Current Genetics

3' untranslated regions, regulation at the end of the road.

--Manuscript Draft--

Manuscript Number:	CUGE-D-18-00164R1
Full Title:	3' untranslated regions, regulation at the end of the road.
Article Type:	Mini-Review
Corresponding Author:	Carlos Balsalobre Universitat de Barcelona Barcelona, SPAIN
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Universitat de Barcelona
Corresponding Author's Secondary Institution:	
First Author:	Youssef El Mouali
First Author Secondary Information:	
Order of Authors:	Youssef El Mouali
	Carlos Balsalobre
Order of Authors Secondary Information:	
Funding Information:	Secretaría de Estado de Investigación, Desarrollo e Innovación (AGL2013-45339R) — Dr. Carlos Balsalobre
	Departament d'Universitats, Recerca i Societat de la Informació (2017SGR499) — Dr. Carlos Balsalobre
Abstract:	Post-transcriptional gene regulation in bacteria plays a major role in the adaptation of bacterial cells to the changing conditions encountered in the environment. In bacteria, most of the regulation at the level of mRNA seem to be targeting the 5'untranslated regions where accessibility to the ribosome-binding site can be modulated to alter gene expression. In recent years, the role of 3'untranslated regions has gained attention also as a site for post-transcriptional regulation. In addition to be a source of trans-encoded small RNAs, the 3'untranslated regions can be targets to modulate gene expression. Taking recent findings in the post-transcriptional regulation of the hilD gene, encoding for the main regulator of virulence in Salmonella enterica serovar Typhimurium, we highlight the role of 3'untranslated regions as targets of post-transcriptional regulation mediated by small RNAs and discuss the implications of transcriptional elongation in the 3'UTR-mediated regulation in bacteria.

3'untranslated regions, regulation at the end of the road.

Youssef El Mouali[1] and Carlos Balsalobre[2]*

[1]RNA Biology Group, Institute for Molecular Infection Biology, University of Würzburg, Würzburg (Germany), [2]Department of Genetics, Microbiology and Statistics, School of Biology, University of Barcelona, Barcelona (Spain).

* Corresponding author. E-mail: cbalsalobre@ub.edu, telephone: +34934034622, ORCID: 0000-0002-4147-219X

Keywords: post-transcriptional regulation, 3'UTR, Hfq, sRNA, ProQ, transcription elongation

Abstract

Post-transcriptional gene regulation in bacteria plays a major role in the adaptation of bacterial cells to the changing conditions encountered in the environment. In bacteria, most of the regulation at the level of mRNA seem to be targeting the 5'untranslated regions where accessibility to the ribosome-binding site can be modulated to alter gene expression. In recent years, the role of 3'untranslated regions has gained attention also as a site for post-transcriptional regulation. In addition to be a source of trans-encoded small RNAs, the 3'untranslated regions can be targets to modulate gene expression. Taking recent findings in the post-transcriptional regulation of the *hilD* gene, encoding for the main regulator of virulence in *Salmonella enterica* serovar Typhimurium, we highlight the role of 3'untranslated regions as targets of post-transcriptional regulation mediated by small RNAs and discuss the implications of transcriptional elongation in the 3'UTR-mediated regulation in bacteria.

Introduction

Bacterial cells, as unicellular organisms, lack the homeostasis present in more complex life systems. When encountering changing environmental conditions, bacteria need to alter their gene expression profile in a fast and accurate manner in order to adapt. Gene expression is a multistep process, while transcription and translation are tightly regulated processes, the stability and activity of the resulting molecules, RNA and proteins, can also be modulated to finely define the proper expression level of the active products. For decades, most expression studies focused on transcriptional regulation, more specifically regulation at the level of transcription initiation. However, rapid changes in the final output of gene expression are more easily achieved when modulatory mechanisms act at later steps. In this context, post-transcriptional regulation provides a plethora of mechanisms that contribute to the tight regulation of gene expression. Its contribution in the control of complex bacterial processes such as intracellular survival, biofilm and persister cell formation has been shown (Mika and Hengge 2014; Quereda and Cossart 2017; Berghoff and Wagner 2017).

Many mRNA molecules carry, in addition to the coding sequence, long untranslated extensions at the 5' and/or 3' ends. In bacteria, it is well established that the 5'untranslated regions (5'UTRs) of the mRNA constitute hubs for post-transcriptional regulation. Different mechanisms, mediated by metabolites, proteins and/or small

RNAs (sRNAs), target the 5'UTR of mRNAs to modulate gene expression (for review see Ignatov and Johansson 2017; Kavita et al. 2018; Holmqvist and Vogel 2018). By contrast, the role of the 3' untranslated regions (3'UTRs) in gene regulation has so far been underestimated. The molecular characterization of the mechanism targeting the *hilD* mRNA from *Salmonella* will be used as an example to discuss the role of mRNA 3'UTRs in post-transcriptional regulation in bacteria.

mRNA 3'UTR, last but not least

The role of the 3'UTRs in prokaryotic RNA biology has gained attention in the last years. Recent studies propose that 3'UTRs in many mRNAs are involved in post-transcriptional regulation where they can modulate gene expression in *cis*, affecting expression from the own mRNA, or in *trans*, altering expression of other loci (as reviewed in Ren et al. 2017).

So far, the main role attributed to 3'UTRs is to act as sRNAs reservoirs. Although it was assumed that most trans-encoded sRNAs are located in intergenic regions, it has been shown that sRNAs can be also encoded in the 3'UTR of certain mRNAs (Kawano et al. 2005; Chao et al. 2012). The resulting sRNAs are generated either by processing from the mRNA or by transcription from an internal promoter within the ORF (for review see Miyakoshi et al. 2015). Although it is not known how ubiquitous is the presence of sRNAs in the mRNA 3'UTRs, some data suggest that it can be highly extended. In Gram-negative bacteria, most characterized sRNAs require association with the global RNA chaperone Hfq (for review see Kavita et al. 2018). Consecutive global analysis of Hfq-bound RNA in *Salmonella* led to the identification of Hfq binding sites within the 3'UTR of many mRNAs (Chao et al. 2012; Holmqvist et al. 2016). Some identified Hfq-bound RNA have been characterized as 3'UTRs that, upon processing, generate functional trans-encoded sRNAs (Chao and Vogel 2016). Interestingly, a putative sRNA (STnc600) was annotated within the *hilD* 3'UTR sequence although its existence could not be confirmed by Northern blot (Sittka et al. 2008).

Some studies also indicate that 3'UTRs can modulate gene expression by acting in *cis*. A direct role for the 3'UTRs influencing protein synthesis was described with the *icaR* mRNA of *Staphylococcus aureus*. The long 3'UTR of *icaR* mRNA interacts with its 5'UTR to regulate translation (Ruiz de los Mozos et al. 2013). A distinct example of *cis*-acting 3'UTR is found in *Salmonella hilD* mRNA, as extensively discussed in the

next section (López-Garrido et al. 2014).**The *hilD* 3'UTR, a target for sRNA regulation.**

The main role attributed to the *hilD* 3'UTR is to regulate *hilD* mRNA levels, since its deletion leads to an increase of *hilD* transcript levels (López-Garrido et al. 2014). To shed light in the mechanism of action of the *hilD* 3'UTR, it should be noted that the RNA chaperone Hfq is required for proper *hilD* expression and it has two interaction sites within the 3'UTR of *hilD* mRNA (Sittka et al. 2007; Holmqvist et al. 2016). Accordingly, genetic analyses demonstrated that the 3'UTR of the *hilD* mRNA is the target sequence for Hfq's contribution to *hilD* expression (López-Garrido et al. 2014). Having in consideration that Hfq assists many sRNA-mediated regulation processes, these observations raised the question whether the *hilD* 3'UTR and, by extension, other mRNA 3'UTRs, can be regulatory domains targeted by sRNAs.

Are 3'UTRs, similar to 5'UTRs, common regulatory targets for trans-encoded sRNAs? Most characterized sRNAs targeting 5'UTRs interact with their target mRNA by base pairing in an Hfq-dependent manner, as was also shown by a recent study describing the global sRNA-mRNA interactome of Hfq in *Escherichia coli* (Melamed et al. 2016). Interactions between 3'UTRs of mRNAs and sRNAs were also detected. Still, whether these interactions correspond to sRNAs targeting mRNAs at their 3'UTR or sRNAs targeting 3'UTR-derived sRNAs requires further characterization. In addition to Hfq, the two other major RNA-binding proteins - CsrA and ProQ - also seem to bind mRNA 3'UTRs. A low number of interaction sites for CsrA within mRNA 3'UTRs were identified both in *E. coli* and *S. enterica* (Holmqvist et al. 2016; Potts et al. 2017). Interestingly, ProQ seems to have a preference for binding to 3'UTRs of mRNAs, pointing towards a defined role of 3'UTRs in post-transcriptional regulation and existing differences in the mechanisms associated to 5' and 3'UTR-mediated regulation (Holmqvist et al. 2018).

hilD mRNA is target of the main RNA chaperones in the cell. In addition to the Hfq binding previously described, it is well known that CsrA binds to its 5'UTR and recently new CsrA binding sites have been reported at the coding sequence upstream of the 3'UTR (Martínez et al. 2011; Holmqvist et al. 2016). Remarkably, it was found that ProQ binds within the *hilD* 3'UTR (annotated as STnc600), near the *hilD* mRNA terminator (Holmqvist et al. 2018) (Fig 1). Overall, the fact that the *hilD* mRNA is targeted by the three major RNA binding proteins strongly indicates that its expression

is under tight post-transcriptional regulation. In a recent study, we described that the sRNA Spot 42 positively regulates the expression of *hilD* through physical interaction with the 3'UTR of *hilD* mRNA. Although the exact interaction site of Spot 42 within *hilD* 3'UTR remains elusive, our findings constitute the first description of a trans-encoded sRNA targeting an mRNA through its 3'UTR (El Mouali et al. 2018).

Some insights on the molecular mechanisms targeting 3'UTRs suggest that RNAses seem to play a pivotal role in post-transcriptional regulation mediated through mRNA 3'UTRs. The 3'UTR of the *gadX* mRNA is the target sequence for the cis-encoded antisense sRNA GadY, a mechanism dependent on the activity of the endonuclease RNAse III (Opdyke et al. 2004, 2011). It is also known that interaction of ProQ with the 3'UTR of *cspE* prevents cleavage by the exonuclease RNAse II (Holmqvist et al. 2018). In the case of *hilD* mRNA, RNAse E seem to play a role in the mechanism by which the *hilD* 3'UTR modulates the levels of *hilD* mRNA (López-Garrido et al. 2014; El Mouali et al. 2018). It has been suggested that mRNA degradation promoted from the 3'UTR might provide a faster response to shut down gene expression as compared to regulation at the 5'UTR (Ren et al. 2017). Blocking of the ribosome-binding site at the 5'UTR by sRNAs only affects translation initiation but it does not abolish the production of proteins whose synthesis was already initiated. However, cleavage of the mRNA through its 3'UTR would block translation elongation and therefore have a sudden response to environmental stimuli. This promptness could be especially critical during modulation of the expression of global regulators, such as the major virulence regulator HilD in *Salmonella*.

Is RNA structure pivotal for the function of the 3'UTRs?

The structure of an mRNA is often crucial for its proper expression. It is well-known that the structure of 5'UTRs might have major effects in post-transcriptional regulation. As mentioned, interaction of 5'UTRs with proteins, sRNAs or metabolites can alter the 5'UTR structure leading to activation or repression of the downstream coding sequence. Regulatory features such as riboswitches and RNA thermometers in 5'UTRs are good examples (for review see Ignatov and Johansson 2017; Loh et al. 2018). mRNA folding is mainly accomplished during transcription elongation (Pan and Sosnick 2006). Elongation kinetics can be relevant during RNA folding. It has been shown that transcriptional pausing assist folding of 5'UTR and non-coding RNAs

(Wong et al. 2007; Nechooshtan et al. 2014) and can coordinate RNA folding and ligand binding in riboswitches (Perdrizet et al. 2012).

In a previous study from our lab, we described that the transcription elongation factors GreA and GreB are involved in *hilD* regulation. Remarkably, the transcriptional pause rescued by the Gre factors occurs within the first 100 nt of the *hilD* 3'UTR (Gaviria-Cantin et al. 2017). The *hilD* 3'UTR is 310 nt long and although the first 100 nt play a major role in the overall negative effect on HilD expression, the interaction of Spot 42, Hfq and ProQ occurs far downstream in the *hilD* 3'UTR (Holmqvist et al. 2016, 2018; El Mouali et al. 2018).

One could speculate that transcriptional pausing nearby UTRs, might affect the mRNA structure and eventually its post-transcriptional regulation. In this scenario, the transcriptional pause within the first 100 nt of the *hilD* 3'UTR could affect mRNA folding, having effects in the overall structure of the *hilD* 3'UTR and, consequently, in the recruitment of regulatory factors such as sRNAs and RNA chaperones. Additional studies will be required for further characterization. For instance, changes in the global structure of mRNAs at nucleotide resolution can be monitored *in vivo* by dimethyl sulfate sequencing (DMS-seq) (Ding et al. 2015). Assessment of the structure of mRNAs in presence and absence of Gre factors could provide information on the changes in the *hilD* 3'UTR structure as well as in the structure of other mRNAs transcribed from genes containing transcriptional pauses. Additionally, it is well known that *Salmonella* cells display virulence gene expression heterogeneity during infection, whether the *hilD* 3'UTR, as target of transcriptional elongation factors, RNA binding proteins and sRNAs, plays a role in generating gene expression heterogeneity remains to be elucidated (Ackermann 2015; Evans and Ling 2018).

Conclusions

mRNA 3'UTRs are involved in post-transcriptional regulation, as trans-acting elements as reservoir of trans-encoded sRNAs or as cis-acting elements regulating either translation or the steady-state level of the transcript. Our recent work showed that mRNA 3'UTRs can act as targets for post-transcriptional regulation mediated by trans-encoded sRNAs. The extent of the 3'UTR-mediated regulation of gene expression in bacteria is still unknown and future studies will clarify if the regulation through mRNAs

3’UTR is a general and pivotal step during modulation of gene expression. In this regard, it was shown in *Staphylococcus aureus* that up to one third of the mRNAs carry 3’UTRs longer than 100 nt (Ruiz de los Mozos et al. 2013). Moreover, further studies on the mechanisms by which mRNA 3’UTRs are targeted at the post-transcriptional level will shed light on whether the regulation occurs similarly to 5’UTRs or distinct mechanisms are also involved. It is stimulating that mechanisms based on structure motifs rather than sequence motifs might drive 3’UTR-mediated post-transcriptional regulation. We highlight here, that mRNA 3’UTRs can be regulatory hubs targeted by trans-encoded sRNAs through physical interaction leading to gene expression regulation.

Figure legends

Fig. 1 Summary of the transcriptional and post-transcriptional regulation of *hilD* expression. The transcriptional expression of the *hilD* gene is tightly regulated. In the mRNA, the 5' and 3' UTRs are labelled in black and the coding sequence in orange. During transcription elongation, GreA and/or GreB are required to alleviate a transcriptional pause that occurs in the 3'UTR. The *hilD* transcript is under post-transcriptional regulation targeting both the 5'UTR and the 3'UTR. Several factors as the RNA binding proteins CsrA, Hfq and ProQ and the sRNA Spot 42 have been involved in *hilD* post-transcriptional regulation. When a factor has an overall effect stimulating or repressing *hilD* expression is indicated with a + or – sign. If the effect is unknown is indicated with a question mark.

Acknowledgments

We thank Jens Hör for critical reading of the manuscript. This work was supported by the Spanish Minister of Economy and Competitiveness (grant AGL2013-45339R) and the Catalonian Government (grant 2017SGR499).

References

Ackermann M (2015) A functional perspective on phenotypic heterogeneity in microorganisms. Nat Rev Microbiol 13:497–508. doi: 10.1038/nrmicro3491

Berghoff BA, Wagner EGH (2017) RNA-based regulation in type I toxin-antitoxin systems and its implication for bacterial persistence. Curr Genet 63:1011–1016. doi: 10.1007/s00294-017-0710-y

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

Chao Y, Vogel J (2016) A 3′ UTR-Derived Small RNA Provides the Regulatory Noncoding Arm of the Inner Membrane Stress Response. Mol Cell 61:352–63. doi: 10.1016/j.molcel.2015.12.023

Ding Y, Kwok CK, Tang Y, et al (2015) Genome-wide profiling of in vivo RNA structure at single-nucleotide resolution using structure-seq. Nat Protoc 10:1050–1066. doi: 10.1038/nprot.2015.064

El Mouali Y, Gaviria-Cantin T, Sánchez-Romero MA, et al (2018) CRP-cAMP

mediates silencing of *Salmonella* virulence at the post-transcriptional level. PLOS Genet 14:e1007401. doi: 10.1371/journal.pgen.1007401

Evans CR, Ling J (2018) Visualizing translational errors: one cell at a time. Curr Genet 64:551–554. doi: 10.1007/s00294-017-0784-6

Gaviria-Cantin T, El Mouali Y, Le Guyon S, et al (2017) Gre factors-mediated control of *hilD* transcription is essential for the invasion of epithelial cells by *Salmonella enterica* serovar Typhimurium. PLOS Pathog 13:e1006312. doi: 10.1371/journal.ppat.1006312

Holmqvist E, Li L, Bischler T, et al (2018) Global Maps of ProQ Binding In Vivo Reveal Target Recognition via RNA Structure and Stability Control at mRNA 3′ Ends. Mol Cell 70:971–982.e6. doi: 10.1016/j.molcel.2018.04.017

Holmqvist E, Vogel J (2018) RNA-binding proteins in bacteria. Nat Rev Microbiol. doi: 10.1038/s41579-018-0049-5

Holmqvist E, Wright PR, Li L, et al (2016) Global RNA recognition patterns of post-transcriptional regulators Hfq and CsrA revealed by UV crosslinking *in vivo*. EMBO J 35:991–1011. doi: 10.15252/embj.201593360

Ignatov D, Johansson J (2017) RNA-mediated signal perception in pathogenic bacteria. Wiley Interdiscip Rev RNA 8:e1429. doi: 10.1002/wrna.1429

Kavita K, de Mets F, Gottesman S (2018) New aspects of RNA-based regulation by Hfq and its partner sRNAs. Curr Opin Microbiol 42:53–61. doi: 10.1016/j.mib.2017.10.014

Kawano M, Reynolds AA, Miranda-Rios J, Storz G (2005) Detection of 5'- and 3'-UTR-derived small RNAs and cis-encoded antisense RNAs in *Escherichia coli.* Nucleic Acids Res 33:1040–50. doi: 10.1093/nar/gki256

Loh E, Righetti F, Eichner H, et al (2018) RNA Thermometers in Bacterial Pathogens. Microbiol Spectr 6:. doi: 10.1128/microbiolspec.RWR-0012-2017

López-Garrido J, Puerta-Fernández E, Casadesús J (2014) A eukaryotic-like 3' untranslated region in *Salmonella enterica hilD* mRNA. Nucleic Acids Res 42:5894–906. doi: 10.1093/nar/gku222

Martínez LC, Yakhnin H, Camacho MI, et al (2011) Integration of a complex regulatory cascade involving the SirA/BarA and Csr global regulatory systems that controls expression of the *Salmonella* SPI-1 and SPI-2 virulence regulons through HilD. Mol Microbiol 80:1637–56. doi: 10.1111/j.1365-2958.2011.07674.x

Melamed S, Peer A, Faigenbaum-Romm R, et al (2016) Global Mapping of Small

RNA-Target Interactions in Bacteria. Mol Cell 63:884–897. doi: 10.1016/j.molcel.2016.07.026

Mika F, Hengge R (2014) Small RNAs in the control of RpoS, CsgD, and biofilm architecture of *Escherichia coli*. RNA Biol 11:494–507. doi: 10.4161/rna.28867

Miyakoshi M, Chao Y, Vogel J (2015) Regulatory small RNAs from the 3′ regions of bacterial mRNAs. Curr Opin Microbiol 24:132–139. doi: 10.1016/j.mib.2015.01.013

Nechooshtan G, Elgrably-Weiss M, Altuvia S (2014) Changes in transcriptional pausing modify the folding dynamics of the pH-responsive RNA element. Nucleic Acids Res 42:622–630. doi: 10.1093/nar/gkt868

Opdyke JA, Fozo EM, Hemm MR, Storz G (2011) RNase III Participates in GadY-Dependent Cleavage of the *gadX-gadW* mRNA. J Mol Biol 406:29–43. doi: 10.1016/j.jmb.2010.12.009

Opdyke JA, Kang J-G, Storz G (2004) GadY, a Small-RNA Regulator of Acid Response Genes in *Escherichia coli*. J Bacteriol 186:6698–6705. doi: 10.1128/JB.186.20.6698-6705.2004

Pan T, Sosnick T (2006) RNA FOLDING DURING TRANSCRIPTION. Annu Rev Biophys Biomol Struct 35:161–175. doi: 10.1146/annurev.biophys.35.040405.102053

Perdrizet GA, Artsimovitch I, Furman R, et al (2012) Transcriptional pausing coordinates folding of the aptamer domain and the expression platform of a riboswitch. Proc Natl Acad Sci 109:3323–3328. doi: 10.1073/pnas.1113086109

Potts AH, Vakulskas CA, Pannuri A, et al (2017) Global role of the bacterial post-transcriptional regulator CsrA revealed by integrated transcriptomics. Nat Commun 8:1596. doi: 10.1038/s41467-017-01613-1

Quereda JJ, Cossart P (2017) Regulating Bacterial Virulence with RNA. Annu Rev Microbiol 71:263–280. doi: 10.1146/annurev-micro-030117-020335

Ren G-X, Guo X-P, Sun Y-C (2017) Regulatory 3′ Untranslated Regions of Bacterial mRNAs. Front Microbiol 8:1276. doi: 10.3389/fmicb.2017.01276

Ruiz de los Mozos I, Vergara-Irigaray M, Segura V, et al (2013) Base pairing interaction between 5'- and 3'-UTRs controls *icaR* mRNA translation in *Staphylococcus aureus*. PLoS Genet 9:e1004001. doi: 10.1371/journal.pgen.1004001

Sittka A, Lucchini S, Papenfort K, et al (2008) Deep sequencing analysis of small

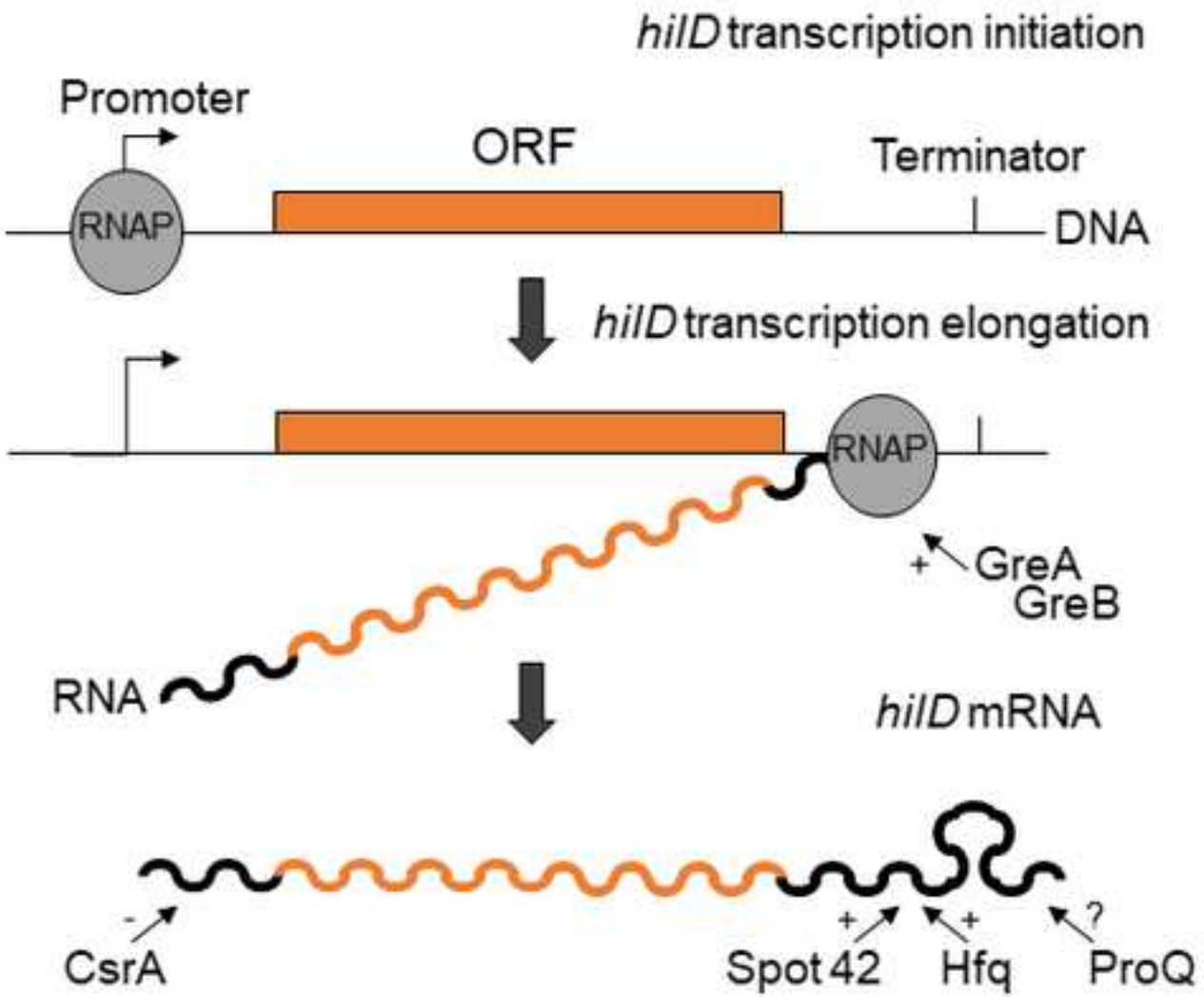
noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq. PLoS Genet 4:e1000163. doi: 10.1371/journal.pgen.1000163

Sittka A, Pfeiffer V, Tedin K, Vogel J (2007) The RNA chaperone Hfq is essential for the virulence of *Salmonella typhimurium*. Mol Microbiol 63:193–217. doi: 10.1111/j.1365-2958.2006.05489.x

Wong TN, Sosnick TR, Pan T (2007) Folding of noncoding RNAs during transcription facilitated by pausing-induced nonnative structures. Proc Natl Acad Sci 104:17995–18000. doi: 10.1073/pnas.0705038104

Figure

Click here to download Figure Figure 1.tif



Cover letter