



HAL
open science

Regulatory interplay between small RNAs and transcription termination factor Rho

Lionello Bossi, Nara Figueroa-Bossi, Philippe Bouloc, Marc Boudvillain

► To cite this version:

Lionello Bossi, Nara Figueroa-Bossi, Philippe Bouloc, Marc Boudvillain. Regulatory interplay between small RNAs and transcription termination factor Rho. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 2020, pp.194546. 10.1016/j.bbagr.2020.194546 . hal-02533337

HAL Id: hal-02533337

<https://hal.science/hal-02533337>

Submitted on 6 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Regulatory interplay between small RNAs and transcription termination factor Rho

Lionello Bossi^{a*}, Nara Figueroa-Bossi^a, Philippe Bouloc^a and Marc Boudvillain^b

^a *Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198, Gif-sur-Yvette, France*

^b *Centre de Biophysique Moléculaire, CNRS UPR4301, rue Charles Sadron, 45071 Orléans cedex 2, France*

* Corresponding author: lionello.bossi@i2bc.paris-saclay.fr

Highlights

- Repression of translation by sRNAs exposes Rho binding sites in nascent mRNA causing transcription termination
- Rho-dependent transcription termination allows sRNAs to downregulate distal genes in polycistronic transcripts
- Regulatory sRNAs can promote anti-termination by competing with Rho for RNA binding

ABSTRACT

The largest and best studied group of regulatory small RNAs (sRNAs) in bacteria act by modulating translation or turnover of messenger RNAs (mRNAs) through base-pairing interactions that typically take place near the 5' end of the mRNA. This allows the sRNA to bind the complementary target sequence while the remainder of the mRNA is still being made, creating conditions whereby the action of the sRNA can extend to transcriptional steps, most notably transcription termination. Increasing evidence corroborates the existence of a functional interplay between sRNAs and termination factor Rho. Two general mechanisms have emerged. One mechanism operates in translated regions subjected to sRNA repression. By inhibiting ribosome binding co-transcriptionally, the sRNA uncouples translation from transcription, allowing Rho to bind the nascent RNA and promote termination. In the second mechanism, which functions in 5' untranslated regions, the sRNA antagonizes termination directly by interfering with Rho binding to the RNA or the subsequent translocation along the RNA. Here, we review the above literature in the context of other mechanisms that underlie the participation of Rho-dependent transcription termination in gene regulation.

1. Transcription-translation coupling, polarity and Rho factor

Early work on operon function in *E. coli* revealed that nonsense mutations causing the premature stop of protein synthesis in genes proximal to the promoter, often led to a decrease in the expression of the downstream genes [1, 2]. That is the case in the *lac* operon where peptide chain terminating mutations in the first structural gene, *lacZ*, lower expression of the second gene, *lacY*. In an effort to elucidate the basis of these polarity effects, J Beckwith isolated suppressor mutations that restored *lacY* expression in strains with nonsense mutations in *lacZ*. He identified a class of mutations that were unlinked to the *lac* operon and mapped in an unknown locus that he named *suA* [3]. These suppressors remained uncharacterized for more than a decade until Richardson and coworkers determined that *suA* mutations affected the gene encoding transcription termination factor Rho [4]. JW Roberts had identified Rho as an activity causing the arrest of RNA synthesis and the release of transcripts from phage lambda DNA template in an *in vitro* system [5]. Finding that premature termination of protein synthesis triggers Rho-mediated termination of transcription elongation provided the first indication that translation and transcription in bacteria are dynamically coupled. This notion has since been confirmed by a large body of evidence. During transcription of protein-coding genes, the lead ribosome is thought to bind to the nascent mRNA as soon as it emerges from the RNA exit channel of RNA polymerase [6]. The ribosome then closely follows, or even pushes [7] the transcription elongation complex becoming physically linked to it through interactions between ribosomal protein S10 (NusE) and elongation factor NusG [8-10]. This tight coupling effectively shields the nascent RNA from the action of Rho (Fig. 1A). The situation changes drastically if the mRNA carries a mutation that generates a stop codon. Loss of ribosome shielding beyond the mutant site leaves the mRNA exposed to the binding of Rho which replaces S10 as a NusG binding partner and terminates transcription [11](Fig. 1B). More recently, it was shown that transcription termination within translated regions can also occur at sites where the newly-made RNA contains sequences that “stick” to RNA polymerase. These RNA sequences, named iRAPs, were proposed to cause ribosomal stalling and thus allow Rho to gain access to the nascent mRNA [12].

2. Rho mechanics, RNA binding specificity and NusG involvement

Bacteria use two main transcription termination pathways: intrinsic termination (or Rho-independent termination) does not require any auxiliary factors and is dictated solely by the structure of the nascent RNA. The main structural elements of an intrinsic terminator are a CG-rich palindrome (forming a hairpin-like structure in the nascent RNA) immediately followed by a track of uridines [11 and references

therein]. The second termination mechanism involves Rho factor. Most of the current knowledge on the structure and function of Rho comes from work with the *E. coli* protein. Rho is an RNA-binding homohexameric protein that can adopt alternative ring-shaped conformations, one in which ring is open and slightly distorted (i.e., ends slightly offset), the other with the ring closed [13-15]. The hexamer is thought to initially interact with the RNA in the open conformation. The RNA follows a crown-like path along one face of the ring contacting the N-terminal domain of each monomer in a narrow cleft that can only accommodate pyrimidine dinucleotides, with a preference for cytosine [16]. The overall shape of the complex directs the RNA chain toward the center of the ring through the open gate. RNA binding triggers a conformational rearrangement that closes the gate and traps the RNA chain in the internal cavity [17]. Establishment of contacts between the RNA backbone and the C-terminal helicase-like domain then activates an ATP-fueled translocase motor [18] that allows Rho to move along the RNA chain. How Rho actually triggers termination remains subject to debate. One model posits that the translocase action either pulls the RNA away from the DNA-RNA hybrid in the active site (hybrid shearing) or pushes the polymerase (hypertranslocation). Either way, the net result is the destabilization of the elongation complex and, ultimately, the dissociation of RNA polymerase from the DNA template and RNA transcript [11]. An alternative model proposes that Rho invades the main channel of the elongation complex inducing an allosteric change that alters both the geometry of the RNA polymerase catalytic center and the stability of the transcription complex [19].

Rho binds RNA sequences that contain an excess of cytosine versus guanine residues (a "C>G bubble") over a region of approximately 50 to 80 nucleotides (nt) [20]. In particular, Rho utilization (*Rut*) sites often contain UC and CC dinucleotide repeats spaced every 9 to 13 nucleotides (to allow concomitant binding of multiple subunits) and lacking secondary structures that could sequester these motifs. Termination occurs downstream from the *Rut* site, within a window that can extend as many as 100 nt from the downstream edge of the *Rut* sequence [21]. Extensive variability in the size and shape of the C>G bubble has prevented establishing a consensus sequence for Rho binding and *Rut* sites remain difficult to identify based solely on sequence composition. A computational method based on multivariate statistics incorporating information from a large number of *in vitro* termination assays was recently developed and showed to have a high success rate at predicting sequences apt to Rho-dependent termination [22]. The statistical determinants of the method are generally consistent with the aforementioned features and composition of *Rut* sites. Still, a persisting problem in *Rut* site definition originates from the participation of NusG in Rho-mediated termination. NusG, a conserved transcription elongation factor, binds Rho via its C-terminal domain when the latter is free from the interaction with the ribosome [8, 23] (Fig. 1B). NusG binding accelerates the transition of Rho from the open RNA-loading configuration to the

closed-ring structure required for the translocase activity [24]. In doing so, NusG relaxes the need for a strong RNA interaction to drive the transition, allowing Rho to act at suboptimal *Rut* sites that are difficult, if not impossible to predict.

Most transcription units in the *E. coli* genome are terminated at intrinsic termination sites. Rho-dependent terminators account for no more than 20% of the termination signals found at the end of genes or operons [25]. The advantage of using one termination mechanism versus the other is not obvious. Of note is that Rho is more often used to terminate transcription at the end of transcription units where the next gene is in the opposite orientation [26]. Intriguingly, in most, if not all, of these cases, the site of termination is actually within the antisense strand of the next gene, meaning that the sequence of this gene must have evolved to contain both coding information in the sense strand and a termination signal in the antisense strand. Whether this arrangement confers some special features to the expression/regulation of head-on transcription units is currently unknown. The specialization toward targeting antisense transcription, either produced by *bone-fide* promoters or by spurious promoter-like sequences within coding regions, is a prominent feature of Rho activity in the bacterial cell [26-31].

3. Rho-dependent polarity extends the range of action of translation-targeting sRNAs

Bacterial non-coding sRNAs comprise a family of regulatory molecules that act by modulating mRNA translation and/or stability through Watson-Crick base-pairing [32-34]. The majority of sRNAs range in length between 70 and 140 nt; they are usually primary transcription products or can sometimes result from 3' end processing of longer mRNA precursors [35, 36]. sRNA transcripts are generally terminated at intrinsic terminators and thus incorporate the 3' CG-rich stem-loop and uridine track in their mature structure. In Gram-negative bacteria, a major group of sRNAs require the binding of chaperone protein Hfq for their stability and function [37]. Hfq, a ring-shaped hexamer, binds both the sRNA and the target mRNA and accelerates the formation of the RNA duplex [38, 39]. In addition, Hfq binding to free sRNA protects the sRNA against degradation by masking ribonuclease recognition sites [40, 41]. The polyU track at the 3' end of sRNA is a major contributor to the sRNA binding affinity for Hfq [42]. Thus, the polyU is important not only for the biogenesis but also for the function of the sRNA, a finding that raises a number of interesting implications that were elegantly discussed in a recent review article [43].

The first line of evidence linking Rho to sRNA-mediated regulation was obtained during the study of the *chiPQ* operon in *Salmonella*. This operon encodes a porin for chitin-derived sugars (ChiP) and a small lipoprotein of unknown function (ChiQ). Expression of the *chiPQ* operon is strongly induced when bacteria grow using chitosugars as carbon and nitrogen sources, but is otherwise tightly repressed [44].

Repression results from the combined action of the NagC repressor, which downregulates but does not shut off transcription from the *chiPQ* promoter, and the small RNA ChiX, which represses translation of the transcripts that escape NagC repression [44, 45]. ChiX acts by base-pairing with a sequence in the 5' untranslated region (UTR) of the *chiPQ* mRNA, blocking ribosome binding at the *chiP* initiation codon. The sRNA action is particularly efficient and results in the complete silencing of *chiP*. Remarkably, ChiX also completely silences *chiQ*, even though the *chiQ* sequence lies nearly 1500 nt downstream from the ChiX binding site in the 5' UTR. A genetic study aimed at elucidating the mechanism responsible for *chiQ* silencing, identified three classes of mutations that restore *chiQ* expression in the presence of ChiX. They included mutations in *rho*, *nusG* and in a presumptive *Rut* site inside the translated portion of *chiP* gene, about 200 nt downstream from the initiating AUG [46]. These findings, combined with *in vitro* evidence, supported the conclusion that ChiX represses *chiQ* "from a distance" by promoting NusG-stimulated Rho-dependent termination early in the *chiP* gene. The most likely explanation for these effects is that by blocking ribosome binding to the mRNA, the regulatory sRNA allows Rho to gain access to the *Rut* site, bind NusG at the same time, and terminate transcription (Fig. 2). Conservation of the *chiP* *Rut* site in *Enterobacteriaceae* suggests that ChiX-mediated Rho recruitment operates throughout this family. In fact, the *Rut* site of the *chiP* gene of *E. coli* shows a larger C>G bubble (Fig. 3A) and a CC/UC repeat disposition that could allow Rho to contact the RNA in two alternative registers (Fig. 3B). Consistent with this observation, the *E. coli* *chiP* *Rut* site elicits Rho termination more efficiently than the *Salmonella* counterpart *in vivo* and *in vitro* and is much less dependent on NusG *in vitro* (Schwartz, Bossi, Figueroa-Bossi, Boudvillain, unpublished).

A second example linking Rho to sRNA regulation has come from studying the mechanism by which the sRNA Spot 42 (also named Spf) causes the discoordinate expression of the galactose operon (*galETKM*) in *E. coli*. Spot 42 accumulates during catabolite repression and specifically inhibits translation of the third cistron in the mRNA, *galk*, by base-pairing with a sequence at the *galt-galk* boundary [47]. This selective action that can be rationalized by the requirement for the other gene products (notably GalE and GalT) in pathways other than galactose catabolism. A recent study found evidence that the translational repression of *galk* by Spot 42 activates Rho-dependent transcription termination near the site of Spot 42 pairing thus contributing to *galk* silencing [48]. Surprisingly, deleting the Spot 42 gene [48] or overproducing the sRNA [47] had no apparent effect on the expression of the distal cistron *galM*. This observation seems at odds with the Rho involvement model and remains unexplained.

Finding an sRNA to downregulate genes that are downstream from its primary base-pairing target within an operon is not uncommon. Does Rho contribute to these effects? And if so, why more examples like those described above have not

been reported? A trivial answer to these questions is that in most instances, the Rho involvement was simply not investigated, the decrease in distal gene expression being generally assumed to result from mRNA degradation or loss of translational coupling. Although mRNA degradation is the nearly obligatory outcome of this type of regulation, the role of degradation in the expression of downstream cistrons is not obvious. The untranslated portion of the mRNA is degraded as a result of the concerted actions of RNase E that cleaves the RNA internally and exonucleases that degrade the resulting fragments in a 3' to 5' direction [49]. Since no 5'-3' exonucleases are known in Gram-negative bacteria, the downstream cistron should escape degradation because shielded by ribosomes as soon as its ribosome-binding site emerges from the exit channel of ARN polymerase (Fig. 4A and 4B). Unless RNase E cleaves this sequence before the first ribosome can bind, degradation is not expected to contribute much to polarity. The situation is different in the case of translational coupling. The latter occurs when the translation of a cistron controls the translation rate of the cistron that is immediately downstream. Such coupling can result from ribosomes reinitiating translation immediately after reaching a stop codon [50] or from ribosomes unfolding mRNA structures that occlude the ribosome-binding site of the downstream cistron [51]. Loss of translational coupling can thus be expected to contribute to polarity alongside Rho-dependent transcription termination (Fig. 4C). One such example can be found in the regulation of the *dppABCDF* operon encoding the major dipeptide transporter of *E. coli* and *Salmonella*. In bacteria growing exponentially in rich medium, the small RNA GcvB inhibits translation of the first cistron, *dppA*, by base-pairing with a sequence near the ribosome-binding site of this cistron [52]. However, GcvB also downregulates the remaining genes even though it does not bind to any place else in the polycistronic mRNAs [52, 53]. The mechanism underlying this extended regulation was never directly analyzed. Recently, evidence was obtained pointing to the participation of Rho in this regulation. The evidence comes from experiments with a *dppF-lacZ* fusion, originally isolated as an Hfq-regulated gene fusion in *Salmonella* [54] and subsequently confirmed to be downregulated by GcvB. Expression of *dppF-lacZ* increases four-fold in a *gcvB*[+] strain carrying a Rho mutation (Rho Y80C). In contrast, Rho Y80C does not have any effect on *dppF-lacZ* expression in a strain deleted for *gcvB*, suggesting that the Rho involvement is intimately linked to GcvB activity (Bossi, Figueroa-Bossi, unpublished). The *dppABCDF* transcript might include one or more *Rut* sites, upstream of *dppF*, that become exposed when *dppA* translation is inhibited by GcvB. None of these sites appears to be located in *dppA* itself, as the expression of a *dppB-lacZ* fusion did not increase significantly in the Rho Y80C mutant (Bossi, Figueroa-Bossi, unpublished). It follows that the GcvB effects on *dppB* do not involve Rho and are most likely ascribable to the loss of transcriptional coupling due the *dppA* repression (Fig. 4C). One can then envision a model whereby the loss of translational coupling contributes to transcriptional polarity by

propagating the rarefaction of ribosomes to downstream cistrons (Fig. 4D). Clearly, this would be an economic and efficient way to coordinate transcriptional and post-transcriptional regulatory inputs throughout the operon. The model can be extrapolated to other GcvB targets, for example the *oppABCDF* operon whose structural features closely resemble those of *dppABCDF* [52, 53].

4. The Rut site as a regulatory switch

Rho-dependent terminators are often found in intergenic or 5' leader regions where they participate in regulatory mechanisms. A paradigm for this type of regulation is phage lambda antitermination mechanism whereby binding of N protein to a sequence within a *Rut* site promotes the assembly of a multiprotein complex that allows RNA polymerase to override the terminator and transcribe replication genes at the onset of the lytic cycle [55]. In this regulatory architecture, shared by many lambdoid phages, the terminator is constitutively active and becomes silenced by the activity of one or more anti-termination factors. A different regulatory design is found in the *E. coli pgaABCD* operon, which encodes proteins needed for the synthesis and export of a polysaccharide adhesin involved in biofilm formation. The 5' UTR of the *pgaABCD* mRNA includes a Rho-dependent terminator which is however inactive because sequestered in a secondary structure that makes it inaccessible to Rho [56]. Upon a shift to a planktonic environment, binding of global regulator protein CsrA to one arm of the secondary structure unmask the *Rut* site making it available for the binding of Rho. Rho-mediated termination prevents expression of the operon (unnecessary during planktonic growth) [56]. This system provides a unique example of a protein, in this case CsrA, acting as a pro-termination factor. Interestingly, CsrA also binds near the *pgaA* initiating AUG codon and it actually does so with higher affinity than that for the upstream *Rut* site [57]. This suggests that activation of the *pgaABCD* operon occurs gradually as a function of the decreased CsrA availability, *i.e.*, transcription termination being relieved first while translation repression is still maintained. The decrease in CsrA levels, in cells moving toward a sessile lifestyle, follows the accumulation of two non-coding RNAs, CsrB and CsrC, that contain multiple CsrA binding sites and sequester the protein [58].

Rho participation in gene regulation comes in a variety of additional flavors. The mechanism regulating expression of tryptophanase in *E. coli* [59] and three separate Mg⁺⁺ transporters in *Salmonella* [60-62] couples Rho activity to a translational sensor. The leader regions of these transcripts include a short open reading frame (orf) and are structured in such a way that if a ribosome stalls during translation of the short orf, Rho binding to the RNA is prevented allowing transcription to read through the terminator. In contrast, if there is no impediment to translation of the orf, ribosome dissociation upon completion of the translational run opens a time window in which Rho can bind to the RNA and terminate transcription. Thus, this

mechanism functions in a manner opposite to classical transcriptional polarity and hence has been termed “reverse polarity” [63]. At the tryptophanase-encoding locus (*tnaAB*), ribosomal stalling is triggered by high tryptophan levels and results from L-Trp molecules invading the ribosomal exit tunnel and blocking the release of the leader peptide [64, 65]. The stalled ribosome sterically hinders the binding of Rho to an adjacent *Rut* site, allowing transcription to proceed into the *tnaAB* operon (thus meeting the need for tryptophanase expression). In the transcripts for Mg⁺⁺ transporters (*mgtA*, *mgtB*, *corA*), the ribosome stalls in the leader region when Mg⁺⁺ is limiting, apparently due to an Mg⁺⁺ requirement for peptidyl transfer at certain codons, particularly proline [60-62]. Ribosomal stalling promotes folding of secondary structures that interfere with the accessibility of a *Rut* site, favoring transcriptional readthrough (thus meeting the need for expression of the transporters).

Reverse polarity is also observed in the autogenous regulation of the *Salmonella tufB* gene encoding translation elongation factor Tu. In this case, the slowdown of the ribosomes, when the Tu concentration is suboptimal, creates a traffic jam in the initial portion of the *tufB* orf that favors formation of a secondary structure that occludes a Rho binding site [66]. The resulting transcriptional deattenuation restores optimal TufB levels. In contrast, if Tu levels are too high, rapid displacement of the initiating ribosome allows the 5' UTR to fold (before the next ribosome can bind) into a structure with the *Rut* site accessible, leading to premature termination of transcription [66].

In a yet different scenario, Rho activity has been linked to the functioning of a flavin mononucleotide (FMN) riboswitch in *E. coli*. FMN binding to the 5' UTR of the *E. coli ribB* gene, coding for a riboflavin biosynthetic enzyme, was found to stimulate Rho-dependent termination within the *ribB* leader region *in vitro* [67]. This observation, combined with the finding that Rho inhibition by the drug bicyclomycin (BCM) causes *ribB* expression to increase *in vivo*, has led to the proposal that RNA remodeling by FMN creates a functional *Rut* site [67]. Since the FMN-bound form of the riboswitch also blocks *ribB* translation by masking the ribosome-binding site [68] a non-mutually exclusive possibility is that translational repression uncovers Rho binding sites in the protein-coding portion of the mRNA leading to premature termination. This was shown to be the case in the regulation of the *thiM* riboswitch, which controls the expression of genes involved in the synthesis of thiamin pyrophosphate (TPP). The TPP-bound form of the riboswitch inhibits translation of the *thiMD* mRNA (by sequestering the ribosome-binding site) and in doing so exposes a *Rut* site in the initial portion of the *thiM* coding region, leading to transcription termination in the window between codons 20 and 34 [69]. In contrast, in the TPP-sensing riboswitch of the *thiCEFSGH* operon (also involved TPP biosynthesis), the *Rut* site lies entirely within the untranslated region and is accessible to Rho in the TPP-bound conformer. Transcription terminates near the

initiating AUG. Although TPP-bound conformer also represses *thiC* translation, this appears to be a secondary, fail-safe mechanism to prevent expression of transcripts that escape termination [70].

5. Regulatory sRNAs as antiterminators

Long 5' untranslated regions are expected to be potential targets for Rho action due to the lack of ribosomal shielding in the nascent RNA. Indeed, a genome-wide survey found that transcripts with 5' UTRs longer than 80 nt were overrepresented among RNAs whose levels increased in *E. coli* cells treated with BCM [71]. One of the BCM-responsive transcripts was from the *rpoS* gene encoding the general stress response/stationary phase sigma factor SigmaS. This suggested that the *rpoS* leader region might contain one or more Rho termination sites, a guess readily confirmed by *in vitro* transcription analysis with purified Rho protein [71]. The *rpoS* mRNA has a particularly long leader sequence (567 nt) initiating at a promoter within the upstream *nlpD* gene [72]. This region includes an extensive stem-loop structure that occludes the ribosome-binding site of *rpoS* and causes the *rpoS* mRNA to be poorly translated. Any of three different Hfq-dependent sRNAs (DsrA, RprA, ArcZ), synthesized in response to different stress conditions, were previously known to activate *rpoS* translation by base-pairing with a sequence in the upstream arm of the inhibitory stem and disrupting the structure [73, 74]. Finding the most prominent Rho termination sites *in vitro* to map in close proximity of the sRNA target sequence prompted a study assessing whether the sRNAs affected the termination mechanism. This work confirmed that all three sRNAs antagonize Rho function in the *rpoS* leader region [71]. Thus, the sRNAs act at two levels: promoting transcriptional readthrough and, at the same time, stimulating translation of the readthrough transcripts. The authors proposed that the sRNAs either interfere with Rho loading onto the RNA or with the subsequent translocation step [71]. In the latter scenario, the bound sRNA would act as a "speed bump" causing Rho to stall long enough to allow RNA polymerase to escape ahead (Fig. 5). Alternatively, the sRNAs might promote formation of secondary structures that mask the *Rut* site or expose sequence motifs that inhibit Rho action [75]. All of the above mechanisms could be further potentiated by the presence of Hfq, which is known to have multiple contact points in the *rpoS* leader RNA including a (AAN)₄ repeat [76]. It is significant in this respect that Hfq has been shown to associate with Rho and mediate antitermination at Rho-dependent terminators both *in vivo* and *in vitro* [77]. Additional lines of evidence suggest that the ability of the three sRNAs to act as antiterminators is not limited to *rpoS* but applies to other putative mRNA targets [71]. More generally, one can speculate that interfering with Rho activity is a widespread strategy employed by sRNAs to function as positive regulators.

A recent study on the regulation of a CRISPR-Cas locus of *Pseudomonas* further supports the above conclusion. CRISPR-Cas loci direct the synthesis of an adaptive immune system that protects bacteria against infection by phages and other mobile genetic elements [78]. The CRISPR moiety comprises an array of short repeated sequences intercalated with unique sequences, (called spacers), that originate from the invader genome. The CRISPR-associated (Cas) genes encode the protein machinery involved in spacer capture and surveillance [78]. CRISPR arrays are transcribed in the form of a polycistronic precursor subsequently processed into small RNAs (crRNAs) that guide a Cas-encoded nuclease (in a complex with other Cas proteins) to the corresponding sequence (called the protospacer) when present in the invader genome and to promote its cleavage [79]. The structure and organization of the Cas genes as well as the number and relative orientation of the associated CRISPR arrays vary extensively even among isolates of the same bacterial species. *Pseudomonas aeruginosa* strain P14 carries two CRISPR arrays, CRISPR-1 and CRISPR-2, flanking the Cas gene cluster [80]. Recently, a high throughput search for sRNAs that might regulate CRISPR expression identified two sRNAs base-pairing with complementary sequences within the leader sequence of CRISPR-2 RNA precursor [81]. One of them was the sRNA PhrS previously known as a quorum sensing regulator in *Pseudomonas*. Further analysis revealed that PhrS activates CRISPR-2 transcription by preventing Rho-dependent termination in the CRISPR-2 leader region. Like in the *rpoS* regulation above, it is not clear whether the mechanism underlying PhrS action involves Rut site occlusion or inhibition of Rho translocation (Fig. 5). Regardless, PhrS antitermination activity appears critically important for CRISPR-2-mediated immunity. Deleting the *phrS* gene caused the *Pseudomonas* strain to become sensitive to infection by a phage normally targeted by a CRISPR-2 spacer suggesting that loss of PhrS completely prevents expression of the CRISPR-2 array; in contrast, the *phrS*-deleted strain remained resistant to a CRISPR-1-targeted phage [81]. The PhrS sRNA was originally characterized due to its ability to upregulate the *pqsR* gene encoding the activator of a quorum-sensing pathway in *Pseudomonas* [82]. PhrS stimulates translation of *pqsR* mRNA via a base-pairing interaction that engages the same seed sequence, the so-called *creg* motif, used for CRISPR 2 antitermination [82]. The recent demonstration that quorum sensing controls *Pseudomonas aeruginosa* CRISPR-Cas adaptive immunity [83] makes PhrS involvement in CRISPR-Cas regulation all of the more intriguing and interesting.

A further example of sRNA-mediated antitermination is found in the *rho* gene itself. In *E. coli*, Rho negatively regulates transcription of its own gene by terminating transcription at multiple sites within the 5' UTR and the initial translated portion of *rho* mRNA [84]. A similar feedback mechanism operates in *Salmonella* where finding that strains with *rho* mutations overproduce Rho protein suggests that a substantial level of termination occurs constitutively during the growth of wild-type cells [28]. A recent study identified *rho* as a regulatory target of the small RNA SraL in

Salmonella. SraL forms a base-pairing interaction with a 12 nt sequence approximately in the middle of the 255 nt long *rho* leader RNA. The interaction antagonizes Rho termination in the region and results in overproduction of full-length *rho* mRNA [85]. The SraL target sequence lies within a putative *Rut* site, suggesting that SraL directly competes with Rho for RNA binding. The biological role of the SraL action remains enigmatic. The sRNA accumulates in the stationary growth phase at a time when the *rho* mRNA is virtually undetectable [85]. Thus, SraL participation in *rho* regulation at this stage seems unlikely. Perhaps SraL is also induced under transient stress conditions that the Rho autogenous control mechanism might be slow in adapting to. For example, if a given stress negatively affects *rho* promoter activity, some time would be required before Rho protein levels are diluted enough to relieve attenuation. SraL could provide a quick way to compensate for a sudden drop in the transcription initiation rate by promoting transcriptional readthrough at the *rho* attenuator.

6. Rho-mediated regulatory mechanisms in Gram-positive bacteria

Most of the current understanding of Rho function and its involvement in gene regulatory mechanisms comes from studies with model *Enterobacteria* *E. coli* and *Salmonella*. In Gram-positive bacteria, relatively little is known concerning the details of Rho-dependent termination possibly because the protein, deemed non-essential, was not considered as relevant as in the Gram-negative group. For example, as early as in the nineteen seventies, it was already known that some Gram-positive bacteria are naturally resistant to the Rho inhibitor bicyclomycin [86]. The lack (or limited) impact of *rho* inactivation in representative species from both the Firmicutes and Actinobacteria taxa [87-89] also contributed to this idea. Furthermore, the *rho* gene is absent from branches of Firmicutes (most notably *Streptococci*) [90] where Rho-dependent termination is thus not a regulatory option. However, the reductive view of Rho function in Gram-positive bacteria has been reconsidered after finding that Rho is in fact essential in some relevant Actinobacteria (*Micrococcus luteus*, *Mycobacterium tuberculosis*) [29, 91] and that *rho* gene inactivation can affect cell differentiation programs in *Streptomyces lividans* [87] and *Bacillus subtilis* [27] or virulence gene expression in *Staphylococcus aureus* [92]. Furthermore, the *S. aureus* Rho protein can functionally replace *E. coli* Rho in complementation assays [89] and the Rho proteins from both *S. aureus* and *B. subtilis* are sensitive to bicyclomycin [89, 93]. Finally, recent transcriptomic studies conducted in *B. subtilis* [27], *S. aureus* [30], and *M. tuberculosis* [29] have revealed that *rho* inactivation causes genome-wide pervasive (mostly antisense) transcription in these species, a pattern remarkably similar to that found in *E. coli* [26] and *Salmonella* [28] despite significantly divergent genome organizations/compositions

(e.g., GC contents of ~33% for *S. aureus*, ~43% for *B. subtilis*, 50-52% for *E. coli* and *Salmonella*, and ~66% for *M. tuberculosis*).

The increasing recognition of the importance of Rho function in Gram-positive bacteria parallels our growing perception of the richness of RNA-based regulatory landscape and RNA degradation machinery in this bacterial group [94-96]. Given the high frequency of regulatory events involving 5' UTRs in *B. subtilis* [95], it is tempting to speculate that Rho-dependent and RNA-based regulatory mechanisms are intimately connected in Gram-positive bacteria as well. A recent study of the FMN riboswitch in *Corynebacterium glutamicum* lends support to this idea. In this species, expression of the riboflavin transporter *ribM* gene is controlled by an FMN riboswitch located in the *ribM* leader. The conformation adopted by the riboswitch in presence of its FMN ligand increases the susceptibility of the *ribM* mRNA towards degradation by RNase E/G but also promotes transcription termination through a Rho-dependent mechanism [97]. Furthermore, work in *Bacillus subtilis* has shown the participation of Rho-dependent transcriptional polarity in gene regulation. In the *trpEDCFBA* operon, binding of the regulatory protein TRAP to the 5' UTR of the mRNA (in the presence of excess tryptophan) promotes the formation of an array of secondary structures that attenuate transcription (by a Rho-independent mechanism) and block *trpE* translation in transcripts escaping attenuation [98]. The translational repression, besides disrupting *trpE-trpD* translational coupling, allows Rho to access the nascent transcript and cause polarity on the expression of the downstream genes [99]. By extrapolation, it seems reasonable to predict that the translation-targeting sRNAs will similarly cause transcriptional polarity in polycistronic mRNAs.

The predicted link between Rho and regulatory sRNA is likely to take novel twists in Gram-positive bacteria in light of their peculiarities as far as the role and/or mode of functioning of key auxiliary players are concerned. For instance, protein Hfq appears dispensable for sRNA-mediated regulation in Firmicutes [100, 101]. The DNA chaperone H-NS, which contributes to Rho-dependent termination at multiple loci in *E. coli* and *Salmonella* [26, 28] is absent in Firmicutes [102]. The contribution of NusG to Rho-dependent termination could also be different in the Gram-positive group. This possibility stems from the observation that while NusG reduces hairpin-dependent pausing and does not affect intrinsic termination in *E. coli* [11], it has opposite effects in *B. subtilis* [103] or *M. tuberculosis* [104] where it stimulates hairpin-dependent pausing and intrinsic termination, respectively. Finally, one should note that about one third of the Rho sequences available in databases exhibit a distinct feature consisting of a large polypeptide insertion in the N-terminus [90]. This feature is notably prevalent in (yet not exclusive to) the *Actinobacteria* phylum. Biochemical studies of the Rho factors from *M. luteus* and *M. tuberculosis* revealed that the arginine-rich insertion domain potentiates the capacity to utilize RNA [105, 106]. Rho variants lacking this domain are no longer able to deal with highly

structured transcripts (characteristic of GC-rich Actinobacteria transcriptomes) and to mediate promoter-proximal termination. In other species such as *Clostridium botulinum*, the insertion domain confers mysterious prion-like properties [107, 108]. Taken together, these observations call for further investigation and suggest that the study of RNA-mediated regulation in Gram-positive bacteria could well represent a fruitful area for Rho research in the years to come.

Acknowledgments

This work was supported by the French National Research Council (CNRS) and by the French National Research Agency (grants ANR-13-BSV3 -0005-01 and ANR-15-CE11-0024-03 to N.F.-B and ANR-13-BSV3 -0005-02 and ANR-15-CE11-0024-02 to M.B).

References

- [1] F. Imamoto, J. Ito, C. Yanofsky, Polarity in the tryptophan operon of *E. coli*, Cold Spring Harb Symp Quant Biol, 31 (1966) 235-249.
- [2] W.A. Newton, J.R. Beckwith, D. Zipser, S. Brenner, Nonsense mutants and polarity in the *lac* operon of *Escherichia coli*, J Mol Biol, 14 (1965) 290-296.
- [3] J. Beckwith, Restoration of Operon Activity by Suppressors, Biochim Biophys Acta, 76 (1963) 162-164.
- [4] J.P. Richardson, C. Grimley, C. Lowery, Transcription termination factor rho activity is altered in *Escherichia coli* with *suA* gene mutations, Proc Natl Acad Sci U S A, 72 (1975) 1725-1728.
- [5] J.W. Roberts, Termination factor for RNA synthesis, Nature, 224 (1969) 1168-1174.
- [6] K. McGary, E. Nudler, RNA polymerase and the ribosome: the close relationship, Curr Opin Microbiol, 16 (2013) 112-117.
- [7] S. Proshkin, A.R. Rahmouni, A. Mironov, E. Nudler, Cooperation between translating ribosomes and RNA polymerase in transcription elongation, Science, 328 (2010) 504-508.
- [8] B.M. Burmann, K. Schweimer, X. Luo, M.C. Wahl, B.L. Stitt, M.E. Gottesman, P. Rosch, A NusE:NusG complex links transcription and translation, Science, 328 (2010) 501-504.
- [9] G. Demo, A. Rasouly, N. Vasilyev, V. Svetlov, A.B. Loveland, R. Diaz-Avalos, N. Grigorieff, E. Nudler, A.A. Korostelev, Structure of RNA polymerase bound to ribosomal 30S subunit, Elife, 6 (2017).
- [10] R. Kohler, R.A. Mooney, D.J. Mills, R. Landick, P. Cramer, Architecture of a transcribing-translating expressome, Science, 356 (2017) 194-197.
- [11] A. Ray-Soni, M.J. Bellecourt, R. Landick, Mechanisms of Bacterial Transcription Termination: All Good Things Must End, Annu Rev Biochem, 85 (2016) 319-347.
- [12] N. Sedlyarova, P. Rescheneder, A. Magan, N. Popitsch, N. Rziha, I. Bilusic, V. Epshtein, B. Zimmermann, M. Lybecker, V. Sedlyarov, R. Schroeder, E. Nudler, Natural RNA Polymerase Aptamers Regulate Transcription in *E. coli*, Mol Cell, 67 (2017) 30-43 e36.
- [13] M.R. Lawson, K. Dyer, J.M. Berger, Ligand-induced and small-molecule control of substrate loading in a hexameric helicase, Proc Natl Acad Sci U S A, 113 (2016) 13714-13719.
- [14] E. Skordalakes, J.M. Berger, Structure of the Rho transcription terminator: mechanism of mRNA recognition and helicase loading, Cell, 114 (2003) 135-146.
- [15] N.D. Thomsen, M.R. Lawson, L.B. Witkowsky, S. Qu, J.M. Berger, Molecular mechanisms of substrate-controlled ring dynamics and substepping in a nucleic acid-dependent hexameric motor, Proc Natl Acad Sci U S A, 113 (2016) E7691-E7700.
- [16] C.E. Bogden, D. Fass, N. Bergman, M.D. Nichols, J.M. Berger, The structural basis for terminator recognition by the Rho transcription termination factor, Mol Cell, 3 (1999) 487-493.
- [17] N.D. Thomsen, J.M. Berger, Running in reverse: the structural basis for translocation polarity in hexameric helicases, Cell, 139 (2009) 523-534.

- [18] J.P. Richardson, Activation of rho protein ATPase requires simultaneous interaction at two kinds of nucleic acid-binding sites, *J Biol Chem*, 257 (1982) 5760-5766.
- [19] V. Epshtein, D. Dutta, J. Wade, E. Nudler, An allosteric mechanism of Rho-dependent transcription termination, *Nature*, 463 (2010) 245-249.
- [20] P. Alifano, F. Rivellini, D. Limauro, C.B. Bruni, M.S. Carlomagno, A consensus motif common to all Rho-dependent prokaryotic transcription terminators, *Cell*, 64 (1991) 553-563.
- [21] L.V. Richardson, J.P. Richardson, Rho-dependent termination of transcription is governed primarily by the upstream Rho utilization (rut) sequences of a terminator, *J Biol Chem*, 271 (1996) 21597-21603.
- [22] C. Nadiras, E. Eveno, A. Schwartz, N. Figueroa-Bossi, M. Boudvillain, A multivariate prediction model for Rho-dependent termination of transcription, *Nucleic Acids Res*, 46 (2018) 8245-8260.
- [23] R.A. Mooney, K. Schweimer, P. Rosch, M. Gottesman, R. Landick, Two structurally independent domains of *E. coli* NusG create regulatory plasticity via distinct interactions with RNA polymerase and regulators, *J Mol Biol*, 391 (2009) 341-358.
- [24] M.R. Lawson, W. Ma, M.J. Bellecourt, I. Artsimovitch, A. Martin, R. Landick, K. Schulten, J.M. Berger, Mechanism for the Regulated Control of Bacterial Transcription Termination by a Universal Adaptor Protein, *Mol Cell*, 71 (2018) 911-922 e914.
- [25] J.M. Peters, R.A. Mooney, P.F. Kuan, J.L. Rowland, S. Keles, R. Landick, Rho directs widespread termination of intragenic and stable RNA transcription, *Proc Natl Acad Sci U S A*, 106 (2009) 15406-15411.
- [26] J.M. Peters, R.A. Mooney, J.A. Grass, E.D. Jessen, F. Tran, R. Landick, Rho and NusG suppress pervasive antisense transcription in *Escherichia coli*, *Genes Dev*, 26 (2012) 2621-2633.
- [27] V. Bidnenko, P. Nicolas, A. Grylak-Mielnicka, O. Delumeau, S. Auger, A. Aucouturier, C. Guerin, F. Repoila, J. Bardowski, S. Aymerich, E. Bidnenko, Termination factor Rho: From the control of pervasive transcription to cell fate determination in *Bacillus subtilis*, *PLoS Genet*, 13 (2017) e1006909.
- [28] L. Bossi, M. Ratel, C. Laurent, P. Kerboriou, A. Camilli, E. Eveno, M. Boudvillain, N. Figueroa-Bossi, NusG prevents transcriptional invasion of H-NS-silenced genes, *PLoS Genet*, 15 (2019) e1008425.
- [29] L. Botella, J. Vaubourgeix, J. Livny, D. Schnappinger, Depleting *Mycobacterium tuberculosis* of the transcription termination factor Rho causes pervasive transcription and rapid death, *Nat Commun*, 8 (2017) 14731.
- [30] U. Mäder, P. Nicolas, M. Depke, J. Pane-Farre, M. Debarbouille, M.M. van der Kooi-Pol, C. Guerin, S. Derozier, A. Hiron, H. Jarmer, A. Leduc, S. Michalik, E. Reilman, M. Schaffer, F. Schmidt, P. Bessieres, P. Noirot, M. Hecker, T. Msadek, U. Volker, J.M. van Dijl, *Staphylococcus aureus* Transcriptome Architecture: From Laboratory to Infection-Mimicking Conditions, *PLoS Genet*, 12 (2016) e1005962.
- [31] A. Magan, F. Amman, F. El-Isa, N. Hartl, I. Shamovsky, E. Nudler, R. Schroeder, N. Sedlyarova, iRAPs curb antisense transcription in *E. coli*, *Nucleic Acids Res*, 47 (2019) 10894-10905.

- [32] S. Gottesman, G. Storz, Bacterial Small RNA Regulators: Versatile Roles and Rapidly Evolving Variations, *Cold Spring Harb Perspect Biol*, doi: 10.1101/cshperspect.a003798 (2011).
- [33] G. Storz, J. Vogel, K.M. Wassarman, Regulation by small RNAs in bacteria: expanding frontiers, *Mol Cell*, 43 (2011) 880-891.
- [34] E.G. Wagner, P. Romby, Small RNAs in bacteria and archaea: who they are, what they do, and how they do it, *Adv Genet*, 90 (2015) 133-208.
- [35] M. Kawano, A.A. Reynolds, J. Miranda-Rios, G. Storz, Detection of 5'- and 3'-UTR-derived small RNAs and cis-encoded antisense RNAs in *Escherichia coli*, *Nucleic Acids Res*, 33 (2005) 1040-1050.
- [36] M. Miyakoshi, Y. Chao, J. Vogel, Cross talk between ABC transporter mRNAs via a target mRNA-derived sponge of the GcvB small RNA, *EMBO J*, 34 (2015) 1478-1492.
- [37] J. Vogel, B.F. Luisi, Hfq and its constellation of RNA, *Nat Rev Microbiol*, 9 (2011) 578-589.
- [38] S. Panja, S.A. Woodson, Hfq proximity and orientation controls RNA annealing, *Nucleic Acids Res*, 40 (2012) 8690-8697.
- [39] A. Zhang, K.M. Wassarman, J. Ortega, A.C. Steven, G. Storz, The Sm-like Hfq protein increases OxyS RNA interaction with target mRNAs, *Mol Cell*, 9 (2002) 11-22.
- [40] M. Folichon, V. Arluison, O. Pellegrini, E. Huntzinger, P. Regnier, E. Hajnsdorf, The poly(A) binding protein Hfq protects RNA from RNase E and exoribonucleolytic degradation, *Nucleic Acids Res*, 31 (2003) 7302-7310.
- [41] I. Moll, T. Afonyushkin, O. Vytvytska, V.R. Kaberdin, U. Bläsi, Coincident Hfq binding and RNase E cleavage sites on mRNA and small regulatory RNAs, *RNA*, 9 (2003) 1308-1314.
- [42] T. Morita, R. Nishino, H. Aiba, Role of the terminator hairpin in the biogenesis of functional Hfq-binding sRNAs, *RNA*, 23 (2017) 1419-1431.
- [43] J. Chen, T. Morita, S. Gottesman, Regulation of Transcription Termination of Small RNAs and by Small RNAs: Molecular Mechanisms and Biological Functions, *Front Cell Infect Microbiol*, 9 (2019) 201.
- [44] N. Figueroa-Bossi, M. Valentini, L. Malleret, F. Fiorini, L. Bossi, Caught at its own game: regulatory small RNA inactivated by an inducible transcript mimicking its target, *Genes Dev*, 23 (2009) 2004-2015.
- [45] J. Plumbridge, L. Bossi, J. Oberto, J.T. Wade, N. Figueroa-Bossi, Interplay of transcriptional and small RNA-dependent control mechanisms regulates chitosugar uptake in *Escherichia coli* and *Salmonella*, *Mol Microbiol*, 92 (2014) 648-658.
- [46] L. Bossi, A. Schwartz, B. Guillemardet, M. Boudvillain, N. Figueroa-Bossi, A role for Rho-dependent polarity in gene regulation by a noncoding small RNA, *Genes Dev*, 26 (2012) 1864-1873.
- [47] T. Moller, T. Franch, C. Udesen, K. Gerdes, P. Valentin-Hansen, Spot 42 RNA mediates discoordinate expression of the *E. coli* galactose operon, *Genes Dev*, 16 (2002) 1696-1706.
- [48] X. Wang, S.C. Ji, H.J. Jeon, Y. Lee, H.M. Lim, Two-level inhibition of *galk* expression by Spot 42: Degradation of mRNA mK2 and enhanced transcription

- termination before the *galk* gene, Proc Natl Acad Sci U S A, 112 (2015) 7581-7586.
- [49] A.J. Carpousis, B.F. Luisi, K.J. McDowall, Endonucleolytic initiation of mRNA decay in *Escherichia coli*, Prog Mol Biol Transl Sci, 85 (2009) 91-135.
- [50] R.A. Spanjaard, J. van Duin, Translational reinitiation in the presence and absence of a Shine and Dalgarno sequence, Nucleic Acids Res, 17 (1989) 5501-5507.
- [51] D.S. Oppenheim, C. Yanofsky, Translational coupling during expression of the tryptophan operon of *Escherichia coli*, Genetics, 95 (1980) 785-795.
- [52] C.M. Sharma, F. Darfeuille, T.H. Plantinga, J. Vogel, A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites, Genes Dev, 21 (2007) 2804-2817.
- [53] C.M. Sharma, K. Papenfort, S.R. Pernitzsch, H.J. Mollenkopf, J.C. Hinton, J. Vogel, Pervasive post-transcriptional control of genes involved in amino acid metabolism by the Hfq-dependent GcvB small RNA, Mol Microbiol, 81 (2011) 1144-1165.
- [54] N. Figueroa-Bossi, S. Lemire, D. Maloriol, R. Balbontín, J. Casadesús, L. Bossi, Loss of Hfq activates the σ^E -dependent envelope stress response in *Salmonella enterica*, Mol Microbiol, 62 (2006) 838-852.
- [55] J. Greenblatt, J.R. Nodwell, S.W. Mason, Transcriptional antitermination, Nature, 364 (1993) 401-406.
- [56] N. Figueroa-Bossi, A. Schwartz, B. Guillemardet, F. D'Heygere, L. Bossi, M. Boudvillain, RNA remodeling by bacterial global regulator CsrA promotes Rho-dependent transcription termination, Genes Dev, 28 (2014) 1239-1251.
- [57] X. Wang, A.K. Dubey, K. Suzuki, C.S. Baker, P. Babitzke, T. Romeo, CsrA post-transcriptionally represses *pgaABCD*, responsible for synthesis of a biofilm polysaccharide adhesin of *Escherichia coli*, Mol Microbiol, 56 (2005) 1648-1663.
- [58] T. Romeo, P. Babitzke, Global Regulation by CsrA and Its RNA Antagonists, Microbiol Spectr, 6 (2018).
- [59] V. Stewart, R. Landick, C. Yanofsky, Rho-dependent transcription termination in the tryptophanase operon leader region of *Escherichia coli* K-12, J Bacteriol, 166 (1986) 217-223.
- [60] A.R. Gall, K.A. Datsenko, N. Figueroa-Bossi, L. Bossi, I. Masuda, Y.M. Hou, L.N. Csonka, Mg²⁺ regulates transcription of *mgtA* in *Salmonella* Typhimurium via translation of proline codons during synthesis of the MgtL peptide, Proc Natl Acad Sci U S A, 113 (2016) 15096-15101.
- [61] M.A. Kriner, E.A. Groisman, The Bacterial Transcription Termination Factor Rho Coordinates Mg(2+) Homeostasis with Translational Signals, J Mol Biol, 427 (2015) 3834-3849.
- [62] M.A. Kriner, E.A. Groisman, RNA secondary structures regulate three steps of Rho-dependent transcription termination within a bacterial mRNA leader, Nucleic Acids Res, 45 (2017) 631-642.
- [63] M.A. Kriner, A. Sevostyanova, E.A. Groisman, Learning from the Leaders: Gene Regulation by the Transcription Termination Factor Rho, Trends Biochem Sci, 41 (2016) 690-699.
- [64] L. Bischoff, O. Berninghausen, R. Beckmann, Molecular basis for the ribosome functioning as an L-tryptophan sensor, Cell Rep, 9 (2014) 469-475.

- [65] A.K. Martinez, N.H. Shirole, S. Murakami, M.J. Benedik, M.S. Sachs, L.R. Cruz-Vera, Crucial elements that maintain the interactions between the regulatory TnaC peptide and the ribosome exit tunnel responsible for Trp inhibition of ribosome function, *Nucleic Acids Res*, 40 (2012) 2247-2257.
- [66] G. Brandis, J.M. Bergman, D. Hughes, Autoregulation of the *tufB* operon in *Salmonella*, *Mol Microbiol*, 100 (2016) 1004-1016.
- [67] K. Hollands, S. Proshkin, S. Sklyarova, V. Epshtein, A. Mironov, E. Nudler, E.A. Groisman, Riboswitch control of Rho-dependent transcription termination, *Proc Natl Acad Sci U S A*, 109 (2012) 5376-5381.
- [68] D. Pedrolli, S. Langer, B. Hobl, J. Schwarz, M. Hashimoto, M. Mack, The *ribB* FMN riboswitch from *Escherichia coli* operates at the transcriptional and translational level and regulates riboflavin biosynthesis, *FEBS J*, 282 (2015) 3230-3242.
- [69] L. Bastet, A. Chauvier, N. Singh, A. Lussier, A.M. Lamontagne, K. Prevost, E. Massé, J.T. Wade, D.A. Lafontaine, Translational control and Rho-dependent transcription termination are intimately linked in riboswitch regulation, *Nucleic Acids Res*, 45 (2017) 7474-7486.
- [70] A. Chauvier, F. Picard-Jean, J.C. Berger-Dancause, L. Bastet, M.R. Naghdi, A. Dube, P. Turcotte, J. Perreault, D.A. Lafontaine, Transcriptional pausing at the translation start site operates as a critical checkpoint for riboswitch regulation, *Nat Commun*, 8 (2017) 13892.
- [71] N. Sedlyarova, I. Shamovsky, B.K. Bharati, V. Epshtein, J. Chen, S. Gottesman, R. Schroeder, E. Nudler, sRNA-Mediated Control of Transcription Termination in *E. coli*, *Cell*, 167 (2016) 111-121 e113.
- [72] R. Lange, D. Fischer, R. Hengge-Aronis, Identification of transcriptional start sites and the role of ppGpp in the expression of *rpoS*, the structural gene for the sigma S subunit of RNA polymerase in *Escherichia coli*, *J Bacteriol*, 177 (1995) 4676-4680.
- [73] P. Mandin, S. Gottesman, Integrating anaerobic/aerobic sensing and the general stress response through the ArcZ small RNA, *EMBO J*, 29 (2010) 3094-3107.
- [74] F. Repoila, N. Majdalani, S. Gottesman, Small non-coding RNAs, co-ordinators of adaptation processes in *Escherichia coli*: the RpoS paradigm, *Mol Microbiol*, 48 (2003) 855-861.
- [75] A. Sevostyanova, E.A. Groisman, An RNA motif advances transcription by preventing Rho-dependent termination, *Proc Natl Acad Sci U S A*, 112 (2015) E6835-6843.
- [76] Y. Peng, T.J. Soper, S.A. Woodson, Positional effects of AAN motifs in *rpoS* regulation by sRNAs and Hfq, *J Mol Biol*, 426 (2014) 275-285.
- [77] M. Rabhi, R. Tuma, M. Boudvillain, RNA remodeling by hexameric RNA helicases, *RNA Biol*, 7 (2010) 655-666.
- [78] L.A. Marraffini, E.J. Sontheimer, CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea, *Nat Rev Genet*, 11 (2010) 181-190.
- [79] L.A. Marraffini, CRISPR-Cas immunity in prokaryotes, *Nature*, 526 (2015) 55-61.
- [80] M.E. Zegans, J.C. Wagner, K.C. Cady, D.M. Murphy, J.H. Hammond, G.A. O'Toole, Interaction between bacteriophage DMS3 and host CRISPR region inhibits group behaviors of *Pseudomonas aeruginosa*, *J Bacteriol*, 191 (2009) 210-219.
- [81] P. Lin, Q. Pu, Q. Wu, C. Zhou, B. Wang, J. Schettler, Z. Wang, S. Qin, P. Gao, R. Li, G. Li, Z. Cheng, L. Lan, J. Jiang, M. Wu, High-throughput screen reveals sRNAs

- regulating crRNA biogenesis by targeting CRISPR leader to repress Rho termination, *Nat Commun*, 10 (2019) 3728.
- [82] E. Sonnleitner, N. Gonzalez, T. Sorger-Domenigg, S. Heeb, A.S. Richter, R. Backofen, P. Williams, A. Huttenhofer, D. Haas, U. Bläsi, The small RNA PhrS stimulates synthesis of the *Pseudomonas aeruginosa* quinolone signal, *Mol Microbiol*, 80 (2011) 868-885.
- [83] N.M. Hoyland-Kroghsbo, J. Paczkowski, S. Mukherjee, J. Broniewski, E. Westra, J. Bondy-Denomy, B.L. Bassler, Quorum sensing controls the *Pseudomonas aeruginosa* CRISPR-Cas adaptive immune system, *Proc Natl Acad Sci U S A*, 114 (2017) 131-135.
- [84] Y. Matsumoto, K. Shigesada, M. Hirano, M. Imai, Autogenous regulation of the gene for transcription termination factor rho in *Escherichia coli*: localization and function of its attenuators, *J Bacteriol*, 166 (1986) 945-958.
- [85] I.J. Silva, S. Barahona, A. Eyraud, D. Lalaouna, N. Figueroa-Bossi, E. Massé, C.M. Arraiano, SraL sRNA interaction regulates the terminator by preventing premature transcription termination of rho mRNA, *Proc Natl Acad Sci U S A*, 116 (2019) 3042-3051.
- [86] M. Nishida, Y. Mine, T. Matsubara, S. Goto, S. Kuwahara, Bicyclomycin, a new antibiotic. 3. *In vitro* and *in vivo* antimicrobial activity, *J Antibiot (Tokyo)*, 25 (1972) 582-593.
- [87] C.J. Ingham, I.S. Hunter, M.C. Smith, Isolation and sequencing of the rho gene from *Streptomyces lividans* ZX7 and characterization of the RNA-dependent NTPase activity of the overexpressed protein, *J Biol Chem*, 271 (1996) 21803-21807.
- [88] P.G. Quirk, E.A. Dunkley, Jr., P. Lee, T.A. Krulwich, Identification of a putative *Bacillus subtilis* rho gene, *J Bacteriol*, 175 (1993) 647-654.
- [89] R.S. Washburn, A. Marra, A.P. Bryant, M. Rosenberg, D.R. Gentry, rho is not essential for viability or virulence in *Staphylococcus aureus*, *Antimicrob Agents Chemother*, 45 (2001) 1099-1103.
- [90] F. D'Heygere, M. Rabhi, M. Boudvillain, Phyletic distribution and conservation of the bacterial transcription termination factor Rho, *Microbiology*, 159 (2013) 1423-1436.
- [91] W.L. Nowatzke, E. Keller, G. Koch, J.P. Richardson, Transcription termination factor Rho is essential for *Micrococcus luteus*, *J Bacteriol*, 179 (1997) 5238-5240.
- [92] A. Nagel, S. Michalik, M. Debarbouille, T. Hertlein, M. Gesell Salazar, H. Rath, T. Msadek, K. Ohlsen, J.M. van Dijl, U. Volker, U. Mader, Inhibition of Rho Activity Increases Expression of SaeRS-Dependent Virulence Factor Genes in *Staphylococcus aureus*, Showing a Link between Transcription Termination, Antibiotic Action, and Virulence, *MBio*, 9 (2018).
- [93] C.J. Ingham, J. Dennis, P.A. Furneaux, Autogenous regulation of transcription termination factor Rho and the requirement for Nus factors in *Bacillus subtilis*, *Mol Microbiol*, 31 (1999) 651-663.
- [94] S. Durand, A. Tomasini, F. Braun, C. Condon, P. Romby, sRNA and mRNA turnover in Gram-positive bacteria, *FEMS Microbiol Rev*, 39 (2015) 316-330.
- [95] R.A. Mars, P. Nicolas, E.L. Denham, J.M. van Dijl, Regulatory RNAs in *Bacillus subtilis*: a Gram-Positive Perspective on Bacterial RNA-Mediated Regulation of Gene Expression, *Microbiol Mol Biol Rev*, 80 (2016) 1029-1057.

- [96] S. Schwenk, K.B. Arnvig, Regulatory RNA in *Mycobacterium tuberculosis*, back to basics, *Pathog Dis*, 76 (2018).
- [97] N. Takemoto, Y. Tanaka, M. Inui, Rho and RNase play a central role in FMN riboswitch regulation in *Corynebacterium glutamicum*, *Nucleic Acids Res*, 43 (2015) 520-529.
- [98] P. Babitzke, Regulation of transcription attenuation and translation initiation by allosteric control of an RNA-binding protein: the *Bacillus subtilis* TRAP protein, *Curr Opin Microbiol*, 7 (2004) 132-139.
- [99] H. Yakhnin, J.E. Babiarez, A.V. Yakhnin, P. Babitzke, Expression of the *Bacillus subtilis* *trpEDCFBA* operon is influenced by translational coupling and Rho termination factor, *J Bacteriol*, 183 (2001) 5918-5926.
- [100] T. Rochat, P. Bouloc, Q. Yang, L. Bossi, N. Figueroa-Bossi, Lack of interchangeability of Hfq-like proteins, *Biochimie*, 94 (2012) 1554-1559.
- [101] T. Rochat, O. Delumeau, N. Figueroa-Bossi, P. Noirot, L. Bossi, E. Dervyn, P. Bouloc, Tracking the Elusive Function of *Bacillus subtilis* Hfq, *PLoS One*, 10 (2015) e0124977.
- [102] R.L. Ohniwa, Y. Ushijima, S. Saito, K. Morikawa, Proteomic analyses of nucleoid-associated proteins in *Escherichia coli*, *Pseudomonas aeruginosa*, *Bacillus subtilis*, and *Staphylococcus aureus*, *PLoS One*, 6 (2011) e19172.
- [103] A.V. Yakhnin, H. Yakhnin, P. Babitzke, Function of the *Bacillus subtilis* transcription elongation factor NusG in hairpin-dependent RNA polymerase pausing in the *trp* leader, *Proc Natl Acad Sci U S A*, 105 (2008) 16131-16136.
- [104] A. Czyz, R.A. Mooney, A. Iaconi, R. Landick, Mycobacterial RNA polymerase requires a U-tract at intrinsic terminators and is aided by NusG at suboptimal terminators, *MBio*, 5 (2014) e00931.
- [105] F. D'Heygere, A. Schwartz, F. Coste, B. Castaing, M. Boudvillain, ATP-dependent motor activity of the transcription termination factor Rho from *Mycobacterium tuberculosis*, *Nucleic Acids Res*, 43 (2015) 6099-6111.
- [106] W.L. Nowatzke, C.M. Burns, J.P. Richardson, Function of the novel subdomain in the RNA binding domain of transcription termination factor Rho from *Micrococcus luteus*, *J Biol Chem*, 272 (1997) 2207-2211.
- [107] I. Pallares, V. Iglesias, S. Ventura, The Rho Termination Factor of *Clostridium botulinum* Contains a Prion-Like Domain with a Highly Amyloidogenic Core, *Front Microbiol*, 6 (2015) 1516.
- [108] A.H. Yuan, A. Hochschild, A bacterial global regulator forms a prion, *Science*, 355 (2017) 198-201.

Legends to the Figures

Fig. 1. Mechanism of transcriptional polarity. (A) During transcription of a protein-coding gene, the leading ribosome is physically connected to the transcription elongation complex *via* the interaction between ribosomal protein S10 and the C-terminal domain of NusG. The close proximity of the ribosome to RNA polymerase prevents Rho factor from gaining access to the nascent RNA. (B) If protein synthesis stops prematurely due to a nonsense mutation, the nascent RNA is no longer covered by the lead ribosome. Rho can load onto the first encountered *Rut* site (yellow dots) and promote termination through a step stimulated by the interaction with the C-terminal domain of NusG.

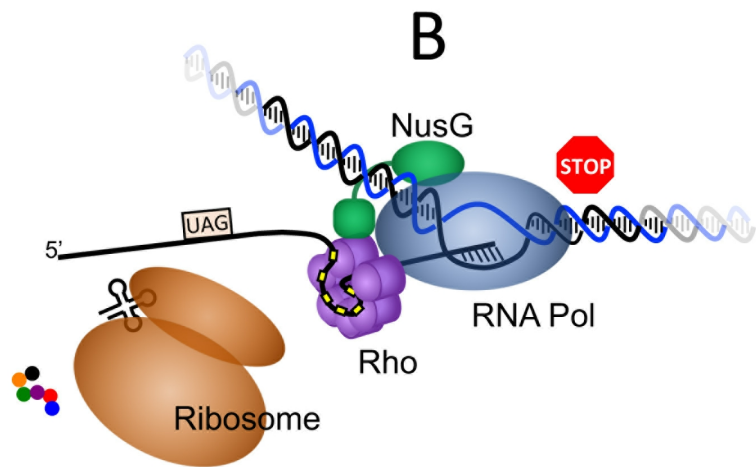
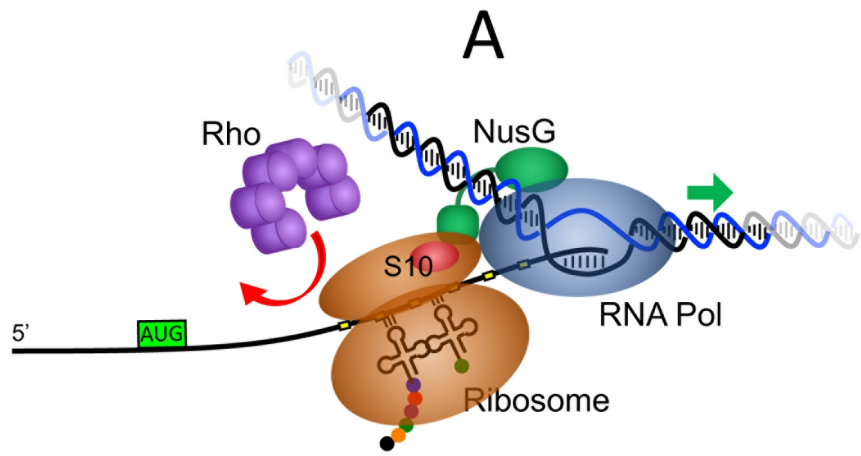
Fig. 2. Small regulatory RNA promotes Rho-dependent termination. sRNA base-pairing with a complementary sequence in the 5' UTR of the target mRNA occludes the ribosome-binding site. No longer shielded by the lead ribosomes, the *Rut* site (yellow dots) becomes accessible to the binding of Rho, which terminates transcription.

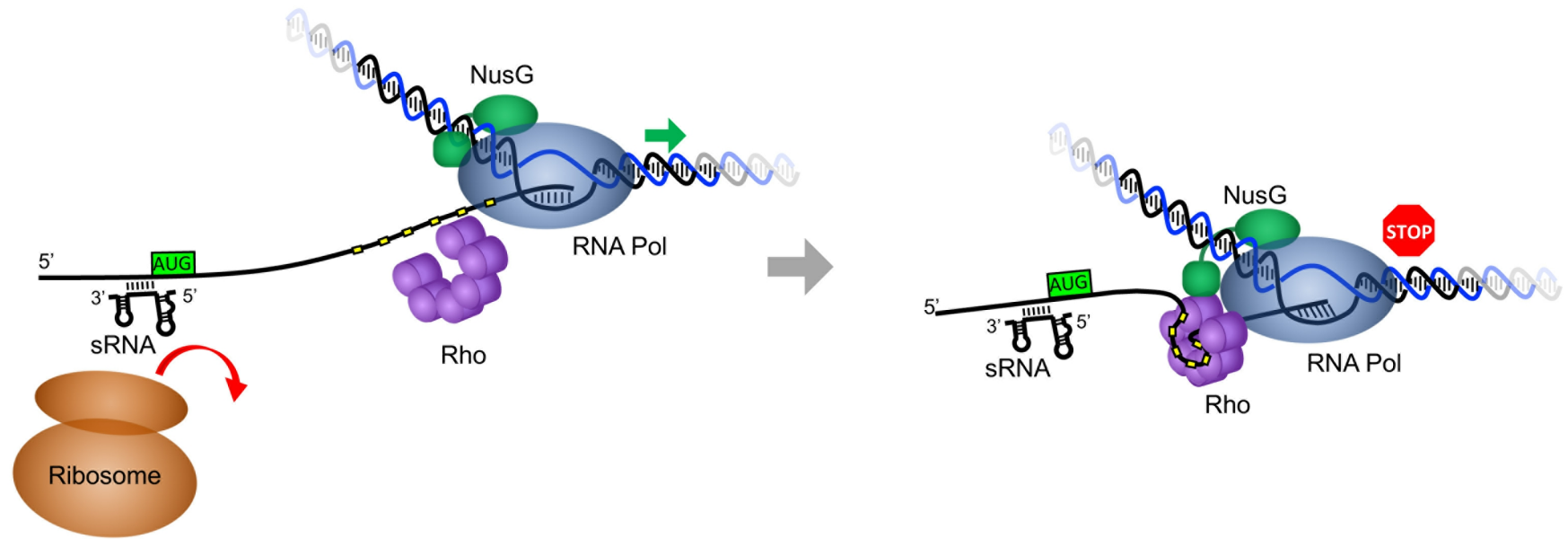
Fig. 3. (A) Cytosine / Guanine frequencies in the initial portion of *chiP* mRNA from *Salmonella* and *E. coli*. The C>G bubble (red arrow) corresponds to the *Rut* site. (B) Nucleotide sequences of the *Rut* site. The CC/UC repeats that make up the *Rut* site are highlighted in yellow. Brackets denote 11-12 nt spacing between the repeats, consistent with the binding of four adjacent Rho subunits (which could be sufficient for productive Rho-*Rut* interaction [13, 22]). Note that the sequence from *E. coli* contains two alternative sets of appropriately spaced repeats (Blue and Purple brackets). The bars above the sequences mark the positions of the 3' ends of the most prominent *in vitro* termination products with Rho alone (black bar) or with Rho and NusG (red bar). No termination occurs with the *Salmonella* template in the presence of Rho alone.

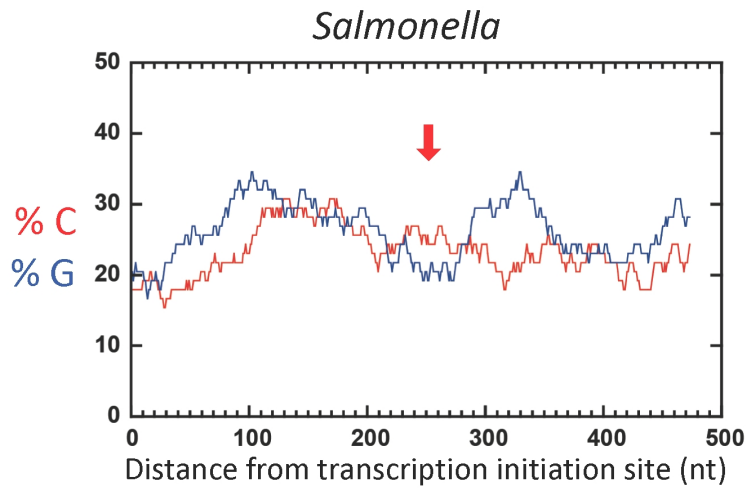
Fig. 4. Models for the effects of sRNA repression on a polycistronic transcript. Depicted are the first two cistrons of a hypothetical polycistronic mRNA. (A) In the absence of the sRNA, translation of cistron 1 is unhampered and stimulates translation of cistron 2 (translational coupling). This stimulation results from two mechanisms: *i*) ribosomes translating cistron 1 remove higher order structures that occlude the ribosome-binding site of cistron 2; *ii*): ribosomes reaching the end of cistron 1 can reinitiate translation of cistron 2. (B) In the presence of an sRNA blocking cistron 1 translation, the nascent RNA, no longer protected by ribosomes, is cleaved by RNase E and further degraded by 3' to 5' exonucleases. As transcription elongation continues, the ribosome-binding site of cistron 2 will emerge from the exit channel of RNA polymerase. Unless RNase E cleaves this sequence before

ribosomes can bind to it, cistron 2 will continue to be translated. (C) Repression of cistron 1 translation by the sRNA causes the nascent mRNA to fold into higher order structures that occlude cistron 2 ribosome-binding site. Cistron 2 will no longer be translated. (D) Repression of cistron 1 translation by the sRNA disrupts the translational coupling between cistron 1 and cistron 2. A *Rut* site in cistron 2 becomes exposed. Rho binds to this site and terminates transcription. All of the downstream cistrons are no longer expressed. Note that the three models above are not mutually exclusive. In particular, the untranslated portions of cistrons 1 and 2 mRNA in C and D will likely undergo degradation as depicted in B.

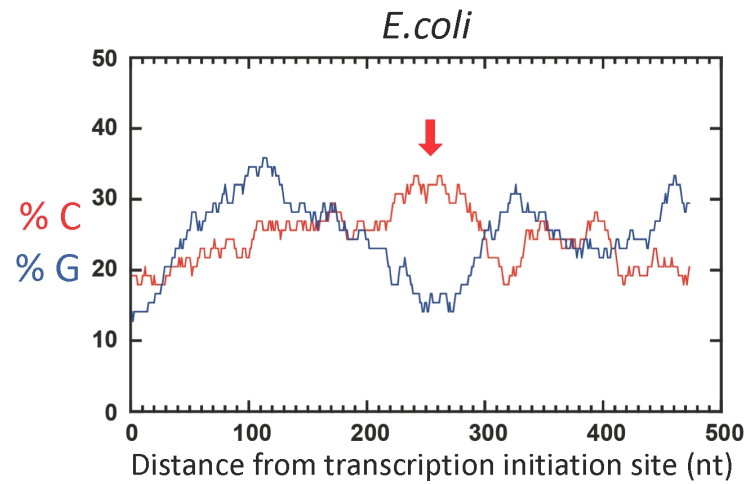
Fig. 5. Model for antitermination by sRNA. (A) The sRNA target sequence is part of the *Rut* site. The sRNA competes with Rho for binding to the RNA. If the sRNA binds first, formation of the sRNA:mRNA duplex prevents Rho binding and results in antitermination. (B) The sRNA target sequence is immediately downstream from the *Rut* site. Rho can load onto the RNA but its subsequent translocation is hindered by the presence of the sRNA:mRNA duplex. The bump gives RNA polymerase enough time to move ahead and exit the termination window.







A

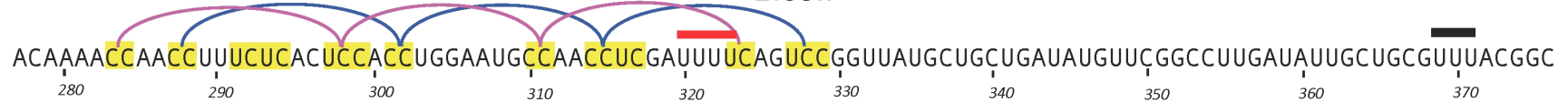


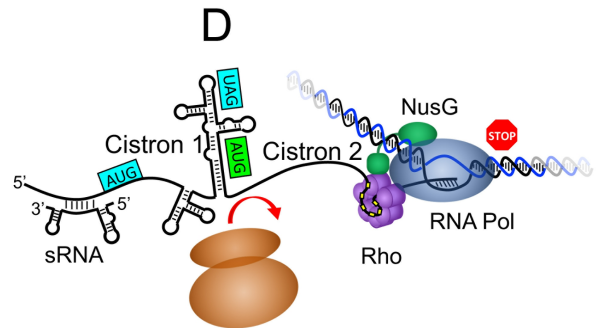
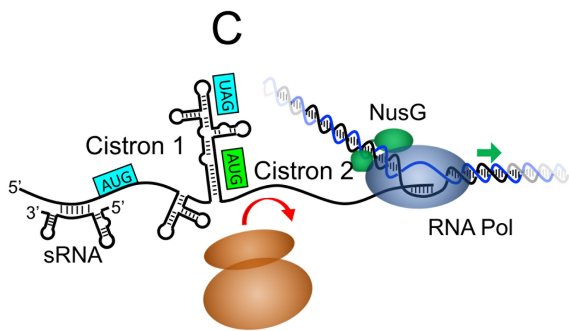
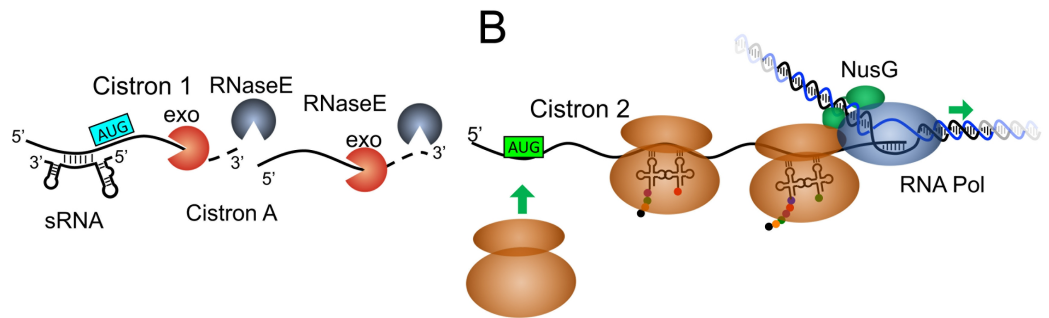
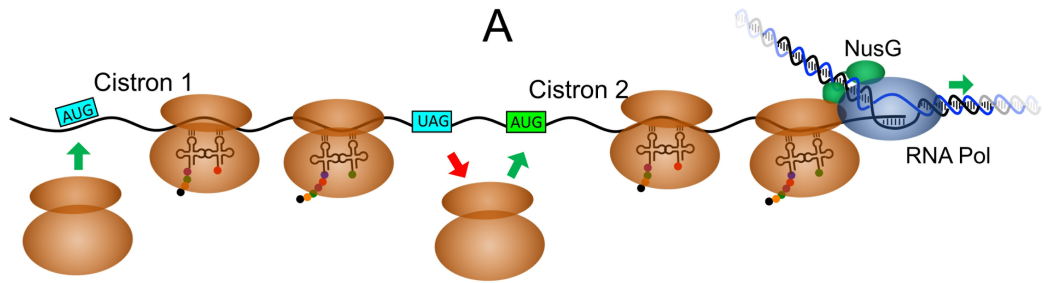
B

Salmonella

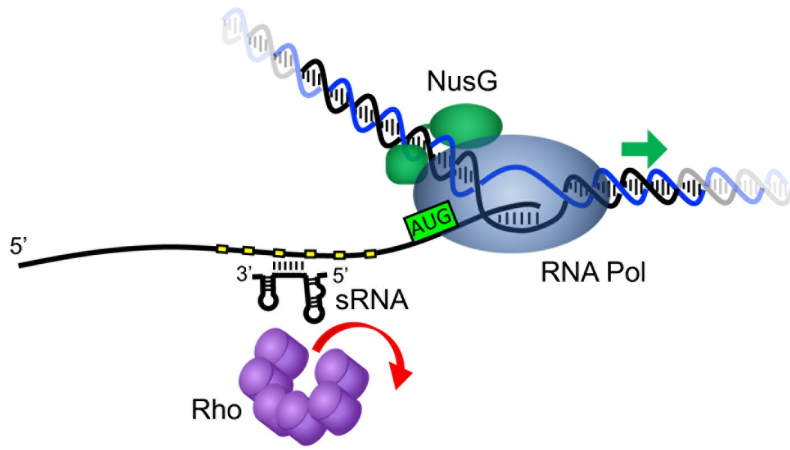


E.coli

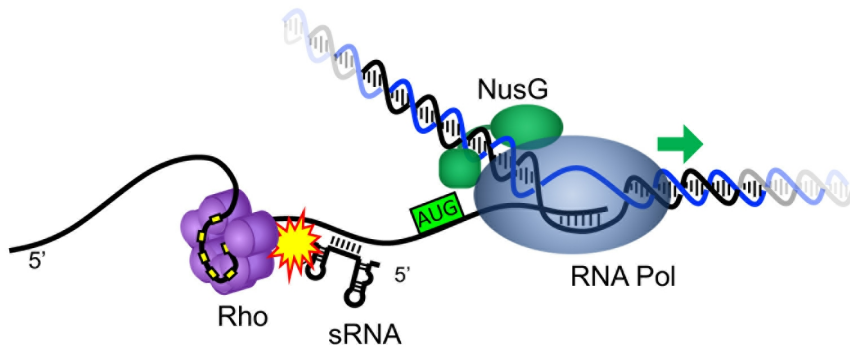




A



B



Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

CRedit author statement

Lionello Bossi: Original draft preparation, Writing-Reviewing and Editing **Nara Figueroa-Bossi:** Writing-Reviewing and Editing **Philippe Boulloc:** Writing-Reviewing and Editing **Marc Boudvillain:** Writing-Reviewing and Editing.