RNA-binding proteins in bacteria

Erik Holmqvist[1] and Jörg Vogel[2,3,*]

[1]Department of Cell and Molecular Biology, Biomedical Center, Uppsala University, Box 596, S-75124 Uppsala, Sweden

[2]Helmholtz Institute for RNA-based Infection Research (HIRI), D-97080 Würzburg, Germany

[3]Institute of Molecular Infection Biology, University of Würzburg, D-97080 Würzburg, Germany

*Corresponding author
E-mail: joerg.vogel@uni-wuerzburg.de
Phone:+49-931-3182575
Fax: +49-931-3182578

ABSTRACT

RNA-binding proteins (RBPs) are at the heart of central cellular processes, dictating the fate of virtually all RNA molecules in the cell. Starting with pioneering work on ribosomal proteins, bacterial RBPs have long served as models for molecular studies of RNA-protein interactions. Work over the years has defined major RBPs that act on cellular transcripts at the various stages of bacterial gene expression and enable their integration into post-transcriptional networks, including RBPs that crucially aid the activities of small noncoding RNAs. However, we argue that bacterial RBP research is entering a new era as RNA sequencing-based methods now allow the study of RBP activity in a truly global manner *in vivo*. Moreover, the soaring interest in understudied members of host-associated microbiota and environmental communities is likely to unveil new RBPs and greatly expand our knowledge of RNA-protein interactions in bacteria.

INTRODUCTION

Interactions of proteins with RNAs are fundamental for life as we know it. Consequently, RNA-binding proteins (RBPs) are found in all living organisms. The most conserved RBPs are a core set of ribosomal proteins (r-proteins) **[G]** that were already present in the last universal common ancestor [1]. Bacteria use RBPs both as structural components of larger complexes, such as the ribosome, and as regulators of many cellular processes, including the synthesis, modification, translation, processing, and degradation of RNA (FIG. 1a).

RBPs interact with their ligands through RNA-binding domains (RBDs) that recognize short RNA sequences. Classic RBDs are widespread among bacterial RBPs. These include the S1 domain and the cold-shock domain (CSD) of the oligonucleotide/oligosaccharide (OB) binding superfamily [2,3], the Sm and Sm-like domains [4], the RNA recognition motif (RRM) [5], the K homology (KH) domain [6], the double-stranded RNA-binding domain (dsRBD) [7], and the PAZ and PIWI domains [8] (FIG. 1b).

Our present knowledge of how RBPs function is dominated by work in eukaryotes in which specificity and versatility of RNA-protein interactions often arise from cooperative binding of several RBPs to one transcript, or by combining multiple RBDs—different ones or repeats of the same—in a single protein. A modular architecture allows RBPs to recognize different RNA motifs in the same RNA ligand, or to simultaneously interact with different RNA ligands [9]. Bacterial RBPs tend to be more streamlined, often possessing only a single RBD per protein (FIG. 1b-c).

On the face of it, the fact that r-proteins make up one third of the ~180 annotated RBPs in a typical bacterium [10], but only one tenth of the annotated 1,500 RBPs in the human genome [11, 12] seems to suggest a lesser role of bacterial RBPs in functions outside pure protein synthesis. Furthermore, bacteria were traditionally thought to control their genes almost exclusively at the level of transcription, whereas extensive post-transcriptional processes such as mRNA splicing involving dozens if not hundreds of different RBPs would be seen as a hallmark of eukaryotic gene expression [13].

Things have changed. The past two decades have shown extensive post-transcriptional control in bacteria, which includes large regulatory networks built of RBPs and small noncoding RNAs (sRNA) [G] [14]. We have also achieved a much better understanding of how RBPs act on nascent bacterial transcripts, modulate the translation rate of mRNAs, or selectively alter the decay rates of specific sRNAs [15] (FIG. 1d). New profiling methods using RNA-seq [G] are providing evidence that bacterial RBPs can act as truly global factors that directly associate with and influence the fate of several hundreds of transcripts from across the genome [16]. At the same time, molecular studies of CRISPR/Cas systems [G] have uncovered bacterial ribonucleoprotein particles (RNPs) [G] of unexpected structural and functional complexity [17]. Similarly, a recent census of cytosolic complexes suggests that many bacterial RBPs are part of larger molecular assemblies [10] (FIG. 1e). In light of these developments, it is timely to take stock of the molecular mechanisms and functions of major RBPs in bacterial gene expression.

This review will follow the path from transcription to protein synthesis, ultimately ending in RNA decay, and highlight the roles of specific RBPs in each process with an emphasis on their molecular properties (FIG. 1d). We will focus on recent findings regarding the molecular interactions between proteins and RNA, and describe emerging technologies for the global characterization of RBPs and their cellular ligands. Despite being proteins that interact with RNA, ribonucleases [G] and RNA-modification proteins [G] are often not referred to as RBPs. Readers interested in these types of enzymes are referred to a number of recent specialized reviews [18-20]. Similarly, the molecular properties of Cas proteins have recently been discussed elsewhere [17].

Regulation of transcription

During RNA synthesis, numerous proteins – in addition to RNA polymerase (RNAP) itself – make contacts with the nascent transcript. Most of them modulate the termination phase of transcription, either by promoting the release of RNAP or by preventing premature termination (FIG. 2a-d). Bacteria use two general mechanisms for transcription termination: intrinsic termination [G] and Rho-dependent termination [21] (FIG. 2a). Intrinsic termination, also called Rho-independent termination, occurs as RNAP encounters a palindromic GC-rich DNA sequence followed by a stretch of T residues. In the other mechanism, the release of the transcribing RNAP from the DNA template is brought about by the protein Rho [G].

Rho is amongst the best-studied RBPs in bacteria. It was first described ~50 years ago as a factor that promotes transcription termination in phage λ [22]. Following early observations that Rho also acted on host genes, this protein was linked to termination in >25% of all operons in *Escherichia coli* [23]. Rho forms large (~280 kDa) homohexameric rings (FIG. 2d), presenting two RNA-binding sites in each of the monomers. Termination initiates as the OB-fold of the primary RNA-binding site recognizes cytosine-rich r̲ho ut̲ilization (*rut*) sites in a cellular transcript. The full molecular mechanism of how Rho translocates along the RNA and finally halts transcription elongation is yet to be settled. Possibly, contacts by the secondary RNA-binding sites induce Rho ring closure and

stimulates ATP-dependent 5'→3' movement along the nascent RNA while still remaining tethered to the *rut* sequences [24]. Alternatively, Rho may associate with RNAP already at the start of the transcription cycle [25]. In any case, termination occurs as Rho encounters the RNA exit channel of the transcribing RNAP and releases the nascent transcript by unwinding the RNA-DNA duplex [26].

Induction of termination

For both induction and prevention of termination, RBPs may either act globally, or regulate only specific genes. Among those that induce termination globally, the well-studied Nus factors are of primary interest.

Originally discovered in *E. coli* as host factors of phage λ (see Text Box 1), Nus factor-like proteins have since been found in many other bacteria [27]. *E. coli* NusA is an essential protein with several different roles in termination: providing antitermination activity for phage transcription [28], while enhancing RNA-hairpin-dependent pausing and termination in chromosomal transcription [27]. In *Bacillus subtilis*, 25% of all termination events are ascribed to NusA, especially at sub-optimal intrinsic terminators [29]. Not only does this protein affect intrinsic termination: *E. coli* NusA may also globally antagonize Rho function [30].

A protein of many functions, NusA exemplifies a bacterial RBP with multiple functional domains: the N-terminal domain (NTD) that contacts RNAP is followed by one S1 and two KH domains with putative RNA-binding activity, and two acidic repeat domains that make contacts with RNAP's α subunit [31] (FIG. 2d). Several of these NusA domains have been implicated in RNA-binding; while the NTD interacts with RNA hairpins to induce transcriptional pausing [32], the two KH domains form a large interaction surface with rRNA antitermination sequences [33] (FIG. 2d).

Contrasting Rho's global effects, numerous RBPs induce termination at specific loci (FIG. 2b). A classic example is the *trp* RNA-binding attenuation protein (TRAP) in *B. subtilis*, a negative regulator of the tryptophan biosynthesis operon [34]. Functional TRAP protein is a ring of eleven subunits [35] (FIG. 2d), which

binds tryptophan in hydrophobic pockets between its monomers when this amino acid is plentiful. The activated TRAP then recognizes (G/U)AG repeats in the *trp* leader and terminates transcription upstream of the *trp* genes [36]. The RBP PyrR regulates the pyrimidine operon (*pyr*) in a similar fashion, and it also provides an excellent example of how an RBP structurally remodels an RNA leader sequence to abort transcription [37].

TRAP and PyrR both modulate intrinsic termination, but others act by modulating Rho-dependent termination. For instance, CsrA binding to leader sequences can expose *rut* sites that then mediate premature termination of transcription via recruting Rho [38].

Antitermination

RNA structure strongly influences the rate of transcription elongation. RNA hairpins that invade the RNAP exit channel induce pausing, the stronger ones termination. Antiterminator RBPs either directly prevent the formation of terminator hairpins, or favour the formation of alternative structures (FIG. 2c). They do this to generally suppress aberrant pre-mature termination events or to regulate gene expression as part of a physiological response.

Cold-shock proteins (CSPs) have been of much interest as a potentially major source of global anti-termination activity. These are small ($\approx$7 kDa) proteins of a ubiquitous family that also includes eukaryotic Y-box proteins, and which interact with single-stranded nucleic acids through their shared cold-shock domain [2]. Purified CSPs decrease termination and pausing *in vitro*, and their overexpression *in vivo* increases the expression of genes preceded by several intrinsic terminators [39].

The structure of the family's founding member, the cold-induced CspA of *E. coli*, revealed five antiparallel β-strands with an overall negatively charged surface [40, 41]. Positively charged amino acid residues of the conserved RNA-binding motifs RNP-1 and RNP-2 mediate RNA binding, as shown for CspB, a CspA homolog in *B. subtilis* [42] (FIG. 2d). Interestingly, structural similarity (and

functional redundancy) between the CSPs and S1 domains hint at a common ancestry from an ancient RBP [43].

While the full scope of CspA activity remains to be described, CspC and CspE are now established as truly globally acting RBPs [44, 45]. Their combined activities impact the expression of 20% of all genes in *Salmonella* and are essential for the virulence of this pathogen [44]. It is unclear, however, how many of these effects involve antitermination, as CSPs can also modulate translation initiation and RNA stability [44, 46, 47] (see below).

Many mechanistic aspects of the CSPs are unclear, including the molecular mechanism of antitermination itself, except for their melting activity on RNA secondary structures [39]. Strikingly, the affinity for RNA targets is typically in the micromolar range [44, 48], whereas many other RBPs described here bind with nanomolar affinities. *In vitro* selection has identified putative consensus sequences for several *E. coli* CSPs [48], but how RNA targets are recognized *in vivo* remains unclear. It will be important to elucidate how these very similar RBPs, of which some enteric bacteria have no fewer than eleven, select targets in a seemingly non-redundant manner *in vivo* [44].

While the CSPs represent global factors, other RBPs inhibit termination only at specific loci, often as part of a specific physiological response [21]. The hexameric antiterminator protein HutP acts on the histidine utilization genes in *B. subtilis*. Its RNA-binding activity is activated by a conformational change that is triggered by sensing the intracellular concentration of L-histidine [49]. By binding to six NAG triplets between its own gene and the downstream *hut* operon, HutP directly prevents the formation of the terminator hairpin to promote transcription elongation [50]. By contrast, RBPs of the BglG/SacY family bind inverted repeat sequences in mRNA leaders to induce RNA structures that are mutually exclusive with terminator formation [51]. Their target genes encode carbohydrate-metabolizing enzymes, and their RNA-binding activity is regulated through sugar availability-dependent phosphorylation.

The latest additions to antiterminator RBPs include LoaP, a Nus-like protein that promotes readthrough of intrinsic terminators in antibiotic gene clusters in Firmicutes [52]. Furthermore, two studies [53, 54] have highlighted how

bacteria integrate RBPs with other regulatory molecules to control antitermination. In *Enterococcus faecalis* and *Listeria monocytogenes*, the phosphorylated form of the RBP EutV binds to hairpin structures in the *eut* RNAs, thereby causing antitermination and thus expression of the ethanolamine utilization (*eut*) locus. Active EutV can be out-titrated by an sRNA that in turn is controlled by a riboswitch responding to adenosylcobalamin, a cofactor for ethanolamine catabolism.

In conclusion, RBPs engage in regulating transcription termination by promoting or inhibiting the formation of intrinsic terminators, or by affecting the availability of nascent transcripts for termination factor Rho.

Regulation of translation

Mechanistically, most RBPs that affect translation operate at the level of translation initiation (FIG. 3a-b), i.e. the process of 30S ribosomal subunit association to the mRNA at the ribosome binding site (RBS) [G]. RBPs may directly compete with the 30S for binding to an mRNA—as do Csr/Rsm proteins on hundreds of different mRNAs—or induce a structural change in the mRNA that alters the accessibility of the RBS. Another major mechanism whereby bacteria control mRNA translation involves RBP-mediated recruitment of sRNAs, the primary example of which is the regulatory network formed by the Hfq protein. Moreover, recent reports suggest that CSPs and FinO-like RBPs also regulate mRNAs globally [44, 55, 56](Holmqvist et al., in revision).

There are also many translation-related RBPs that control their own mRNA. The r-protein S1 as an RBP that both autoregulates its own and affects global translation will be discussed at the end of this section.

Csr/Rsm

Csr/Rsm proteins are highly conserved RBPs in the ~7 kDa range; predicted homologs of *E. coli* CsrA—the founding member of this class—are found in almost all bacterial phyla [57]. Their physiological importance reveals itself by the usually strong pleiotropic phenotypes if not lethality upon genetic inactivation.

For example, CsrA (<u>C</u>arbon <u>s</u>torage <u>r</u>egulator <u>A</u>) was discovered through its major impact on glycogen production and carbon storage in *E. coli* [58]; the former results from direct mRNA repression of the first gene in a glycogen biosynthesis operon [59] in which a dimer of CsrA recognizes two single-stranded GGA triplets in the *glgC* 5'UTR to obstruct 30S binding [59, 60].

Csr/Rsm proteins are truly global regulators that target several hundreds of mRNAs in some Gram-negative model species [61-63]. In addition to the canonical mechanism in which binding in 5'UTRs inhibits translation [61, 62, 64, 65], regulatory mechanisms of Csr/Rsm affecting transcription or RNA decay have been described [38, 66].

CLIP-seq studies capturing hundreds of CsrA-RNA interactions inside live *E. coli* and *Salmonella* [61, 62] have helped to refine the protein's recognition motif as AUGGA, preferably located in the apical loops of RNA hairpins (FIG. 3c). This motif is not only found in mRNAs but also abounds in a particular set of sRNAs that act as decoys of Csr/Rsm proteins [59, 67]. For example, the 369 nucleotide CsrB sRNA contains no fewer than 22 GGA sequences, of which 18 are high-affinity CsrA binding sites [67]. One such sponge, the RsmZ sRNA of *Pseudomonas fluorescens*, was structurally analyzed and revealed how the RBP RsmE binds RNA in an ordered, sequential, and cooperative fashion. Specifically, sequential binding of RsmE to RsmZ induces allosteric changes in the RNA which strongly enhance the affinity for subsequent RsmE binding events. The resulting highly stable RNP of five RsmE dimers not only sequesters the protein but also protects the RsmZ sRNA from nucleolytic degradation [68]. It is easy to envision the same process in the 5'UTRs of target mRNAs leading to stable complexes that fully inhibit translation initiation.

While studies in Gram-negative bacteria revealed intricate regulatory circuits composed of Csr/Rsm proteins and antagonistic sRNAs [69], these proteins can also be regulated by protein-protein interactions. For example, the protein CesT in enteropathogenic *E. coli* not only functions as a cytosolic chaperone of secreted virulence effectors, but also moonlights as a direct inhibitor of CsrA to lift the CsrA-dependent repression of effector mRNAs [70]. Likewise, CsrA

interactions with the Hag and FliW proteins precisely time the expression and secretion of the flagellin Hag in *B. subtilis* [71, 72].

Back to RNA binding, the recognition mode of Csr/Rsm proteins seems highly conserved because homologs from distant species can often complement each other's function *in vivo* [73]. However, yet other bacteria express multiple Csr/Rsm homologs with partly non-redundant activities [74]. One open question is whether there may be auxiliary proteins that modulate the affinity of these RBPs for certain mRNAs, either directly or by changing the subcellular localization of a target [75]. The growing atlases of *in vivo* binding sites [61, 62] should assist in determining whether the Csr/Rsm proteins act on certain targets in a combinatorial fashion with any of the other major RBPs described below.

Hfq

If bacteria needed a poster child for RBPs that facilitate global RNA networks of eukaryotic-like complexity, the choice would be Hfq. This protein's history, from its discovery as a host factor that unwinds phage RNA for efficient replication, to its increasing implication in endogenous gene control, has been reviewed [76]. To date, the primary role of Hfq is seen in aiding sRNAs finding and regulating *trans-encoded* **[G]** mRNAs through short, imperfect base pairing interactions. In *E. coli* and *Salmonella*, up to one hundred sRNAs may recognize more than 25% of all mRNAs via Hfq, comprising thousands of potential RNA interactions [61, 77-79]. In the current standard model, Hfq simultaneously binds sRNAs and mRNAs and, if these possess sufficient complementarity, an sRNA-mRNA complex forms. The resulting RNA duplex usually sequesters the RBS, silencing translation of the mRNA; however, other modes of inhibition as well as various mechanisms of target activation are known [14, 80, 81].

How does Hfq work as a global RNA matchmaker? Hfq possesses a conserved Sm domain—an N-terminal core that harbors two sequence motifs named Sm1 and Sm2—and a variable C terminus. To be functional, six Hfq protomers assemble into the typical ring-like architecture of the Hfq–Sm–LSm family found in both prokaryotes and eukaryotes [76]. The homohexameric Hfq ring binds single-stranded RNA at four different sites: the proximal and the distal

face, the rim and the C-terminal tail. Although we still lack a high-resolution structure of full-length Hfq in complex with two paired RNAs, the results of different *in vitro* and *in vivo* analysis have led to a model of how regulation is brought about (FIG. 3c).

The proximal face of Hfq preferentially interacts with single-stranded U-rich sequences at sRNA 3' ends [82,83]; additional rim interactions help to position the sRNA such that its seed region is poised for contacting mRNA targets [82]. By contrast, the distal face interacts with single-stranded A-rich sequences present in mRNA 5'UTRs [84-86], bringing potential mRNA target sites close to the sRNA-binding proximal face [87]. When Hfq is loaded with both sRNA and cognate mRNA, the rim contacts UA-rich sequences in them to promote RNA pairing [88]. While this is the canonical view of how Hfq contacts sRNAs and mRNAs, the enormous sequence diversity of RNA partners predicts case-to-case variations. Indeed, mutational studies suggest two major classes of Hfq-associated sRNAs that differ by their dependencies for face or rim contacts [89].

The function of the fourth major site on Hfq, the poorly-conserved and intrinsically disordered C-terminus, has been a matter of controversy [90,91]. However, recent data suggest that it helps to displace bound transcripts which may not only help to rid Hfq of non-specific RNA binders [92], but also promote a rapid cycling of cellular RNA whose total concentration exceeds that of Hfq [93]. One can imagine how the six flexible carboxy-termini protruding from the ring brush RNA off the core. In addition, their acidic tip may transiently bind the basic Sm core residues necessary for RNA annealing, thereby auto-regulating RNA binding to the Sm ring [94].

Following the pioneering structure analysis of *Staphylococcus aureus* Hfq [85], much of the above model has been pieced together from the *E. coli* and *Salmonella* proteins whose hexamers are ~70 kDa. However, the Hfq family is diverse, containing proteins with predicted hexamers of up to 140 kDa [95]. There are more reasons to suspect that we are yet to understand the true functional diversity of these proteins. On the one hand, heterologous expression of diverse Hfq proteins in *Salmonella* showed that members of this family have an intrinsic propensity to bind sRNAs [96]. On the other hand, the collective results of Hfq

studies in Gram-positive bacteria so far argue against the prominent sRNA-mRNA matchmaking function that is a hallmark in Gram-negative bacteria [97].

Even in Gram-negative species, Hfq can modulate mRNA translation by sRNA-independent mechanisms, the latest example being from the *E. coli* mismatch repair pathway [98]. Moreover, studies of how *Pseudomonas* Hfq directly represses translation of mRNAs under conditions of catabolite repression have suggested a direct interaction with the Crc protein [99]. Crc was initially reported to inhibit translation by binding A-rich sequences in mRNAs required for growth on non-preferred carbon sources; this activity would be counteracted by the CrcY/Z sRNAs [100]. Yet, more recent work suggests that the translation inhibitor is Hfq, whereas Crc acts to stimulate the mRNA-binding activity of Hfq [99, 101]. Crc and Hfq form a complex when RNA is bound to the distal side of Hfq, which results in prolonged life-time of Hfq-RNA interactions [99]. How Crc turns Hfq into a better mRNA repressor is not yet understood but the finding itself raises the possibility that other proteins exist that may guide Hfq to mRNA targets in an sRNA-independent fashion.

RBPs of the ProQ/FinO family

While Hfq and CsrA have well-established roles in translational control, those of ProQ and other FinO-domain containing RBPs are just emerging. The founding member of this class, FinO, has been studied for its function as an RNA chaperone in antisense regulation of F plasmid conjugation in *E. coli* [102].

The ~25 kDa ProQ protein is a chromosomal homolog of FinO. However, whereas FinO has only two known RNA targets, ProQ stably associates with several hundreds of cellular transcripts in *Salmonella* and *E. coli* [10] (Holmqvist et al., in revision). These ProQ targets include >70 sRNAs, one of which base pairs with the RBS of the mRNA of histone-like protein HU [56]. Not only does ProQ stabilize this sRNA, it also accelerates RNA duplex formation [56], a general property that had already been seen with artificial substrates [103]. RocC, a FinO-like protein expressed from the chromosome of *Legionella pneumophila*, also mediates translational repression, helping the RocR sRNA to bind multiple mRNAs of this organism's DNA uptake system [55].

While these examples seem to reiterate the functions of Hfq, what happens after target recognition may differ: whereas Hfq, being at limiting concentration, quickly releases a matched sRNA-mRNA couple, ProQ seems to hold onto sRNA-mRNA duplexes for a longer time as if assuring that 30S is always excluded [56]. Cellular abundance argues for this scenario: there are 5-10 times more ProQ monomers (the active form) than Hfq hexamers [10], albeit the number of RNA targets is similar [61](Holmqvist et al., in revision).

Potential FinO homologues are found in the chromosomes and on plasmids of α-, β-, and γ-proteobacteria, and Acidithiobacilli [10, 55], promising a rich diversity of physiological functions and molecular mechanisms. Moreover, results from evolutionarily distant bacteria such as *E. coli* and *Salmonella* [61](Holmqvist et al., in revision), *Vibrio cholera* (K. Papenfort, unpublished) and *Neisseria meningitidis* (J. Vogel, unpublished) have consistently shown that the cellular RNA target suites of CsrA, Hfq and ProQ are distinct. Intriguingly, whereas both Hfq and CsrA target defined sequence motifs in single-stranded RNA, FinO and its siblings may preferentially recognize targets by RNA structure. For example, CLIP-seq analysis in *E. coli* and *Salmonella* predicts ProQ to typically recognize a stable RNA hairpin, often a transcription terminator (Holmqvist et al, in revision). Similarly, FinO and RocC bind to the terminator hairpins of the FinP and RocR sRNAs, respectively [55, 102]. In other words, ProQ appears to govern its global network of sRNAs and mRNAs by reading RNA structure rather than sequence. At the same time, it successfully discriminates against the abundant and hairpin-rich rRNAs and tRNAs.

Structural analysis of FinO-like RBPs is in its infancy (FIG. 3c) but it has great potential for cracking their structural recognition code. Such analyses may also help explaining why some members such as FinO and RocC have only few cellular RNA targets whereas ProQ binds hundreds, although all these RBPs carry the same FinO-domain. Is their selectivity determined by their distinct N- and/or C-terminal extensions [104, 105]? A structural analysis of *E. coli* ProQ indicates that in addition to the FinO-domain, other regions contribute to RNA binding and so may modulate target selectivity [105]. For *E. coli* ProQ, there is another open question: this protein was stumbled upon as being required for full expression of

the proline transporter ProP (hence its name) [106] but the underlying mechanism remains unsolved.

The rich world of r-proteins

The largest functional class of bacterial RBPs are those involved in protein synthesis, such as r-proteins, ribosome-associated proteins, tRNA synthetases and enzymes that modify tRNA and/or rRNA. There are 57 bacterial r-proteins, 34 of which are conserved in all domains of life [1]. Both the biogenesis of ribosomes and the ribosome-mediated process of translation are highly complex and dynamic processes that involve countless transient and stable interactions of r-proteins with RNA. As such, the rapidly growing information on ribosome assembly pathways [107] and molecular structures [108] should provide a treasure trove for the understanding of RBP functions. Upon binding, an r-protein may induce a local structural change in the rRNA for other proteins to associate [109]. Interestingly, however, r-proteins also help to move the mRNA within the ribosome and secure translational accuracy through selection of the correct aminoacyl-tRNAs [110].

One interesting aspect of r-proteins involves their general use of positively charged residues, rather than of defined amino acids, to specifically recognize the shape and charge of the rRNA backbone [109]. Many of these proteins possess globular domains with classical RBDs that are extended by elongated tails and internal loops to allow for contacts with distinct, sometimes multiple rRNA regions [111]. However, another common RNA target has helped to better understand the specificity of r-proteins: their own mRNA [112].

As a substantial fraction of cellular resources is devoted to produce ribosomes, the synthesis of each r-protein is controlled to match the stoichiometry of the other ribosome constituents. In most cases, this is achieved through auto-inhibition at the level of translation initiation. Auto-inhibition by r-proteins usually involves molecular mimicry; the mRNA presents a binding motif akin to the cognate site in the rRNA [113]. A difference in the affinity of the two sites ensures that r-protein mediated auto-repression only occurs when rRNA synthesis is reduced, for example under starvation. Mechanistically, the

molecular mimicry can entrap the ribosome, as shown for r-protein S15 [114], or outcompete the 30S subunit [115].

The r-protein S1 generally promotes protein synthesis by unfolding structured mRNAs to allow for 30S docking and optimal positioning of the start codon [116]. For autoregulation, it uses a mechanism distinct from the above: its C-terminal region anchors S1 on the ribosome via protein-protein interactions, while its N-terminal region binds AU-rich RNA sequences. When S1 is in molar excess over ribosomes, free S1 binds its own mRNA to inhibit translation [117].

Autoinhibition via RNA binding has become a recurrent theme for bacterial RBPs, well beyond r-proteins. Other well-studied examples are the translation initiation factor IF-3 which suppresses translation initiation at the non-canonical AUU start codon of its own mRNA [118], and threonyl-tRNA synthetase (ThrRS) that translationally represses its synthesis by binding to its mRNA at a 5'UTR-located RNA motif that structurally mimics its substrate, tRNA^{Thr} [119]. Moreover, several nucleases cleave their own mRNA [120, 121].

Other RBPs from translation processes

Proper ribosome biogenesis and function relies upon several additional well-conserved RBPs. Of several ribosome maturation factors, the GTPase RsgA uses its OB-fold to interact with the 16S rRNA to suppress the formation of kinetically trapped 30S intermediates [122]. Turning to ribosome function, the RBP Ffh together with 4.5S RNA constitutes the bacterial signal recognition particle (SRP), the RNP that delivers translating ribosomes to the membrane. Structural analyses of SRP have revealed an unexpected arrangement in which a methionine-rich (M) domain of Ffh holds on to one end of 4.5S RNA, whereas another domain contacts the opposite end [123]. Considering their structural robustness, protein synthesis-related RBPs will continue to provide intricate details of RNA-protein interactions, in particular, as cryogenic electron microscopy (cryoEM) has begun to tackle higher-order complexes of bacterial gene expression [124].

RNA turnover

Bacterial RNA turnover is generally sequential: transcripts are first attacked by single-strand-specific endoribonucleases such as RNase E or RNase Y, or the double-strand-specific endoribonuclease RNase III, before decay is completed by exo- and oligoribonucleases [20]. These activities can be combined in multi-protein RNA degradation complexes called degradosomes [G] which are found in many bacterial species.

For RBPs that regulate the stability of cellular transcripts, major mechanisms include the direct competition with RNases for the same site, the RBP-mediated recruitment of RNases to specific RNA ligands, and RBP-dependent positioning of RNase cleavage sites (FIG. 4) [20]. Regarding CsrA, Hfq, and ProQ, we will focus on direct mechanisms where changes in RNA stability are more than a consequence of altered translation rates.

RNA decay

No matter how degradosomes from different bacteria may vary in composition, all contain members of the DEAD-box family of RNA helicases [125]. RNA helicases are RBPs that unwind RNA duplexes powered by ATP hydrolysis [126]. Because most RNases act on single-stranded RNA, the activity of an RNA helicase usually promotes RNA degradation.

The helicase in the *E. coli* degradosome is RhlB [127]. Stimulated by direct contact with RNase E, RhlB unwinds RNA secondary structures to allow for full decay by the exoribonuclease PNPase, another degradosome component [128]. CshA is the corresponding DEAD-box protein in the degradosomes of Gram-positive species. Interestingly, work in *S. aureus* suggests a more specific function of CshA in accelerating the degradation of mRNAs from the *agr* quorum-sensing system to generally promote biofilm formation [129].

Another way for RBPs to promote RNA decay is the active recruitment of nucleases. For example, it was proposed that Hfq-associated sRNAs, as they base pair with mRNAs to inhibit translation, also actively recruit RNase E to render silencing irreversible. Recruitment would involve the formation of an sRNA-Hfq-

RNase E complex that is distinct from the degradosome [130]. Such a triple complex in which the RBP Hfq recruits RNase E is yet to be proven structurally; in fact, others have argued that the sRNA rather than the RBP recruits RNase E [131]. Regardless of the nature of this elusive complex, there is evidence that Hfq and sRNAs can guide RNase E to cleave mRNAs, even in the absence of translational inhibition [132, 133].

A well-characterized example of RBP-mediated nuclease recruitment is RapZ (a.k.a. YhbJ). Originally identified through a phenotype in amino sugar metabolism, the 32 kDa RapZ protein of *E. coli* was shown to selectively bind and present the GlmZ sRNA to RNase E for inactivation [134]. Structural analysis has confirmed that RapZ directly binds to RNase E, and also indicated a molecular origin of this RBP from metabolic enzymes [135]. The Csr pathway contributes another putative adaptor protein for specific RNA decay: the ~73 kDa CsrD protein stimulates CsrB decay by counteracting CsrA binding to an RNase E site [136]. Acting on a more global level, a bacterial homolog of Ro (a well-studied eukaryotic RBP) was recently shown to promote the degradation of highly structured RNAs in *Deinococcus radiodurans*. This RBP is tethered by the noncoding Y RNAs to PNPase, yielding a larger RNP that displays different substrate specificity than PNPase alone [137]. Together, these examples provide molecular models for a better understanding of how RBPs contribute to shaping the selectivity of RNA decay.

RNA stabilization and processing

Two major principles of how RBPs positively affect transcript stability have emerged: direct competition with an endoribonuclease for an internal cleavage site, and competition with an exonuclease activity at the transcript termini. The former has been well-studied for Hfq-associated RNAs which are stabilized by Hfq occupying an internal RNase E recognition site [138]. A comparison of recent global maps of Hfq occupancy [61] and RNase E binding sites [139] in *Salmonella* indicates Hfq-mediated protection to be common amongst sRNAs. Moreover, by suppressing certain RNase E sites in precursor transcripts, Hfq guides RNase E to the correct processing site in some sRNAs [139]. This role of Hfq in processing

seems to be particularly important for those sRNAs that are derived from mRNA 3'UTRs [139, 140]. By contrast, Hfq-mediated protection of mRNAs against RNase E largely depends on associated sRNAs that sequester critical cleavage sites in the target, or a higher ribosome density due to sRNA-induced translational activation [81]. YbeY, an RNase that carries a putative MID domain of Argonaute proteins, has also been predicted to have links with Hfq- and/ or sRNA-dependent RNA decay [141].

That major RBPs directly protect mRNAs against endonuclease activity was demonstrated with CsrA [66] and some CSPs [44]. CsrA shields RNase E cleavage sites not only in its decoy, the CsrB sRNA [136], but also in the mRNA of the *E. coli* master activator of motility genes, FhlDC [66]. Interestingly, one of the two CsrA sites lies at the very 5' end of the *flhDC* mRNA [66], raising the possibility that CsrA blocks RNase E access at the earliest possible stage (for RNase E, 5' initiated decay is faster than internal entry [18]). One promising example is the short *ecnB* transcript which possesses twelve RNase E sites that all cluster in its 5' end [139]. Complementing *in vivo* evidence, the *ecnB* mRNA fully resists RNase E attack when pre-incubated with recombinant CspC or CspE proteins *in vitro* [44].

Although endonucleolytic cleavage is the rate-limiting step in bacterial RNA turnover, exoribonucleolytic activity can influence RNA decay rates, too. Consequently, RBPs can exert stabilizing effects at transcript termini. Hfq does so by binding to RNA 3' ends where it antagonizes exonuclease activity or stimulates polyadenylation [142]. Examples include sRNAs whose half-lives decrease upon loss of Hfq but are restored when PNPase is simultaneously inactivated [143]. CspC and CspE may also have a general effect on RNA degradation as they can counteract PNPase activity [46]. Adding to this list, a fine-mapping of ProQ sites in *E. coli* and *Salmonella* revealed massive binding of this RBP to mRNA 3' ends. Genetic evidence suggests that ProQ stabilizes at least one of these targets, the *cspE* mRNA, against 3'→5' degradation by yet another major exoribonuclease, RNase II (Holmqvist et al, in revision).

Collectively, these examples highlight how RBPs may directly interfere with RNases at both internal and terminal transcript positions. Considering that nucleases may not be evenly distributed in a bacterial cell—for example, RNase E

accumulates at the inner membrane [144]—one may speculate about additional mechanisms of modulating RNA decay, such as an RBP-mediated relocalization of specific transcripts into or out of zones of high nuclease activity.

Emerging aspects of bacterial RBPs

Bifunctional RBPs

The recent global RBP discovery studies in eukaryotes predicted an unexpected number of new RBPs of so far unknown function, many even lacking identifiable RBDs. Equally surprising, however, was that many proteins with established functions, especially metabolic enzymes, appear to moonlight as RBPs [11]. Although analogous global RBP screens are missing in bacteria, several observations suggest that bifunctional RBPs may be present there as well.

Of metabolic enzymes, the iron-containing protein aconitase, which converts citrate to isocitrate in the TCA cycle, also acts as a post-transcriptional regulator. In *B. subtilis*, this non-metabolic activity regulates important physiological processes such as sporulation [145]. Both bacterial and eukaryotic aconitases tend to auto-inhibit their expression via their RNA-binding activity when intracellular iron becomes scarce. Intriguingly, *E. coli* aconitase binds in the 3' region of its own mRNA, thereby masking an RNase E cleavage site that is essential for sRNA-induced degradation [146]. Other good candidates for bifunctional RBPs include the cytoskeletal protein RodZ that promotes mRNA decay of a virulence regulator in *Shigella sonnei* [147], the translocator protein YopD that also inhibits the translation of secreted virulence factors in *Yersinia* species [148], and the (p)ppGpp synthetase RelQ of *Enterococcus faecalis* whose enzymatic activity is inhibited by single-stranded RNA [149].

Some global DNA-binding proteins such as HU, HN-S, and StpA also bind RNA *in vivo*. While HU acts as a chaperone for DNA condensation through interacting with noncoding RNA species [150], HN-S and StpA alter the half-lives of the sRNAs DsrA and MicF, respectively [151, 152]. As of yet, none of the above proteins has had its entire spectrum of potential RNA ligands profiled. Thus, approaching these proteins with now available RNA-seq methodology [16] may

yield unexpected new targets and broaden our knowledge about how bacterial proteins combine enzymatic or structural functions with RNA-binding regulatory activities.

Cellular localisation of RBPs

Thanks to good progress in bacterial cell biology, it is now clear that even these "simple" organisms show ample spatial and temporal organization of their proteome. Likewise, examples of RNA molecules with particular subcellular localization keep accumulating [153]. A new RNP with membrane localization was recently shown to confer resistance to alcohol- and cold-induced stress in certain Firmicutes. The RNP contains the abundant noncoding RNA OLE that is anchored to the inner membrane through a direct interaction with OLA, an RBP that has a trans-membrane domain [154]. Other recently described examples of RBPs implicated in RNA localization include the aforementioned RodZ protein in *Shigella* [147], and CsrA in *Campylobacter jejuni* in which it counteracts FliW-dependent localization of flagellin mRNA [75].

These recent findings notwithstanding, intracellular localization remains an understudied aspect of bacterial RBP biology. As to available resources, the ASKA collection containing all *E. coli* ORFs as GFP tagged variants lends itself to obtain a bird's eyes view of the cellular localization of an RBP of interest [155]. However, such results need to be interpreted with caution because of mislocalization issues caused by the presence of fluorescent tags, as demonstrated for Hfq [156].

RNA-protein complexes

Clear differences between bacterial and eukaryotic RBPs lie in their overall propensity to form larger complexes. Eukaryotic RBPs often participate in higher-order, often dynamic, complexes such as the spliceosome or the RNA-induced silencing complexes [12]. Except for the ribosome and SRP, none of these complexes are conserved in bacteria. Conversely, bacteria contain several larger complexes that are absent from eukaryotes, for example, the degradosome [18],

several CRISPR-Cas related RNPs, and highly specialized complexes formed by certain toxin-antitoxin systems [157].

If and how many other bacterial RNPs are out there is currently difficult to predict. A recent proteome analysis of glycerol gradient fractions in *Salmonella* [10] showed numerous proteins with RNA-related functions migrate faster than predicted by their molecular weight (FIG. 5). The extent to which these putative complexes contain other proteins, or are only driven by the molecular weight of associated RNA, remains to be determined. Other resources yet to be fully explored for the discovery of additional higher-order RNPs are the many global protein-protein interaction data that have become available for different bacteria over the past two decades. While the latest one for *E. coli* [158] predicted no protein partners of Hfq and only three for CsrA, ProQ was associated with an astonishing forty non-ribosomal proteins. While these numbers should be treated with caution, the presence of higher-order ProQ-containing complexes is supported by the aforementioned gradient profiling [10].

OUTLOOK

This review aimed to provide a high-level view of where and how bacterial RBPs function on the levels of transcription, translation and RNA decay. As to the true number of RBPs in bacteria, technical obstacles have hampered the transfer of game-changing *in vivo* crosslinking and purification technology for genome-wide cataloguing of eukaryotic RBPs [11]. However, it may only be a matter of time, before bacteriologists devise global screens of similar success. Already, the results of the eukaryotic screens have provided hints that more RBPs await to be discovered in bacteria as well. By contrast, almost all of the RBPs covered here were stumbled upon in general phenotypic analyses rather than having been identified in dedicated searches for new RBPs.

New RBPs may carry out fully unexpected functions, as illustrated by those of CRISPR-Cas systems that few bacteriologists had on their radar. Once understood, they can inform the design of artificial RBPs for synthetic types of gene regulation. Moreover, our general approach taken here for the sake of brevity naturally neglected that "the bacteria" are in fact a very large and diverse

group of organisms. Inasmuch as Jacques Monod's quote that "What is true for *E. coli* is true for the elephant" has amply been proven too simple, we have seen plenty of evidence suggesting that the major RBP functions described here on the basis of work in very few model species, predominantly γ-proteobacteria, may not be representative in the many poorly characterized, often unculturable bacteria from the human microbiota and environmental communities.

The growing information about which RBPs are expressed when, and about their cellular RNA ligands, now enables a systems-level understanding of how these regulatory proteins function. Integrating such data with large-scale functional and phenotypic screens [159] will certainly be a fruitful approach to further understand RBP functionality. For instance, comparing phenotypes of RBP mutant strains with RBP binding site data may reveal direct links between binding events and physiological consequences. Such comparisons would highly benefit from the development of databases that allow comparisons of large-scale datasets in a user-friendly manner.

TEXT BOX 1: RBPs expressed from plasmids and phages

Plasmids and bacteriophages often use RBPs to regulate gene expression, and many early discoveries of RBP-mediated regulation were made in these systems. Phages λ and Mu revealed RBP-based positive regulation [22, 160], whereas examples of negative regulation came from bacteriophages MS2 and T4 [161, 162]. Protein N from phage λ was the first antiterminator protein to be discovered [22], which together with Nus factors A, B, C, and G, and RNAP forms a transcription antitermination complex (TAC) [28]. The structural basis for this form of antitermination was recently described [163]. As the elongating TAC encounters an intrinsic terminator, the S1 domain of NusA prevents terminator formation by sequestering the upstream arm of the terminator hairpin. The RNA-binding coat protein of MS2 has a dual role in genome encapsidation and in translation repression. The high affinity of this RBP for its RNA ligand has been utilized for RNA-protein capture techniques in bacteria [10, 164].

Plasmids use RBPs to regulate processes such as DNA replication and conjugation. On F plasmids, the antisense RNA FinP controls the synthesis of the positive regulator of conjugation, TraJ. This regulation requires the RBP FinO which both stabilizes FinP and aids its annealing to the *traJ* mRNA [102]. FinO-like proteins have recently garnered much attention because of the unexpected functions of family members from bacterial chromosomes (see section *Regulation of translation*) [55, 165]. Other plasmid-expressed RBPs may have chromosomal homologs with a broader physiological function than currently thought. The Rop/Rom proteins which facilitate antisense regulation on plasmid ColE1 seem good candidates to start with [165].

TEXT BOX 2: Characterizing RBP interactomes (includes Figure 6: Methods to identify RBP ligands and binding sites)

For characterising the full cellular repertoire of RNA ligands of an RBP of interest, the analyses of co-immunoprecipitated RNA by microarrays (RIP-chip) or RNA seq (RIP-seq) have been popular [16] (FIG. 6). Pioneering applications in bacteria were directed towards understanding global functions of Hfq [166-168]. More recently, the CLIP-seq technique in which cellular RNA-protein interactions

are stabilized *in vivo*, prior to lysate preparation, by UV-crosslinking have gained momentum (FIG. 6). The UV-induced covalent bonds between protein and RNA not only permit purification under more stringent denaturing conditions, but they also often yield characteristic mutations in the final cDNA pinpointing the nucleotides in close proximity with an RPB's residues [169]; for example, frequent T→C transitions in Hfq sites [61]. Global CLIP-seq maps are now available for Hfq, CsrA, and ProQ [61, 62, 78](Holmqvist et al., in revision). Continuous development has resulted in improved CLIP-seq protocols that require less input material, permit more convenient labeling and more efficient library preparation schemes, the sum of which permits characterization of many RBPs in parallel [169].

In contrast to CLIP-seq, which aims at identifying RNA-protein interactions *in vivo*, RNA sequencing technology has also been exploited for high-throughput *in vitro* binding approaches (FIG. 6). Exposing RBPs to libraries of RNA molecules in a test tube or on a flow cell of a sequencer has been shown to inform on binding motifs and binding kinetics of eukaryotic RBPs [170-172]. Another approach is to combine systematic evolution of ligands by exponential enrichment (SELEX) with RNA-seq [173]. These methods provide a versatile analytical toolkit to understand bacterial RBP functions.

TEXT BOX 3: Identifying RBPs as a systems-level approach (includes Figure 7: Methods to identify RBP ligands and binding sites)

Several recent protocols for the identification of bacterial RBPs rely on co-purification with a tagged bait RNA (FIG. 7). Tagging an RNA of interest with an MS2 aptamer provides a robust strategy for co-purification of sRNA-binding RBPs [10, 164]. The StreptoTag aptamer has also been used [174]. Alternatively, cDNA baits capture specific RNAs and associated proteins from UV-crosslinked bacterial cultures [175].

In contrast to methods that identify protein partners of a single RNA, the Grad-seq method collectively probes the biochemical behaviour of the full soluble RNA and protein content of a bacterial cell [10]. By coupling classic glycerol gradient fractionation with mass spectrometry and RNA-seq, gradient migration

patterns can predict RNA-protein interactions or even larger complexes (FIG. 7). Grad-seq guided the discovery of ProQ as a global RBP in *Salmonella* [10].

RBP discovery in eukaryotes has been accelerated by interactome capture and similar methods (FIG. 7). UV-crosslinked proteins are co-purified via the poly(A)-tails of cellular transcripts and subsequently identified by mass spectrometry [11]. Because the abundant rRNAs and tRNAs lack poly(A)-tails, this strategy preferentially identifies mRNA-associated proteins, such as regulatory RBPs. For several reasons, these approaches are not directly applicable to bacteria: their RNAs lack extensive poly(A)-tails, which precludes the typical oligo(dT)-based purification. Without that, the high concentration of housekeeping RNAs such as rRNAs and tRNAs can be expected to preclude a sufficient capture of RBPs that sit on mRNAs or sRNAs. Bacteria-specific protocols may utilize the 5'PPP group of unprocessed mRNAs and sRNAs for initial transcript capture. In eukaryotes, RBP discovery has also successfully used whole proteome microarrays [176, 177]. The truly high-throughput character and its obvious applicability to bacteria notwithstanding, these screens may come with more false-positive interactions than *in vivo*-based assays.

FIGURE LEGENDS

Figure 1. **Overview of bacterial RBPs. a**) Venn diagram of all *E. coli* proteins classified as "RNA-binding" according to the Gene Ontology divided into the indicated functional categories. **b**) Representative examples of well-characterized bacterial RBPs to demonstrate their variability in length and number of RBDs. RBDs are represented as rounded boxes that are colour-coded as indicated in the box. "Misc. RBDs" indicate RBDs that do not belong to a major RBD family. CSD: cold-shock domain. **c**) Size distribution of all *E. coli* proteins with Gene Ontology term "RNA-binding". **d**) Cellular processes commonly regulated by bacterial RBPs. Well-characterized regulatory RBPs within each process are given by name.

Figure 2. **RBP-mediated regulation of transcription termination. a**) Bacteria use two different processes to terminate transcription: intrinsic termination and Rho-dependent termination. **b**) Regulatory mechanisms by which RBPs stimulate termination. **c**) Regulatory mechanisms of RBP-dependent antitermination. **d**) Structures of four major RBPs that control transcription termination: *E. coli* Rho bound to UUUUUUU RNA (PDB ID: 5JJI), *M. tuberculosis* NusA bound to GAACUCAAUAG RNA (PDB ID: 2ASB), *G. stearothermophilus* TRAP bound to a 53 nucleotide long single-stranded RNA (PDB ID: 1C9S), and *B. subtilis* CspB bound to UUUUUU RNA (PDB ID: 3PF5).

Figure 3. **RBP-based regulation of translation. a**) Major mechanisms by which bacterial RBPs mediate translation repression. **b**) RBP-dependent mechanisms of translation activation. **c**) Molecular structures and RNA recognition for CsrA, Hfq, and ProQ. Protein structures are based on PDB ID 2BTI (CsrA), PDB ID 2YLB (Hfq), and Figure 10 in (Gonzalez et al. 2017).

Figure 4. **RBP-dependent regulation of RNA decay.** Mechanisms by which RBPs activate or inhibit RNA degradation by endoribonucleases (**a**), or exoribonucleases (**b**), respectively.

Figure 5. **Distribution of RBPs as part of cellular complexes.** Heat map showing relative abundance of *Salmonella* RBPs in fractionated bacterial lysates based on data in (Smirnov et al, 2016). For each protein (rows), the fraction in which the protein was most abundant was set to unity. For the RBPs indicated with names and arrows, the heat map colour was changed to red for better visualization.

GLOSSARY

r-proteins

All proteins that together with the ribosomal RNA form the two subunits of the ribosome, the molecular machine that synthesizes polypeptides from an mRNA template.

sRNA

Bacterial, in most cases noncoding, RNAs in the size range of 50-300 nucleotides that carry out regulatory functions, either by base-pairing to complementary sequences in other RNAs, or by serving as binding partners for proteins.

RNA-seq

A method for massive parallel determination of sequence and abundance of RNA molecules in a biological sample.

CRISPR/Cas systems

Bacterial adaptive immune system consisting of short RNAs that guide Cas proteins to target invading nucleic acids for destruction.

RNP

Macromolecular complex consisting of RNA and RNA-binding proteins.

Ribonuclease

A class of protein enzymes that catalyze the cleavage of RNA. Endoribonucleases cleave RNA internally, while exoribonucleases degrade RNA from either the 5' or 3' end.

RNA-modification proteins

Proteins that introduce chemical modifications in RNA.

Intrinsic termination

A mechanism used by bacteria to stop transcription elongation and release the newly synthesized RNA. Intrinsic terminators are RNA motifs consisting of a stem-loop structure followed by several consecutive uracil residues.

Rho

A homohexameric bacterial protein that promotes termination of transcription elongation by the means of ATP hydrolysis.

RBS

The ribosome binding site refers to an mRNA sequence that recruits the 30S ribosomal subunit to initiate translation. The Shine-Dalgarno sequence of the RBS is complementary to the 16S rRNA and enables 30S-mRNA interaction.

***trans*-encoded**

A concept to describe the relation between two genetic elements. An sRNA is *trans*-encoded with respect to its mRNA target if the two are encoded by different genetic loci. In contrast, a *cis*-encoded sRNA is encoded by the same locus as its mRNA target, for instance when the respective genes overlap.

Degradosome

A multi-protein complex that carries out RNA degradation in bacteria.

REFERENCES

1. Fox, G.E. Origin and evolution of the ribosome. *Cold Spring Harb Perspect Biol* **2**, a003483 (2010).
2. Chaikam, V. & Karlson, D.T. Comparison of structure, function and regulation of plant cold shock domain proteins to bacterial and animal cold shock domain proteins. *BMB Rep* **43**, 1-8 (2010).
3. Hajnsdorf, E. & Boni, I.V. Multiple activities of RNA-binding proteins S1 and Hfq. *Biochimie* **94**, 1544-53 (2012).
4. Updegrove, T.B., Zhang, A. & Storz, G. Hfq: the flexible RNA matchmaker. *Curr Opin Microbiol* **30**, 133-8 (2016).
5. Koonin, E.V. & Makarova, K.S. CRISPR-Cas: evolution of an RNA-based adaptive immunity system in prokaryotes. *RNA Biol* **10**, 679-86 (2013).
6. Nicastro, G., Taylor, I.A. & Ramos, A. KH-RNA interactions: back in the groove. *Curr Opin Struct Biol* **30**, 63-70 (2015).
7. Masliah, G., Barraud, P. & Allain, F.H. RNA recognition by double-stranded RNA binding domains: a matter of shape and sequence. *Cell Mol Life Sci* **70**, 1875-95 (2013).
8. Swarts, D.C. et al. The evolutionary journey of Argonaute proteins. *Nat Struct Mol Biol* **21**, 743-53 (2014).
9. Helder, S., Blythe, A.J., Bond, C.S. & Mackay, J.P. Determinants of affinity and specificity in RNA-binding proteins. *Curr Opin Struct Biol* **38**, 83-91 (2016).
10. Smirnov, A. et al. Grad-seq guides the discovery of ProQ as a major small RNA-binding protein. *Proc Natl Acad Sci U S A* **113**, 11591-11596 (2016).
11. Hentze, M.W., Castello, A., Schwarzl, T. & Preiss, T. A brave new world of RNA-binding proteins. *Nat Rev Mol Cell Biol* (2018).
12. Gerstberger, S., Hafner, M. & Tuschl, T. A census of human RNA-binding proteins. *Nat Rev Genet* **15**, 829-45 (2014).
13. Morris, K.V. & Mattick, J.S. The rise of regulatory RNA. *Nat Rev Genet* **15**, 423-37 (2014).
14. Wagner, E.G.H. & Romby, P. Small RNAs in bacteria and archaea: who they are, what they do, and how they do it. *Adv Genet* **90**, 133-208 (2015).
15. Van Assche, E., Van Puyvelde, S., Vanderleyden, J. & Steenackers, H.P. RNA-binding proteins involved in post-transcriptional regulation in bacteria. *Front Microbiol* **6**, 141 (2015).
16. Hör, J., Gorski, S.A. & Vogel, J. Bacterial RNA Biology on a Genome Scale. *Mol Cell* (2018).
17. van der Oost, J., Westra, E.R., Jackson, R.N. & Wiedenheft, B. Unravelling the structural and mechanistic basis of CRISPR-Cas systems. *Nat Rev Microbiol* **12**, 479-92 (2014).
18. Hui, M.P., Foley, P.L. & Belasco, J.G. Messenger RNA degradation in bacterial cells. *Annu Rev Genet* **48**, 537-59 (2014).
19. Marbaniang, C.N. & Vogel, J. Emerging roles of RNA modifications in bacteria. *Curr Opin Microbiol* **30**, 50-7 (2016).
20. Mohanty, B.K. & Kushner, S.R. Regulation of mRNA Decay in Bacteria. *Annu Rev Microbiol* **70**, 25-44 (2016).
21. Ray-Soni, A., Bellecourt, M.J. & Landick, R. Mechanisms of Bacterial Transcription Termination: All Good Things Must End. *Annu Rev Biochem* **85**, 319-47 (2016).
22. Roberts, J.W. Termination factor for RNA synthesis. *Nature* **224**, 1168-74 (1969).
23. Cardinale, C.J. et al. Termination factor Rho and its cofactors NusA and NusG silence foreign DNA in E. coli. *Science* **320**, 935-8 (2008).
24. Koslover, D.J., Fazal, F.M., Mooney, R.A., Landick, R. & Block, S.M. Binding and translocation of termination factor rho studied at the single-molecule level. *J Mol Biol* **423**, 664-76 (2012).

25. Epshtein, V., Dutta, D., Wade, J. & Nudler, E. An allosteric mechanism of Rho-dependent transcription termination. *Nature* **463**, 245-9 (2010).

26. Mitra, P., Ghosh, G., Hafeezunnisa, M. & Sen, R. Rho Protein: Roles and Mechanisms. *Annu Rev Microbiol* **71**, 687-709 (2017).

27. Zhang, J. & Landick, R. A Two-Way Street: Regulatory Interplay between RNA Polymerase and Nascent RNA Structure. *Trends Biochem Sci* **41**, 293-310 (2016).

28. Nudler, E. & Gottesman, M.E. Transcription termination and anti-termination in E. coli. *Genes Cells* **7**, 755-68 (2002).

29. Mondal, S., Yakhnin, A.V., Sebastian, A., Albert, I. & Babitzke, P. NusA-dependent transcription termination prevents misregulation of global gene expression. *Nat Microbiol* **1**, 15007 (2016).

30. Qayyum, M.Z., Dey, D. & Sen, R. Transcription Elongation Factor NusA Is a General Antagonist of Rho-dependent Termination in Escherichia coli. *J Biol Chem* **291**, 8090-108 (2016).

31. Worbs, M., Bourenkov, G.P., Bartunik, H.D., Huber, R. & Wahl, M.C. An extended RNA binding surface through arrayed S1 and KH domains in transcription factor NusA. *Mol Cell* **7**, 1177-89 (2001).

32. Ha, K.S., Toulokhonov, I., Vassylyev, D.G. & Landick, R. The NusA N-terminal domain is necessary and sufficient for enhancement of transcriptional pausing via interaction with the RNA exit channel of RNA polymerase. *J Mol Biol* **401**, 708-25 (2010).

33. Beuth, B., Pennell, S., Arnvig, K.B., Martin, S.R. & Taylor, I.A. Structure of a Mycobacterium tuberculosis NusA-RNA complex. *EMBO J* **24**, 3576-87 (2005).

34. Gollnick, P., Babitzke, P., Antson, A. & Yanofsky, C. Complexity in regulation of tryptophan biosynthesis in Bacillus subtilis. *Annu Rev Genet* **39**, 47-68 (2005).

35. Antson, A.A. et al. Structure of the trp RNA-binding attenuation protein, TRAP, bound to RNA. *Nature* **401**, 235-42 (1999).

36. Babitzke, P., Stults, J.T., Shire, S.J. & Yanofsky, C. TRAP, the trp RNA-binding attenuation protein of Bacillus subtilis, is a multisubunit complex that appears to recognize G/UAG repeats in the trpEDCFBA and trpG transcripts. *J Biol Chem* **269**, 16597-604 (1994).

37. Turnbough, C.L., Jr. & Switzer, R.L. Regulation of pyrimidine biosynthetic gene expression in bacteria: repression without repressors. *Microbiol Mol Biol Rev* **72**, 266-300, table of contents (2008).

38. Figueroa-Bossi, N. et al. RNA remodeling by bacterial global regulator CsrA promotes Rho-dependent transcription termination. *Genes Dev* **28**, 1239-51 (2014).

39. Bae, W., Xia, B., Inouye, M. & Severinov, K. Escherichia coli CspA-family RNA chaperones are transcription antiterminators. *Proc Natl Acad Sci U S A* **97**, 7784-9 (2000).

40. Newkirk, K. et al. Solution NMR structure of the major cold shock protein (CspA) from Escherichia coli: identification of a binding epitope for DNA. *Proc Natl Acad Sci U S A* **91**, 5114-8 (1994).

41. Schindelin, H., Jiang, W., Inouye, M. & Heinemann, U. Crystal structure of CspA, the major cold shock protein of Escherichia coli. *Proc Natl Acad Sci U S A* **91**, 5119-23 (1994).

42. Sachs, R., Max, K.E., Heinemann, U. & Balbach, J. RNA single strands bind to a conserved surface of the major cold shock protein in crystals and solution. *RNA* **18**, 65-76 (2012).

43. Xia, B., Ke, H. & Inouye, M. Acquirement of cold sensitivity by quadruple deletion of the cspA family and its suppression by PNPase S1 domain in Escherichia coli. *Mol Microbiol* **40**, 179-88 (2001).

44. Michaux, C. et al. RNA target profiles direct the discovery of virulence functions for the cold-shock proteins CspC and CspE. *Proc Natl Acad Sci U S A* **114**, 6824-6829 (2017).

45. Phadtare, S., Tadigotla, V., Shin, W.H., Sengupta, A. & Severinov, K. Analysis of Escherichia coli global gene expression profiles in response to overexpression and deletion of CspC and CspE. *J Bacteriol* **188**, 2521-7 (2006).

46. Feng, Y., Huang, H., Liao, J. & Cohen, S.N. Escherichia coli poly(A)-binding proteins that interact with components of degradosomes or impede RNA decay mediated by polynucleotide phosphorylase and RNase E. *J Biol Chem* **276**, 31651-6 (2001).

47. Phadtare, S. & Severinov, K. RNA remodeling and gene regulation by cold shock proteins. *RNA Biol* **7**, 788-95 (2010).

48. Phadtare, S. & Inouye, M. Sequence-selective interactions with RNA by CspB, CspC and CspE, members of the CspA family of Escherichia coli. *Mol Microbiol* **33**, 1004-14 (1999).

49. Kumarevel, T., Mizuno, H. & Kumar, P.K. Structural basis of HutP-mediated anti-termination and roles of the Mg2+ ion and L-histidine ligand. *Nature* **434**, 183-91 (2005).

50. Gopinath, S.C. et al. Insights into anti-termination regulation of the hut operon in Bacillus subtilis: importance of the dual RNA-binding surfaces of HutP. *Nucleic Acids Res* **36**, 3463-73 (2008).

51. Amster-Choder, O. The bgl sensory system: a transmembrane signaling pathway controlling transcriptional antitermination. *Curr Opin Microbiol* **8**, 127-34 (2005).

52. Goodson, J.R., Klupt, S., Zhang, C., Straight, P. & Winkler, W.C. LoaP is a broadly conserved antiterminator protein that regulates antibiotic gene clusters in Bacillus amyloliquefaciens. *Nat Microbiol* **2**, 17003 (2017).

53. DebRoy, S. et al. Riboswitches. A riboswitch-containing sRNA controls gene expression by sequestration of a response regulator. *Science* **345**, 937-40 (2014).

54. Mellin, J.R. et al. Riboswitches. Sequestration of a two-component response regulator by a riboswitch-regulated noncoding RNA. *Science* **345**, 940-3 (2014).

55. Attaiech, L. et al. Silencing of natural transformation by an RNA chaperone and a multitarget small RNA. *Proc Natl Acad Sci U S A* **113**, 8813-8 (2016).

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

57. Zere, T.R. et al. Genomic Targets and Features of BarA-UvrY (-SirA) Signal Transduction Systems. *PLoS One* **10**, e0145035 (2015).

58. Romeo, T., Gong, M., Liu, M.Y. & Brun-Zinkernagel, A.M. Identification and molecular characterization of csrA, a pleiotropic gene from Escherichia coli that affects glycogen biosynthesis, gluconeogenesis, cell size, and surface properties. *J Bacteriol* **175**, 4744-55 (1993).

59. Liu, M.Y. & Romeo, T. The global regulator CsrA of Escherichia coli is a specific mRNA-binding protein. *J Bacteriol* **179**, 4639-42 (1997).

60. Baker, C.S., Morozov, I., Suzuki, K., Romeo, T. & Babitzke, P. CsrA regulates glycogen biosynthesis by preventing translation of glgC in Escherichia coli. *Mol Microbiol* **44**, 1599-610 (2002).

61. Holmqvist, E. et al. Global RNA recognition patterns of post-transcriptional regulators Hfq and CsrA revealed by UV crosslinking in vivo. *EMBO J* **35**, 991-1011 (2016).

62. Potts, A.H. et al. Global role of the bacterial post-transcriptional regulator CsrA revealed by integrated transcriptomics. *Nat Commun* **8**, 1596 (2017).

63. Sahr, T. et al. The Legionella pneumophila genome evolved to accommodate multiple regulatory mechanisms controlled by the CsrA-system. *PLoS Genet* **13**, e1006629 (2017).

64. Dubey, A.K., Baker, C.S., Romeo, T. & Babitzke, P. RNA sequence and secondary structure participate in high-affinity CsrA-RNA interaction. *RNA* **11**, 1579-87 (2005).

65. Schubert, M. et al. Molecular basis of messenger RNA recognition by the specific bacterial repressing clamp RsmA/CsrA. *Nat Struct Mol Biol* **14**, 807-13 (2007).

66. Yakhnin, A.V. et al. CsrA activates flhDC expression by protecting flhDC mRNA from RNase E-mediated cleavage. *Mol Microbiol* **87**, 851-66 (2013).

67. Weilbacher, T. et al. A novel sRNA component of the carbon storage regulatory system of Escherichia coli. *Mol Microbiol* **48**, 657-70 (2003).

68. Duss, O. et al. Structural basis of the non-coding RNA RsmZ acting as a protein sponge. *Nature* **509**, 588-92 (2014).

69. Kusmierek, M. & Dersch, P. Regulation of host-pathogen interactions via the post-transcriptional Csr/Rsm system. *Curr Opin Microbiol* **41**, 58-67 (2017).

70. Katsowich, N. et al. Host cell attachment elicits posttranscriptional regulation in infecting enteropathogenic bacteria. *Science* **355**, 735-739 (2017).

71. Yakhnin, H. et al. CsrA of Bacillus subtilis regulates translation initiation of the gene encoding the flagellin protein (hag) by blocking ribosome binding. *Mol Microbiol* **64**, 1605-20 (2007).

72. Mukherjee, S. et al. CsrA-FliW interaction governs flagellin homeostasis and a checkpoint on flagellar morphogenesis in Bacillus subtilis. *Mol Microbiol* **82**, 447-61 (2011).

73. Agaras, B., Sobrero, P. & Valverde, C. A CsrA/RsmA translational regulator gene encoded in the replication region of a Sinorhizobium meliloti cryptic plasmid complements Pseudomonas fluorescens rsmA/E mutants. *Microbiology* **159**, 230-42 (2013).

74. Marden, J.N. et al. An unusual CsrA family member operates in series with RsmA to amplify posttranscriptional responses in Pseudomonas aeruginosa. *Proc Natl Acad Sci U S A* **110**, 15055-60 (2013).

75. Dugar, G. et al. The CsrA-FliW network controls polar localization of the dual-function flagellin mRNA in Campylobacter jejuni. *Nat Commun* **7**, 11667 (2016).

76. Vogel, J. & Luisi, B.F. Hfq and its constellation of RNA. *Nat Rev Microbiol* **9**, 578-89 (2011).

77. Melamed, S. et al. Global Mapping of Small RNA-Target Interactions in Bacteria. *Mol Cell* **63**, 884-97 (2016).

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

79. Waters, S.A. et al. Small RNA interactome of pathogenic E. coli revealed through crosslinking of RNase E. *EMBO J* **36**, 374-387 (2017).

80. Kavita, K., de Mets, F. & Gottesman, S. New aspects of RNA-based regulation by Hfq and its partner sRNAs. *Curr Opin Microbiol* **42**, 53-61 (2017).

81. Papenfort, K. & Vanderpool, C.K. Target activation by regulatory RNAs in bacteria. *FEMS Microbiol Rev* **39**, 362-78 (2015).

82. Dimastrogiovanni, D. et al. Recognition of the small regulatory RNA RydC by the bacterial Hfq protein. *Elife* **3** (2014).

83. Sauer, E. & Weichenrieder, O. Structural basis for RNA 3'-end recognition by Hfq. *Proc Natl Acad Sci U S A* **108**, 13065-70 (2011).

84. Link, T.M., Valentin-Hansen, P. & Brennan, R.G. Structure of Escherichia coli Hfq bound to polyriboadenylate RNA. *Proc Natl Acad Sci U S A* **106**, 19292-7 (2009).

85. Schumacher, M.A., Pearson, R.F., Moller, T., Valentin-Hansen, P. & Brennan, R.G. Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: a bacterial Sm-like protein. *EMBO J* **21**, 3546-56 (2002).

86. Mikulecky, P.J. et al. Escherichia coli Hfq has distinct interaction surfaces for DsrA, rpoS and poly(A) RNAs. *Nat Struct Mol Biol* **11**, 1206-14 (2004).

87. Peng, Y., Curtis, J.E., Fang, X. & Woodson, S.A. Structural model of an mRNA in complex with the bacterial chaperone Hfq. *Proc Natl Acad Sci U S A* **111**, 17134-9 (2014).

88. Panja, S., Schu, D.J. & Woodson, S.A. Conserved arginines on the rim of Hfq catalyze base pair formation and exchange. *Nucleic Acids Res* **41**, 7536-46 (2013).

89. Schu, D.J., Zhang, A., Gottesman, S. & Storz, G. Alternative Hfq-sRNA interaction modes dictate alternative mRNA recognition. *EMBO J* **34**, 2557-73 (2015).

90. Olsen, A.S., Moller-Jensen, J., Brennan, R.G. & Valentin-Hansen, P. C-terminally truncated derivatives of Escherichia coli Hfq are proficient in riboregulation. *J Mol Biol* **404**, 173-82 (2010).

91. Vecerek, B., Rajkowitsch, L., Sonnleitner, E., Schroeder, R. & Bläsi, U. The C-terminal domain of Escherichia coli Hfq is required for regulation. *Nucleic Acids Res* **36**, 133-43 (2008).

92. Santiago-Frangos, A., Kavita, K., Schu, D.J., Gottesman, S. & Woodson, S.A. C-terminal domain of the RNA chaperone Hfq drives sRNA competition and release of target RNA. *Proc Natl Acad Sci U S A* **113**, E6089-E6096 (2016).

93. Fender, A., Elf, J., Hampel, K., Zimmermann, B. & Wagner, E.G.H. RNAs actively cycle on the Sm-like protein Hfq. *Genes Dev* **24**, 2621-6 (2010).

94. Santiago-Frangos, A., Jeliazkov, J.R., Gray, J.J. & Woodson, S.A. Acidic C-terminal domains autoregulate the RNA chaperone Hfq. *Elife* **6** (2017).

95. Attia, A.S. et al. Moraxella catarrhalis expresses an unusual Hfq protein. *Infect Immun* **76**, 2520-30 (2008).

96. Sittka, A., Sharma, C.M., Rolle, K. & Vogel, J. Deep sequencing of Salmonella RNA associated with heterologous Hfq proteins in vivo reveals small RNAs as a major target class and identifies RNA processing phenotypes. *RNA Biol* **6**, 266-75 (2009).

97. Bouloc, P. & Repoila, F. Fresh layers of RNA-mediated regulation in Gram-positive bacteria. *Curr Opin Microbiol* **30**, 30-35 (2016).

98. Chen, J. & Gottesman, S. Hfq links translation repression to stress-induced mutagenesis in E. coli. *Genes Dev* (2017).

99. Sonnleitner, E. et al. Interplay between the catabolite repression control protein Crc, Hfq and RNA in Hfq-dependent translational regulation in Pseudomonas aeruginosa. *Nucleic Acids Res* (2017).

100. Moreno, R., Fonseca, P. & Rojo, F. Two small RNAs, CrcY and CrcZ, act in concert to sequester the Crc global regulator in Pseudomonas putida, modulating catabolite repression. *Mol Microbiol* **83**, 24-40 (2012).

101. Sonnleitner, E. & Bläsi, U. Regulation of Hfq by the RNA CrcZ in Pseudomonas aeruginosa carbon catabolite repression. *PLoS Genet* **10**, e1004440 (2014).

102. Glover, J.N. et al. The FinO family of bacterial RNA chaperones. *Plasmid* **78**, 79-87 (2015).

103. Chaulk, S.G. et al. ProQ is an RNA chaperone that controls ProP levels in Escherichia coli. *Biochemistry* **50**, 3095-106 (2011).

104. Attaiech, L., Glover, J.N. & Charpentier, X. RNA Chaperones Step Out of Hfq's Shadow. *Trends Microbiol* **25**, 247-249 (2017).

105. Gonzalez, G.M. et al. Structure of the Escherichia coli ProQ RNA-binding protein. *RNA* **23**, 696-711 (2017).

106. Kunte, H.J., Crane, R.A., Culham, D.E., Richmond, D. & Wood, J.M. Protein ProQ influences osmotic activation of compatible solute transporter ProP in Escherichia coli K-12. *J Bacteriol* **181**, 1537-43 (1999).

107. Davis, J.H. & Williamson, J.R. Structure and dynamics of bacterial ribosome biogenesis. *Philos Trans R Soc Lond B Biol Sci* **372** (2017).

108. Ramakrishnan, V. The ribosome emerges from a black box. *Cell* **159**, 979-984 (2014).

109. Melnikov, S. et al. One core, two shells: bacterial and eukaryotic ribosomes. *Nat Struct Mol Biol* **19**, 560-7 (2012).

110. Voorhees, R.M. & Ramakrishnan, V. Structural basis of the translational elongation cycle. *Annu Rev Biochem* **82**, 203-36 (2013).

111. Shajani, Z., Sykes, M.T. & Williamson, J.R. Assembly of bacterial ribosomes. *Annu Rev Biochem* **80**, 501-26 (2011).

112. Meyer, M.M. rRNA mimicry in RNA regulation of gene expression. *Microbiol Spectr* (2018).

113. Merianos, H.J., Wang, J. & Moore, P.B. The structure of a ribosomal protein S8/spc operon mRNA complex. *RNA* **10**, 954-64 (2004).

114. Marzi, S. et al. Structured mRNAs regulate translation initiation by binding to the platform of the ribosome. *Cell* **130**, 1019-31 (2007).

115. Babitzke, P., Baker, C.S. & Romeo, T. Regulation of translation initiation by RNA binding proteins. *Annu Rev Microbiol* **63**, 27-44 (2009).

116. Duval, M. et al. Escherichia coli ribosomal protein S1 unfolds structured mRNAs onto the ribosome for active translation initiation. *PLoS Biol* **11**, e1001731 (2013).

117. Boni, I.V., Artamonova, V.S., Tzareva, N.V. & Dreyfus, M. Non-canonical mechanism for translational control in bacteria: synthesis of ribosomal protein S1. *EMBO J* **20**, 4222-32 (2001).

118. Butler, J.S., Springer, M., Dondon, J., Graffe, M. & Grunberg-Manago, M. Escherichia coli protein synthesis initiation factor IF3 controls its own gene expression at the translational level in vivo. *J Mol Biol* **192**, 767-80 (1986).

119. Romby, P. & Springer, M. Bacterial translational control at atomic resolution. *Trends Genet* **19**, 155-61 (2003).

120. Jain, C. & Belasco, J.G. Autoregulation of RNase E synthesis in Escherichia coli. *Nucleic Acids Symp Ser*, 85-8 (1995).

121. Jarrige, A.C., Mathy, N. & Portier, C. PNPase autocontrols its expression by degrading a double-stranded structure in the pnp mRNA leader. *EMBO J* **20**, 6845-55 (2001).

122. Lopez-Alonso, J.P. et al. RsgA couples the maturation state of the 30S ribosomal decoding center to activation of its GTPase pocket. *Nucleic Acids Res* (2017).

123. Ataide, S.F. et al. The crystal structure of the signal recognition particle in complex with its receptor. *Science* **331**, 881-6 (2011).

124. Kohler, R., Mooney, R.A., Mills, D.J., Landick, R. & Cramer, P. Architecture of a transcribing-translating expressome. *Science* **356**, 194-197 (2017).

125. Laalami, S., Zig, L. & Putzer, H. Initiation of mRNA decay in bacteria. *Cell Mol Life Sci* **71**, 1799-828 (2014).

126. Redder, P., Hausmann, S., Khemici, V., Yasrebi, H. & Linder, P. Bacterial versatility requires DEAD-box RNA helicases. *FEMS Microbiol Rev* **39**, 392-412 (2015).

127. Py, B., Higgins, C.F., Krisch, H.M. & Carpousis, A.J. A DEAD-box RNA helicase in the Escherichia coli RNA degradosome. *Nature* **381**, 169-72 (1996).

128. Coburn, G.A., Miao, X., Briant, D.J. & Mackie, G.A. Reconstitution of a minimal RNA degradosome demonstrates functional coordination between a 3' exonuclease and a DEAD-box RNA helicase. *Genes Dev* **13**, 2594-603 (1999).

129. Oun, S. et al. The CshA DEAD-box RNA helicase is important for quorum sensing control in Staphylococcus aureus. *RNA Biol* **10**, 157-65 (2013).

130. Morita, T., Maki, K. & Aiba, H. RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs. *Genes Dev* **19**, 2176-86 (2005).

131.   Bruce, H.A. et al. Analysis of the natively unstructured RNA/protein-recognition core in the Escherichia coli RNA degradosome and its interactions with regulatory RNA/Hfq complexes. *Nucleic Acids Res* (2017).

132.   Bandyra, K.J. et al. The seed region of a small RNA drives the controlled destruction of the target mRNA by the endoribonuclease RNase E. *Mol Cell* **47**, 943-53 (2012).

133.   Pfeiffer, V., Papenfort, K., Lucchini, S., Hinton, J.C. & Vogel, J. Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nat Struct Mol Biol* **16**, 840-6 (2009).

134.   Göpel, Y., Papenfort, K., Reichenbach, B., Vogel, J. & Görke, B. Targeted decay of a regulatory small RNA by an adaptor protein for RNase E and counteraction by an anti-adaptor RNA. *Genes Dev* **27**, 552-64 (2013).

135.   Gonzalez, G.M. et al. Structural insights into RapZ-mediated regulation of bacterial amino-sugar metabolism. *Nucleic Acids Res* **45**, 10845-10860 (2017).

136.   Vakulskas, C.A. et al. Antagonistic control of the turnover pathway for the global regulatory sRNA CsrB by the CsrA and CsrD proteins. *Nucleic Acids Res* **44**, 7896-910 (2016).

137.   Chen, X. et al. An RNA degradation machine sculpted by Ro autoantigen and noncoding RNA. *Cell* **153**, 166-77 (2013).

138.   Moll, I., Afonyushkin, T., Vytvytska, O., Kaberdin, V.R. & Bläsi, U. Coincident Hfq binding and RNase E cleavage sites on mRNA and small regulatory RNAs. *RNA* **9**, 1308-14 (2003).

139.   Chao, Y. et al. In Vivo Cleavage Map Illuminates the Central Role of RNase E in Coding and Non-coding RNA Pathways. *Mol Cell* **65**, 39-51 (2017).

140.   Chao, Y. & Vogel, J. A 3' UTR-Derived Small RNA Provides the Regulatory Noncoding Arm of the Inner Membrane Stress Response. *Mol Cell* **61**, 352-363 (2016).

141.   Pandey, S.P. et al. Central role for RNase YbeY in Hfq-dependent and Hfq-independent small-RNA regulation in bacteria. *BMC Genomics* **15**, 121 (2014).

142.   Regnier, P. & Hajnsdorf, E. The interplay of Hfq, poly(A) polymerase I and exoribonucleases at the 3' ends of RNAs resulting from Rho-independent termination: A tentative model. *RNA Biol* **10**, 602-9 (2013).

143.   Andrade, J.M., Pobre, V., Matos, A.M. & Arraiano, C.M. The crucial role of PNPase in the degradation of small RNAs that are not associated with Hfq. *RNA* **18**, 844-55 (2012).

144.   Liou, G.G., Jane, W.N., Cohen, S.N., Lin, N.S. & Lin-Chao, S. RNA degradosomes exist in vivo in Escherichia coli as multicomponent complexes associated with the cytoplasmic membrane via the N-terminal region of ribonuclease E. *Proc Natl Acad Sci U S A* **98**, 63-8 (2001).

145.   Alen, C. & Sonenshein, A.L. Bacillus subtilis aconitase is an RNA-binding protein. *Proc Natl Acad Sci U S A* **96**, 10412-7 (1999).

146.   Benjamin, J.A. & Massé, E. The iron-sensing aconitase B binds its own mRNA to prevent sRNA-induced mRNA cleavage. *Nucleic Acids Res* **42**, 10023-36 (2014).

147.   Mitobe, J. et al. RodZ regulates the post-transcriptional processing of the Shigella sonnei type III secretion system. *EMBO Rep* **12**, 911-6 (2011).

148.   Chen, Y. & Anderson, D.M. Expression hierarchy in the Yersinia type III secretion system established through YopD recognition of RNA. *Mol Microbiol* **80**, 966-80 (2011).

149.   Beljantseva, J. et al. Negative allosteric regulation of Enterococcus faecalis small alarmone synthetase RelQ by single-stranded RNA. *Proc Natl Acad Sci U S A* **114**, 3726-3731 (2017).

150.   Qian, Z., Zhurkin, V.B. & Adhya, S. DNA-RNA interactions are critical for chromosome condensation in Escherichia coli. *Proc Natl Acad Sci U S A* **114**, 12225-12230 (2017).

151.   Brescia, C.C., Kaw, M.K. & Sledjeski, D.D. The DNA binding protein H-NS binds to and alters the stability of RNA in vitro and in vivo. *J Mol Biol* **339**, 505-14 (2004).

152. Deighan, P., Free, A. & Dorman, C.J. A role for the Escherichia coli H-NS-like protein StpA in OmpF porin expression through modulation of micF RNA stability. *Mol Microbiol* **38**, 126-39 (2000).

153. Buskila, A.A., Kannaiah, S. & Amster-Choder, O. RNA localization in bacteria. *RNA Biol* **11**, 1051-60 (2014).

154. Wallace, J.G., Zhou, Z. & Breaker, R.R. OLE RNA protects extremophilic bacteria from alcohol toxicity. *Nucleic Acids Res* **40**, 6898-907 (2012).

155. Kuwada, N.J., Traxler, B. & Wiggins, P.A. Genome-scale quantitative characterization of bacterial protein localization dynamics throughout the cell cycle. *Mol Microbiol* **95**, 64-79 (2015).

156. Landgraf, D., Okumus, B., Chien, P., Baker, T.A. & Paulsson, J. Segregation of molecules at cell division reveals native protein localization. *Nat Methods* **9**, 480-2 (2012).

157. Short, F.L. et al. Selectivity and self-assembly in the control of a bacterial toxin by an antitoxic noncoding RNA pseudoknot. *Proc Natl Acad Sci U S A* **110**, E241-9 (2013).

158. Rajagopala, S.V. et al. The binary protein-protein interaction landscape of Escherichia coli. *Nat Biotechnol* **32**, 285-290 (2014).

159. Nichols, R.J. et al. Phenotypic landscape of a bacterial cell. *Cell* **144**, 143-56 (2011).

160. Hattman, S. Unusual transcriptional and translational regulation of the bacteriophage Mu mom operon. *Pharmacol Ther* **84**, 367-88 (1999).

161. Romaniuk, P.J., Lowary, P., Wu, H.N., Stormo, G. & Uhlenbeck, O.C. RNA binding site of R17 coat protein. *Biochemistry* **26**, 1563-8 (1987).

162. McPheeters, D.S., Stormo, G.D. & Gold, L. Autogenous regulatory site on the bacteriophage T4 gene 32 messenger RNA. *J Mol Biol* **201**, 517-35 (1988).

163. Said, N. et al. Structural basis for lambdaN-dependent processive transcription antitermination. *Nat Microbiol* **2**, 17062 (2017).

164. Said, N. et al. In vivo expression and purification of aptamer-tagged small RNA regulators. *Nucleic Acids Res* **37**, e133 (2009).

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

166. Chao, Y., Papenfort, K., Reinhardt, R., Sharma, C.M. & Vogel, J. An atlas of Hfq-bound transcripts reveals 3' UTRs as a genomic reservoir of regulatory small RNAs. *EMBO J* **31**, 4005-19 (2012).

167. Sittka, A. et al. Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq. *PLoS Genet* **4**, e1000163 (2008).

168. Zhang, A. et al. Global analysis of small RNA and mRNA targets of Hfq. *Mol Microbiol* **50**, 1111-24 (2003).

169. Wheeler, E.C., Van Nostrand, E.L. & Yeo, G.W. Advances and challenges in the detection of transcriptome-wide protein-RNA interactions. *Wiley Interdiscip Rev RNA* (2017).

170. Lambert, N. et al. RNA Bind-n-Seq: quantitative assessment of the sequence and structural binding specificity of RNA binding proteins. *Mol Cell* **54**, 887-900 (2014).

171. Buenrostro, J.D. et al. Quantitative analysis of RNA-protein interactions on a massively parallel array reveals biophysical and evolutionary landscapes. *Nat Biotechnol* **32**, 562-8 (2014).

172. Tome, J.M. et al. Comprehensive analysis of RNA-protein interactions by high-throughput sequencing-RNA affinity profiling. *Nat Methods* **11**, 683-8 (2014).

173. Sedlyarova, N. et al. Natural RNA Polymerase Aptamers Regulate Transcription in E. coli. *Mol Cell* **67**, 30-43 e6 (2017).

174.   Windbichler, N., von Pelchrzim, F., Mayer, O., Csaszar, E. & Schroeder, R. Isolation of small RNA-binding proteins from E. coli: evidence for frequent interaction of RNAs with RNA polymerase. *RNA Biol* **5**, 30-40 (2008).

175.   Osborne, J., Djapgne, L., Tran, B.Q., Goo, Y.A. & Oglesby-Sherrouse, A.G. A method for in vivo identification of bacterial small RNA-binding proteins. *Microbiologyopen* **3**, 950-60 (2014).

176.   Scherrer, T., Mittal, N., Janga, S.C. & Gerber, A.P. A screen for RNA-binding proteins in yeast indicates dual functions for many enzymes. *PLoS One* **5**, e15499 (2010).

177.   Tsvetanova, N.G., Klass, D.M., Salzman, J. & Brown, P.O. Proteome-wide search reveals unexpected RNA-binding proteins in Saccharomyces cerevisiae. *PLoS One* **5** (2010).