedtools (Quinlan and Hall, 2010). The statistical significance of these overlaps was evaluated using a random permutation approach. We have randomly shuffled the genomic locations of RNase E cleavage sites in the *Salmonella* Typhimurium genome while maintaining the same feature length and number of the ProQ peaks. This permutation was done 1000 times. A Z-score was calculated to represent the standard deviation from the mean. A p value was derived for evaluating the significance of the Z-score.

Comparison of Hfq and ProQ Terminators

To generate terminator consensus motifs for Hfq- and ProQ-bound sRNAs, sequences predicted by TransTermHP (Kingsford et al., 2007) and supported by RNA-seq that overlapped with CLIP peaks were extracted and used as input for motif search using CMfinder. To analyze the length of unpaired U-tails at the 3′ end of of Hfq- and ProQ-bound sRNAs, the secondary structure of these two different sRNAs classes were predicted using RNAfold (Lorenz et al., 2011), and single-stranded 3′-located U-tails were manually extracted. The length of these unpaired tails was compared with p values calculated by Welch Two sample t test and Wilcoxon rank sum test.

Northern Blot

Bacterial cultures were mixed with 0.2 volumes of stop solution (95% EtOH, 5% phenol) and snap-frozen in liquid nitrogen, thawed on ice, and pelleted by centrifugation at 4°C. Pellets were resuspended in lysis buffer (TE buffer at pH 8.0 with 0.5 mg/ml lysozyme and 1% SDS) and incubated 2 min at 65°C followed by addition of one volume of phenol. After vigorous mixing, samples were incubated 6 min at 65°C and centrifuged 15 min at 13000 rpm and 4°C. The aqueous phase was thoroughly mixed with one volume of chloroform, centrifuged for 15 min followed by precipitation of the aqueous phase in three volumes of EtOH and 1:20 volume of 3 M NaOAc (pH 5.2) for 30 min at −80°C. After 30 min centrifugation, pellets were washed once in 80% EtOH and centrifuged for 10 min. Pellets were dried 2 min at room temperature and resuspended in sterile water for 3 min at 65°C with agitation at 900 rpm. RNA integrity was analyzed on 1% agarose gels.

Five to ten micrograms of RNA was diluted in denaturing loading buffer, heated at 95°C for 5 min, cooled on ice for 2 min and loaded on denaturing gels (6% polyacrylamide, 7 M urea, 1xTBE). RNA was transferred from gels to nitrocellulose membranes (Hybond-XL, GE Healthcare). After crosslinking at 120 mJ, membranes were equilibrated in hybridization buffer (Hybri-Quick, Roti) for o1ne hr at 42°C where after radioactively 5′ end-labeled and denatured DNA probes were added to the hybridization buffer. After over night incubation at 42°C, membranes were washed one time in 1 × SSC/0.1%SDS and one time in 0.5 × SSC/0.1% SDS, followed by exposure to phosphor screens. Radioactive signals were detected with a phosphorimager device (Typhoon FLA 7000, GE Healthcare).

Western Blot

Bacterial cultures were pelleted by centrifugation at 4°C, resuspended in protein loading buffer, heated to 95°C for 5 min, and separated on 12% SDS-polyacrylamide gels. Proteins were transferred to PVDF membranes followed by blocking in 5% milk powder in TBST for 1 hr at room temperature, washing in TBST, incubation with an anti-FLAG antibody (Sigma) for 1 hr at room temperature, washing in TBST, incubation an with anti-mouse-HRP antibody (ThermoScientific) for 1 hr at room temperature, washing in TBST, and imaging on a ImageQuant device (GE Healthcare).

Electromobility Shift Assays

Five nanomolar of ^{32}P-5′ labeled *in vitro* transcribed RNA (MEGAscript, Life Technologies) was mixed with varying concentrations of purified ProQ (Fender et al., 2010; Gonzalez et al., 2017; Smirnov et al., 2016) in reaction buffer (25 mM TrisHCl, pH 7,4, 100 mM NaCl, 1 mM MgCl2) at 37°C for 5 min. The samples were thereafter separated in native 6% polyacrylamide / 0.5xTBE gels at 4°C. After vacuum drying of the gels, the radioactive signals were visualized using a Phosphorimager.

coIP

A 50 mL bacterial culture was grown was grown until an OD_{600} of 2.0. Cells were pelleted by centrifugation at 4°C, resuspended in lysis buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100) and mixed with 1 mL glass beads (0.1 mm radius). Cells were lysed on a Retsch MM400 instrument at 30 Hz for 10 min followed by two rounds of centrifugation for 15 min at 13000 rpm and 4°C. Forty microliter anti-FLAG magnetic beads (Sigma) was pre-washed three times in TBS (50 mM Tris-HCl pH 7.4, 150 mM NaCl), added to the cleared lysate, and the mixture was rotated for 1 hr at 4°C. The beads were washed four times in 1 mL TBS, followed by a 30 min incubation in 200 μL TBS supplemented with FLAG peptide (Sigma) at a final concentration of 150 ng/μl to elute proteins bound to the anti-FLAG antibody. For analysis of eluted proteins, eluates were mixed with three volumes ice-cold acetone, incubated 2 h at −20°C, centrifuged for 30 min at 13 000 *g*, and washed twice in 500 μL acetone. Air-dried pellets where resuspended in 50 μL protein loading buffer. For analysis of co-precipitated RNA, eluates where mixed with P:C:I followed by phase separation using centrifugation. RNA was precipitated in ethanol and resupended in sterile water.

DATA AND SOFTWARE AVAILABILITY

The accession number for the raw sequencing reads in FASTQ format and coverage files normalized by DEseq2 size factors reported in this paper is GEO: GSE106633.