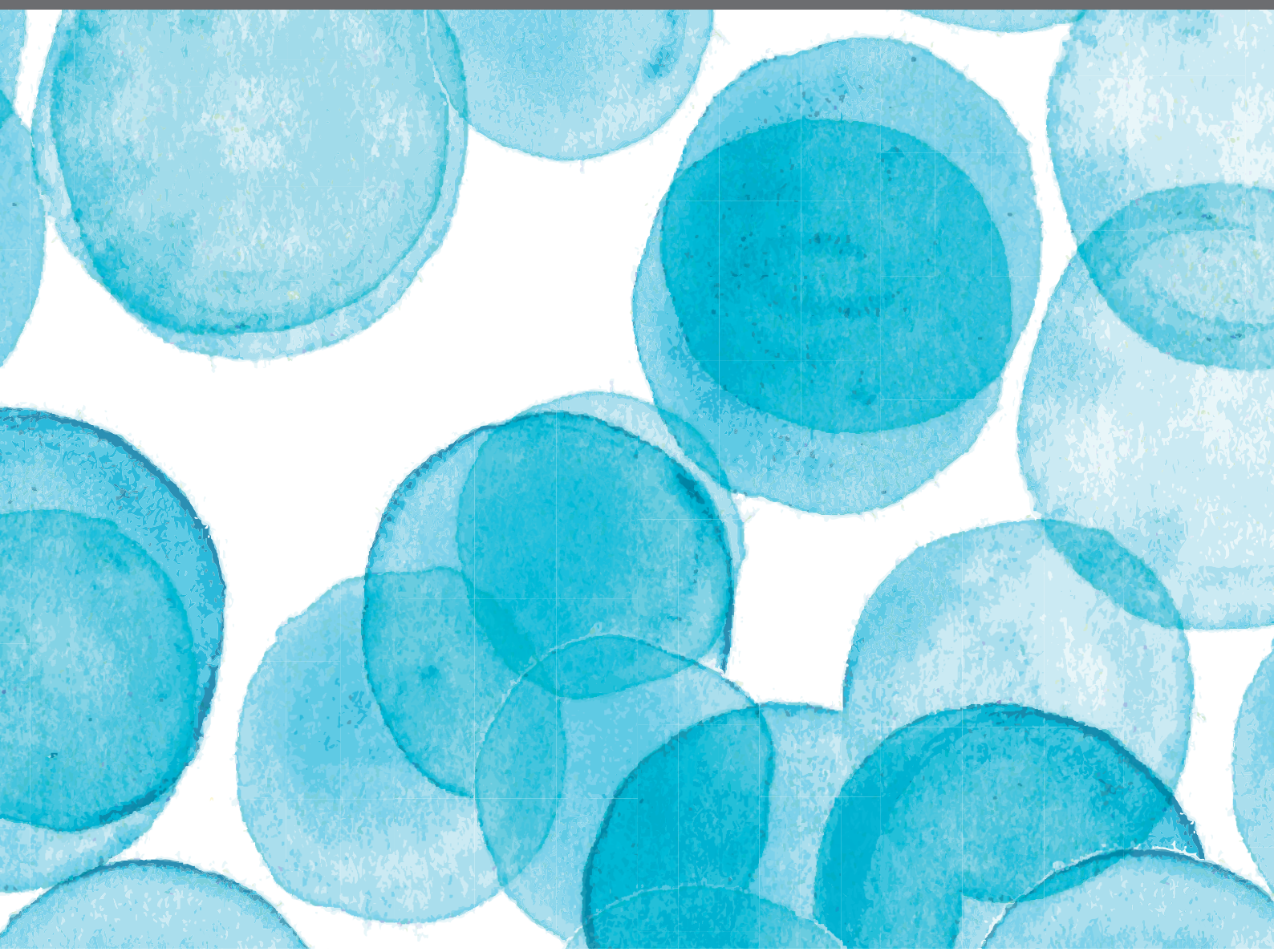
RNA BIOLOGY OF MICROORGANISMS

EDITED BY: Omar Orellana, Orna Amster-Choder, Jiqiang Ling and Rajat Banerjee

PUBLISHED IN: Frontiers in Microbiology



frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88971-918-1
DOI 10.3389/978-2-88971-918-1

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

RNA BIOLOGY OF MICROORGANISMS

Topic Editors:
Omar Orellana, University of Chile, Chile
Orna Amster-Choder, Hebrew University of Jerusalem, Israel
Jiqiang Ling, University of Maryland, College Park, United States
Rajat Banerjee, University of Calcutta, India

Citation: Orellana, O., Amster-Choder, O., Ling, J., Banerjee, R., eds. (2021). RNA Biology of Microorganisms. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88971-918-1

Table of Contents

frontiers in Microbiology

EDITORIAL
published: 29 October 2021
doi: 10.3389/fmicb.2021.754109

Editorial: RNA Biology of Microorganisms

Omar Orellana[1*†], *Orna Amster-Choder*[2†], *Rajat Banerjee*[3†] *and Jiqiang Ling*[4†]

[1] *Program of Cellular and Molecular Biology, Faculty of Medicine, Institute of Biomedical Sciences, University of Chile, Santiago, Chile,* [2] *Department of Microbiology and Molecular Genetics, Faculty of Medicine, The Institute for Medical Research, Israel-Canada (IMRIC), The Hebrew University, Jerusalem, Israel,* [3] *Department of Biotechnology and Dr. Bires Chandra Guha Centre for Genetic Engineering and Biotechnology, University of Calcutta, Kolkata, India,* [4] *Department of Cell Biology and Molecular Genetics, The University of Maryland, College Park, MD, United States*

Keywords: mRNA, sRNA, tRNA, canonical and non-canonical roles, degradation, interactions

Editorial on the Research Topic

RNA Biology of Microorganisms

The intention behind the Research Topic on RNA Biology of Microorganisms was to provide a broad view of the roles of RNA in the physiology of microorganisms, rather than focusing on specific functions. This purpose was mostly accomplished, since the published articles cover subjects from: (i) the life cycle of RNA transcripts, particularly of mRNAs, from birth to degradation; (ii) the various canonical and non-canonical roles of tRNAs, including translation and control of protein fate, as well as influencing genome variability; and (iii) regulatory non-coding RNAs: their identification by novel experimental strategies, the role of redundant RNA versions and the relationship between sRNAs of pathogenic bacteria with host RNAs. Putting all this information together, sheds light on the interplay of events that occur since the birth of an RNA, its distribution in the cell to accomplish its specific roles and its fate upon changes in environmental conditions. In the next paragraphs we briefly describe the contribution of each article to this global view of the biology of microorganisms.

THE BIRTH, FUNCTION, AND FATE OF MESSENGER RNAs

Several transcription factors regulate the binding of RNA polymerase to specific transcription promoters, among them NusG, a general transcription factor that is highly conserved in all organisms. Wang and Arstimovitch describe in their review article that NusG plays a number of roles in different cellular contexts, from the classical transcription antitermination of untranslated RNAs by counteracting Rho function on naked RNAs, through stimulation of transcription by avoiding backtracking of RNA polymerase, slowing down of transcription by halting RNAP at certain sequences and connecting RNAP to the ribosome during translation, to recruiting the ribosome to some mRNAs with weak ribosome binding sites. Paralogs of NusG are widely distributed in nature playing roles on specific genes using certain residues at conserved domains (NGN and KOW), while making similar contacts with RNA polymerase through conserved residues.

In a review article presented by Irastortza-Olaziregi and Amster-Choder, a detailed description of the strategies employed by bacterial cells to ensure association between transcription and translation in the apparently not physically separated milieu that encompasses the genetic material and cytoplasm. Physical coupling by direct or indirect interactions (bridged by proteins like NusG) between transcribing RNA polymerase and translating ribosomes are discussed. Lo and behold.

OPEN ACCESS

Edited by:
Patricia Coutinho Dos Santos, Wake Forest University, United States

Reviewed by:
Yuchen Liu, ExxonMobil, United States

****Correspondence:***
Omar Orellana
oorellan@med.uchile.cl

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *05 August 2021*
Accepted: *08 October 2021*
Published: *29 October 2021*

Citation:
Orellana O, Amster-Choder O, Banerjee R and Ling J (2021) Editorial: RNA Biology of Microorganisms. Front. Microbiol. 12:754109. doi: 10.3389/fmicb.2021.754109

Other strategies, such as the control of transcription rate by interaction of (p)ppGpp with RNA polymerase to coordinate transcription and translation, the localization of transcribing DNA to certain locations in the cell away from the bulk nucleoid (at the inner membrane or the surface of the nucleoid) where translation can take place, are also described. However, an important part of the article is devoted to the analysis of emerging data which promote the idea that uncoupled transcription-translation occurs in the context of the apparently non-separated nucleoid-cytoplasm organization of bacterial cells. Several different evidences reveal that this can happen by several tactics, e.g., ribosome free transcription of 5'UTRs, translation-independent localization of mRNAs to different locations in the cell, and faster transcription than translation in *B. subtilis*. A number of strategies, including association of RNA binding proteins (cold shock or ribosomal proteins among other) with the 5′ ends of mRNAs or sRNAs, as well as formation of secondary structures (riboswitches), might aid in the protection of free mRNAs in the cytoplasm against ribonucleases together with delaying translation until an mRNA reaches the proper location. Other proteins that are probably involved in uncoupled transcription-translation are RNA chaperones (Hfq and ProQ).

The review published by Vargas-Blanco and Shell gives a detailed description of the many different mechanisms that give rise to degradation of RNAs, particularly under stress conditions. It is accepted that steady state of RNAs is a balance between synthesis and degradation. Exo- and endo-ribonucleases are responsible for the degradation of RNAs, and their activity can be modulated in several ways to interfere with the RNA substrates, for example, sRNAs and RNA-binding proteins can block the access of RNases. Also, the chemical nature of the 5'end (chemical modifications, such as addition of NAD, Npn, m^6A or triphosphates) can influence the access of exonucleases. Moreover, secondary structures can also interfere with the access of nucleases. At the 3'end, the formation of poly (A) can facilitate the action of exonucleases. As a general rule, under stress RNAs are stabilized, but their cellular level decreases. This apparent contradiction might be explained by lowering transcription under these conditions.

Setlow and Christie reviewed the data obtained concerning the existence of RNAs in *B. subtilis* spores, as well as other spore-forming bacteria. As opposed to previous assumptions, evidence for the existence of RNAs in spores has been obtained for 1,800 different mRNAs, with no more that 50 mRNAs present in more than 1 copy/spore in the population, with most of them present in <10% of the population. The majority of these mRNAs and rRNAs present in spores are fragmented, and more fragmentation is observed with time post-sporulation. No evidence for the participation of these RNAs in translation after spore germination has been reported, except for *malS* mRNA, coding for malic enzyme, which might provide substrates for the synthesis of ATP. Hence, the data obtained thus far suggest that the main role of RNAs in spores is to provide nucleotides for the synthesis of new transcripts during germination.

TRANSFER RNAs IN TRANSLATION AND EVOLUTION

Three articles describe different ways by which transfer RNAs (tRNAs) play unorthodox roles in translation, as well as new roles in processes other than translation. In the article by Arias et al., the authors report that replacement of non-optimal glycine codons by optimal counterparts in the Cdc13 cyclin gene completely impairs the proliferation of cells. Overexpressing the tRNA decoding the non-optimal codons in wild type *cdc13* strains gave rise to severe defects in cell division, accompanied by an increase of Cdc13 aggregation, without affecting the levels of mRNA. The authors propose that the presence of rare glycine codons in *cdc13* plays a role in adjusting the translation efficiency of Cdc13 mRNA to a level competent with proper folding of the protein to guarantee its functionality.

A review article by Krahn et al. points to non-canonical structures formed by tRNAs. The classical and well-known secondary structure of tRNAs resembles a clover leaf, in which the acceptor stem, D loop arm, anticodon arm, variable arm and TφC arm are conserved between all organisms. However, some variations of this conserved structure have been shown to accomplish specific roles in the charging and translation of specific amino acids that are decoded by specialized termination codons. Examples, provided by selenocysteinyl, pyrrolysyl-tRNA, and others, show that these tRNAs are aminoacylated by canonical amino acids, which are later transformed into the derivatives that are incorporated to proteins. These tRNAs show extended acceptor, TφC and D stems. These modification are relevant for their recognition by the enzymes that modify the amino acids, as well as by the specialized translation elongation factors that allow recognition of the specific coding termination codons, thus facilitating the incorporation of these amino acids into proteins. Special emphasis is placed on the deviation of mitochondrial tRNAs from canonical structures, where different versions are found, including the lack of either the TφC or the D arms, or both, in tRNAs that are still functional in nematodes.

Other dimensions of tRNAs utility are highlighted in the review article by Guimarães et al. They depict the other roles, beside translation, of tRNA and their genes (tDNAs). The authors first describe the participation of transcription factor TFIIIC, one of the RNA polymerase III general transcription factors, which transcribes tDNAs (and 5S rRNA) in eukaryotes by recognizing internal transcription promoters. The transcription machinery helps in the recruitment of architectural chromatin proteins, such as condensin and adhesin, that help limiting the spreading of heterochromatin into the active euchromatin. A second case described by the authors is the effect of tRNAs on genome stability. The tDNAs are known as sites where R loops are formed, particularly when replication collides with transcription of tDNAs. These R loops are precursors of genome instability. Additionally, tDNAs in yeast, as well as in bacteria, are prone to insertion of transposable elements (retrotransposons). Besides the insertion of these genetic elements, they create a framework for duplications or deletions of segments of the genome within inserted genetic elements. Another interesting effect of tRNAs

on genome stability and evolution occurs in yeast when tRNA pools are altered, generating variability in the decoding capacity of certain tRNAs, imposing consequences on proteome stability. In the long run, these effects cause instability of chromosomes, particularly in regions encoding stress response genes as adaptive evolutionary events.

NON-CODING RNAS AS REGULATORY ELEMENTS

During the last couple of decades, a vast number of publications described the identification of non-coding RNAs as molecules that regulate gene expression. In this Research Topic, four articles, all reporting original research, expose different aspects of the regulatory role of non-coding RNAs. The articles by Li et al., and by El Mouali et al. describe studies of duplication of regulatory RNAs and their roles in different states of the bacterial life. The first example is that of the two copies of 6S RNAs (6S-1 and 6S-2) from *Bacillus thuringensis* that are co-transcribed from an operon, whose transcript requires processing. The deletion of the gene encoding 6S-1, and not 6S-2, affects bacterial growth in stationary phase, as well as spores production. The role of the 6S-2 is not known yet. The second example is that of the two copies of the RNAs, CsrB, and CsrC, that regulate the function of CsrA, the protein controlling translation of mRNAs of the carbon storage regulon. These regulatory RNAs, existing in a number of bacteria, bind to CsrA, titrating the protein and thus avoiding its interaction with GGA motifs in the target mRNAs. While CsrB is constitutively expressed in *Salmonella*, CsrC is regulated at the transcription level by the master carbon regulator CRP-cAMP. It is repressed at the logarithmic phase with the contribution of Spot42, an sRNA that is also controlled at the transcription level by CRP-cAMP.

Novel technologies for global analysis of RNAs have uncovered a vast variety of small non-coding RNAs (sRNAs) that play diverse functions in the regulation gene expression, mostly at the post-transcriptional level by binding to target mRNAs and altering their fate, both in bacteria and in eukaryotes. First identified as curiosities decades ago and later, due to advancements in bioinformatics and genome sequencing, many RNAs were predicted and identified in intergenic regions. Recent development of RNA-seq-accompanied methodologies have made possible the identification of many more sRNAs, in both intergenic and within coding regions of various organisms. Different approaches for isolating and identifying sRNAs in non-coding and coding regions of genomes have been developed. Their interaction with certain proteins has made it feasible to isolate and identify bound sRNAs and their regulated target mRNAs, permitting the prediction of their function in the cell. The article by Bar et al. describes a new computational approach, which uses machine-learning-based algorithm for the prediction of novel sRNAs, to define previously unknown sRNAs found by the recently developed RIL-sec technology for identifying sRNA that bind to the Hfq chaperone. Currently, around 2,800 sRNAs in *E. coli* have been found to interact with Hfq. The ∼1,000 identified in this work are encoded by different regions in the genome, either from intergenic regions or as part of the sense or antisense coding sequences in mRNAs transcripts.

Interest in the interaction between commensal bacteria and their host(s) has increased in recent years, mainly due to the analysis of the human microbiome. The article by Legüe et al. sheds light on the interaction between the model worm *C. elegans* and bacterial pathogens. The authors report that chronic exposure to the pathogen induces diapause formation (PIDF) in *C. elegans*. This defense mechanisms is transgenerational, that is, it is inherited by the worm and can be recalled upon exposure to pathogens a few generations forward. This mechanism requires the RNA interference machinery of the worm and sRNAs from the pathogenic bacteria. The authors used bioinformatic analyses of transcriptomic data from holobionts (bacteria and worm) to predict intergenerational and transgenerational interactions of sRNAs from both organisms. Their data points at pairing of RNAs from the two organisms, RNAs with similar sequences from the two organisms and eukaryotic sequence motifs found in bacterial RNAs. Surprisingly, chemical modification of tRNAs emerged as crucial events controlling the formation of the defensive form of the worm and, hence, relevant for the PIDF.

CONCLUDING REMARKS

The topics described in this Research Topic expose the many faces, roles and occurrence of RNA molecules. These studies show that the dynamic changes in mRNAs level toward a steady state concentration, which allow the cell to balance gene expression, is accomplished at the level of transcription, finely tuning the activity of RNA polymerases, but also by controlling the decay of RNAs. Also shown is the interplay between mRNAs and ribosome in bacteria, which is not limited to coupled transcription-translation, but is also relevant for uncoupled processes that allow adaptation to specific cellular requirements. Transfer RNAs, for a long time regarded as vehicles for transporting amino acids to the ribosome during translation, are emerging as active regulators of gene expression, making codon usage a tool for responding to specific physiological events, as well as molecules that play roles in evolutionary processes at the level of genome stability. The diversity of small non-coding RNAs continues to unveil, and their ability to play important regulatory roles is found not to be constrained to particular organisms, but to extend to the interaction of bacteria with their eukaryotic hosts. Together, these articles bring many exciting findings that sheds new light on bacterial RNA biology.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

This work was supported by Grants No. 1274/19 from the Israel Science Foundation (ISF) founded by the Israel Academy of Sciences and Humanities to OA-C, DST-SERB (order no. EMR/2016/002247), Govt of India to RB, National Institute of

General Medical Sciences R35GM136213 to JL, and Fondecyt, Chile 1190552 to OO.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Microbiology

REVIEW
published: 09 September 2020
doi: 10.3389/fmicb.2020.02111

Regulation of mRNA Stability During Bacterial Stress Responses

Diego A. Vargas-Blanco[1†] *and Scarlet S. Shell*[1,2*†]

[1] *Department of Biology and Biotechnology, Worcester Polytechnic Institute, Worcester, MA, United States,* [2] *Program in Bioinformatics and Computational Biology, Worcester Polytechnic Institute, Worcester, MA, United States*

Bacteria have a remarkable ability to sense environmental changes, swiftly regulating their transcriptional and posttranscriptional machinery as a response. Under conditions that cause growth to slow or stop, bacteria typically stabilize their transcriptomes in what has been shown to be a conserved stress response. In recent years, diverse studies have elucidated many of the mechanisms underlying mRNA degradation, yet an understanding of the regulation of mRNA degradation under stress conditions remains elusive. In this review we discuss the diverse mechanisms that have been shown to affect mRNA stability in bacteria. While many of these mechanisms are transcript-specific, they provide insight into possible mechanisms of global mRNA stabilization. To that end, we have compiled information on how mRNA fate is affected by RNA secondary structures; interaction with ribosomes, RNA binding proteins, and small RNAs; RNA base modifications; the chemical nature of 5′ ends; activity and concentration of RNases and other degradation proteins; mRNA and RNase localization; and the stringent response. We also provide an analysis of reported relationships between mRNA abundance and mRNA stability, and discuss the importance of stress-associated mRNA stabilization as a potential target for therapeutic development.

Keywords: ribonucleic acid, stress response, carbon starvation, nutrient starvation, hypoxia, mRNA degradation, mRNA stability, bacteria

OPEN ACCESS

Edited by:
Omar Orellana, University of Chile, Chile

Reviewed by:
Irina Artsimovitch, The Ohio State University, United States
Harald Putzer, UMR 8261 Expression Génétique Microbienne, France

***Correspondence:**
Scarlet S. Shell
sshell@wpi.edu

†ORCID:
Diego A. Vargas-Blanco
orcid.org/0000-0002-3559-2902
Scarlet S. Shell
orcid.org/0000-0003-1136-1728

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *28 May 2020*
Accepted: *11 August 2020*
Published: *09 September 2020*

Citation:
Vargas-Blanco DA and Shell SS (2020) Regulation of mRNA Stability During Bacterial Stress Responses. Front. Microbiol. 11:2111. doi: 10.3389/fmicb.2020.02111

INTRODUCTION

Bacterial adaptation to stress is orchestrated by complex responses to specific environmental stimuli, capable of rapidly regulating transcription, transcript degradation, and translation, which increases the organism's survival opportunities. Historically, regulation mechanisms for transcriptional and translational pathways have been the most studied, providing insight into the genes and protein products needed for bacterial adaptation to unfavorable growth environments. These findings have been key for our understanding of bacterial biology, allowing us, for example, to develop tools to tune bacterial machinery for biotechnology processes (such as Tao et al., 2011; Courbet et al., 2015; Daeffler et al., 2017; Martinez et al., 2017; Riglar et al., 2017), and to discover and develop new antibacterial drugs (for example, Yarmolinsky and Haba, 1959; Wolfe and Hahn, 1965; Maggi et al., 1966; Olson et al., 2011). However, the role of RNA degradation in stress responses is not well understood.

Modulation of mRNA degradation has been associated with various stress conditions in bacteria, such as temperature changes, growth rate, nutrient starvation, and oxygen limitation (see **Table 1**). Transcript stability – also referred as mRNA or transcript half-life – was shown to be globally altered in response to some stressors, while in other cases, gene-specific modulation of transcript stability contributes to specific expression changes that bacteria need to adapt to and survive in new environments (**Figure 1**).

In this review, we will discuss a range of reported situations in which bacterial mRNA stability is modulated in response to various stress conditions, with a focus on known and suspected mechanisms underlying such regulation. We will also discuss the ways in which known gene-specific mechanisms shape our thinking on the unanswered question of how mRNA pools are globally stabilized in response to energy stress. Furthermore, we will discuss the ways in which regulation of mRNA stability in clinically relevant bacteria, such as *Mycobacterium tuberculosis*, shape their responses to the host environment.

RNases AND OTHER DEGRADATION PROTEINS

The Degradosome

RNA degradation is carried out by a wide range of RNases, enzymes with strong activities and relatively low specificities toward their targets (reviewed in Carpousis, 2007). There are two main types of RNases: endonucleases and exonucleases. The former cleave RNA sequences at internal points, while the latter carry out nucleolytic attacks from either end of the RNA chain (deemed 5′ or 3′ exonucleases based on their enzymatic directionality). Some bacteria possess both 5′ and 3′ exonucleases – *M. tuberculosis* and *Mycobacterium smegmatis*, for example – while others such as *E. coli* have only 3′ exonucleases.

With respect to RNA degradation systems, *E. coli* is perhaps the most studied organism. In fact, it was in *E. coli* that a multiprotein complex, deemed the degradosome (**Figure 2**), was first reported (Carpousis et al., 1994; Py et al., 1994). In *E. coli*, the main degradosome components are two RNases (RNase E and PNPase), a DEAD-box RNA helicase (RhlB), and a glycolytic enzyme (enolase) (Carpousis et al., 1994; Py et al., 1994; Marcaida et al., 2006; Carpousis, 2007). RhlB facilitates RNase activity by unwinding stem-loops within RNA targets (Py et al., 1996). Both RNases carry out RNA degradation (Mohanty and Kushner, 2000; Deutscher, 2006; Unciuleac and Shuman, 2013). Moreover, in this bacterium the C-terminal region of RNase E acts as a scaffold for other degradosome components (Kido et al., 1996; Vanzo et al., 1998; Lopez et al., 1999; Morita et al., 2004). However, not all of the degradosome components are well defined or have known roles. For example, enolase is suspected to have a regulatory role in mRNA degradation under low phosphosugar levels (Morita et al., 2004; Chandran and Luisi, 2006) and anaerobic conditions (Murashko and Lin-Chao, 2017).

While RNases can degrade RNA substrates on their own, it has been suggested that degradosomes increase the efficiency of RNA degradation, for example by facilitating processing of structures such as stem-loops and repeated extragenic palindromic sequences (Newbury et al., 1987; McLaren et al., 1991; Py et al., 1996). Alteration of the degradosome components leads to changes in transcriptome stability; for example, deletion of RhlB in *E. coli* results in longer mRNA half-lives (Bernstein et al., 2004). Similarly, mRNA stability is dramatically increased when the arginine−rich RNA binding region or the scaffolding region of RNase E are deleted (Kido et al., 1996; Ow et al., 2000). While the RNA degradosome of *E. coli* has been extensively studied, the composition and function of degradosomes in other gram-negatives and in gram-positives may differ, and new studies are still uncovering this information. In the Firmicute *Bacillus subtilis,* there is no RNase E homolog. Instead, RNase Y serves as a degradosome scaffold for PNPase, the helicase CshA (Lehnik-Habrink et al., 2010), phosphofructokinase (Commichau et al., 2009), and RNase J1 and RNase J2 – two bifunctional enzymes with both endonucleolytic and 5′–3′ exoribonuclease activity (Even et al., 2005; Shahbabian et al., 2009; Mathy et al., 2010; Durand et al., 2012). Interestingly, the *B. subtilis* degradosome interactions have been shown mainly by bacterial 2-hybrid assays and immunoprecipitation of complexes stabilized by formaldehyde crosslinking (Commichau et al., 2009; Lehnik-Habrink et al., 2010), in contrast to the *E. coli* degradosome which can be immunoprecipitated without a crosslinking agent (Carpousis et al., 1994; Py et al., 1994, 1996). This suggests that *B. subtilis* degradosomes could be more transient in nature. A recent report on the Actinomycete *M. tuberculosis* provided insight into its elusive degradosome structure, which appears to be composed of RhlE (an RNA helicase), PNPase, RNase E, and RNase J (Plocinski et al., 2019). Overall, the degradosome is considered to be the ultimate effector of bulk mRNA degradation in bacterial cells, but it has also been implicated in regulating the stability of specific mRNAs and sRNAs, as will be discussed in later sections. For further details on the degradosome, we encourage reading the following reviews (Carpousis, 2007; Bandyra et al., 2013; Ait-Bara and Carpousis, 2015; Cho, 2017; Tejada-Arranz et al., 2020).

An Overview of RNase Regulation

There are multiple ways in which transcript levels can be regulated. Alteration of mRNA steady-state abundance is ultimately a consequence of changes in transcription, changes in mRNA half-life, or both. In the process of mRNA degradation, the roles of different RNases may be defined in part by their preferred cleavage sequences. In *Staphylococcus aureus*, RNase Y cleavage is usually in the R↓W sequence, near AU rich regions (Khemici et al., 2015). This pattern seems to be conserved in *B. subtilis* (Shahbabian et al., 2009). Furthermore, in these two gram-positive organisms, RNase Y cleavage appears to be influenced by proximity to a secondary structure. In *E. coli*, RNase E cleaves single-stranded RNA with a strong preference for the +2 sites in RN↓AU (Mackie, 1992; McDowall et al., 1994), or in RN↓WUU in *Salmonella enterica* (Chao et al., 2017). In *M. smegmatis*, a strong preference for cleavage 5′ of cytidines was detected in a transcriptome-wide RNA cleavage analysis (Martini et al., 2019). RNase E could be

TABLE 1 | Transcriptome-wide studies on mRNA half-life in bacteria.

Organism	Growth/stress condition	Response to stress/condition (transcriptome stability)	mRNA quantification method	Correlation between mRNA abundance and half-life	References
Bacillus cereus ATCC 10987, ATCC 14579	Exponential phase	–	RNA-seq	Positive	Kristoffersen et al., 2012
Bacillus subtilis	Early stationary phase	Stable*	Microarray	Not calculated	Hambraeus et al., 2003
Chlamydia trachomatis biovars: trachoma, lymphogranuloma venereum	Mid-phase stage of developmental cycle	–	RNA-seq	None	Ferreira et al., 2017
Escherichia coli	Exponential phase	–	Microarray	Negative	Bernstein et al., 2002
Escherichia coli	Exponential phase	–	Microarray	Not calculated	Selinger et al., 2003
Escherichia coli	0.1 h^{-1} growth rate	Stabilization at slower growth rates	Microarray	Negative	Esquerre et al., 2014, 2015
	0.2 h^{-1} growth rate				
	0.4 h^{-1} growth rate				
	0.63 h^{-1} growth rate				
Escherichia coli	Exponential phase	Stabilization in stationary phase	RNA-seq	Positive	Chen et al., 2015
	Stationary phase				
Escherichia coli	Exponential phase	Destabilization in Δ*csrA51*	Microarray	Negative	Esquerre et al., 2016
	Exponential phase (Δ*csrD*)				
	Exponential phase (Δ*csrA51*)				
Escherichia coli	Exponential phase	Stabilization in Ksm	RNA-seq	None for either condition†	Moffitt et al., 2016
	Exponential phase + Ksm (initiation inhibitor)				
Escherichia coli, Lactococcus lactis	Multiple‡	Stabilization at low growth rates and stress	Microarray, Nylon membrane-based macroarray	Negative	Nouaille et al., 2017
Escherichia coli	Exponential phase	Stabilization in *rne*Δ*MTS*	Microarray	Not calculated	Hadjeras et al., 2019
	Exponential phase (*rne*Δ*MTS*)				
Escherichia coli	Exponential phase	Stabilization in stress	Microarray	Negative†	Morin et al., 2020
	Glucose exhaustion				
	Acetate consumption				
	Carbon starvation				
Lactococcus lactis	Exponential phase	Stabilization at slower growth rates	Nylon membrane-based macroarray	Negative	Redon et al., 2005a,b
	Deceleration phase			None	
	Starvation phase			Positive	
Lactococcus lactis	Isoleucine limitation, 0.11 h^{-1} growth rate	Stabilization at slower growth rates	Nylon membrane-based macroarrays	Negative	Dressaire et al., 2013
	Isoleucine limitation, 0.51 h^{-1} growth rate				
	Isoleucine limitation, 0.8 h^{-1} growth rate				

(Continued)

TABLE 1 | Continued

Organism	Growth/stress condition	Response to stress/condition (transcriptome stability)	mRNA quantification method	Correlation between mRNA abundance and half-life	References
Stenotrophomonas maltophilia	Exponential phase	–	RNA-seq	None	Bernardini and Martinez, 2017
	Exponential phase (*rng*-defective mutant)	Stabilization in mutant			
Mycobacterium tuberculosis	Exponential phase	Stabilization in stress	Microarray	Negative	Rustad et al., 2013
	Hypoxic stress			Not calculated	
	Cold-induced stress			Not calculated	
Prochlorococcus MED4	0.325 day^{-1} growth rate	–	Microarray	Not calculated	Steglich et al., 2010
Staphylococcus aureus	Exponential phase	Stabilization in stress	Microarray	Not calculated	Anderson et al., 2006
	Cold-induced stress				
	Heat-induced stress				
	Mupirocin (isoleucyl-tRNA synthetase inhibitor, induces stringent response)				
	DNA damage (SOS response)	Destabilization in stress			

**Not compared to an exponential phase culture within the same study. Stabilization report based on previously reported studies. †Our analysis of the source data. ‡Includes data for L. lactis at growth rates of 0.09, 0.24, 0.35, and 0.47 h^{-1} (Dressaire et al., 2013); and unpublished and previously published data for E. coli at growth rates of 0.04, 0.11, 0.38, 0.51, and 0.80 h^{-1} and stationary phase (Esquerre et al., 2014).*

responsible for these cleavage events, given its major role in mycobacteria, however, we cannot yet exclude the possibility that they are produced by another endonuclease. In contrast, RNase III in *E. coli* has optimal activity on double-stranded RNA, where the cleavage site is specified by both positive and negative sequence and secondary structure determinants (Pertzev and Nicholson, 2006). While the preferred cleavage sites of various RNases seem highly represented in the mRNA pool, some transcripts are more resistant to cleavage than others, indicating the presence of mechanisms that regulate not only bulk RNA stability, but also differential stabilities among transcripts.

Studies of various mRNAs have identified multiple features that confer protection against RNase cleavage (**Figures 3**, **4A**). These include stem-loops (Emory et al., 1992; McDowall et al., 1995; Arnold et al., 1998; Hambraeus et al., 2002), 5′ UTRs and leader/leaderless status (Chen et al., 1991; Arnold et al., 1998; Unniraman et al., 2002; Nguyen et al., 2020), subcellular compartmentalization (Khemici et al., 2008; Montero Llopis et al., 2010; Murashko et al., 2012; Khemici et al., 2015; Moffitt et al., 2016), 5′ triphosphate groups (Bouvet and Belasco, 1992; Emory et al., 1992; Arnold et al., 1998; Mackie, 1998), 5′ NAD^+/NADH/dephospho-coenzyme A caps (Chen et al., 2009; Kowtoniuk et al., 2009; Bird et al., 2016; Frindert et al., 2018), Np_nN caps (Luciano et al., 2019; Hudecek et al., 2020), and association with regulatory proteins and sRNAs (Braun et al., 1998; Gualerzi et al., 2003; Moll et al., 2003; Afonyushkin et al., 2005; Daou-Chabo et al., 2009; Nielsen et al., 2010; Morita and Aiba, 2011; Faner and Feig, 2013; Liang and Deutscher, 2013; Deng et al., 2014; Sinha et al., 2018; Zhao et al., 2018; Cameron et al., 2019; Chen H. et al., 2019; Richards and Belasco, 2019). For example, in *Streptococcus pyogenes* the sRNA FasX binds to the 5′ end of *ska* – a transcript coding for streptokinase – increasing its mRNA half-life, thus allowing an extended period of time in which translation of streptokinase can occur (Ramirez-Pena et al., 2010). In other cases, the product of an mRNA can regulate its own transcript stability. In *E. coli*, the fate of the *lysC* transcript is regulated by a dual-acting riboswitch that, under low levels of lysine, promotes translation initiation while simultaneously sequestering RNase E cleavage sites. In the presence of lysine, the riboswitch folds into an alternative conformation that exposes RNase E cleavage motifs, in addition to blocking translation (Caron et al., 2012). In these examples, it is ultimately the conformational structure of the mRNA that allows regulation of its half-life, independently from the stability of the bulk mRNA pool.

The activity of RNases does not always result in RNA decay. Some mRNA precursors can be processed by RNases to create mature, functional forms of the transcript (Condon et al., 1996). In a similar manner, polycistronic transcripts can be cleaved by endonucleases to produce transcripts with varying degrees of stability; some examples include (Belasco et al., 1985; Baga et al., 1988; Nilsson and Uhlin, 1991; Nilsson et al., 1996; Ludwig et al., 2001; Esquerre et al., 2014; Xu et al., 2015). While this is a fascinating mechanism of gene-specific regulation, it is beyond the scope of this review.

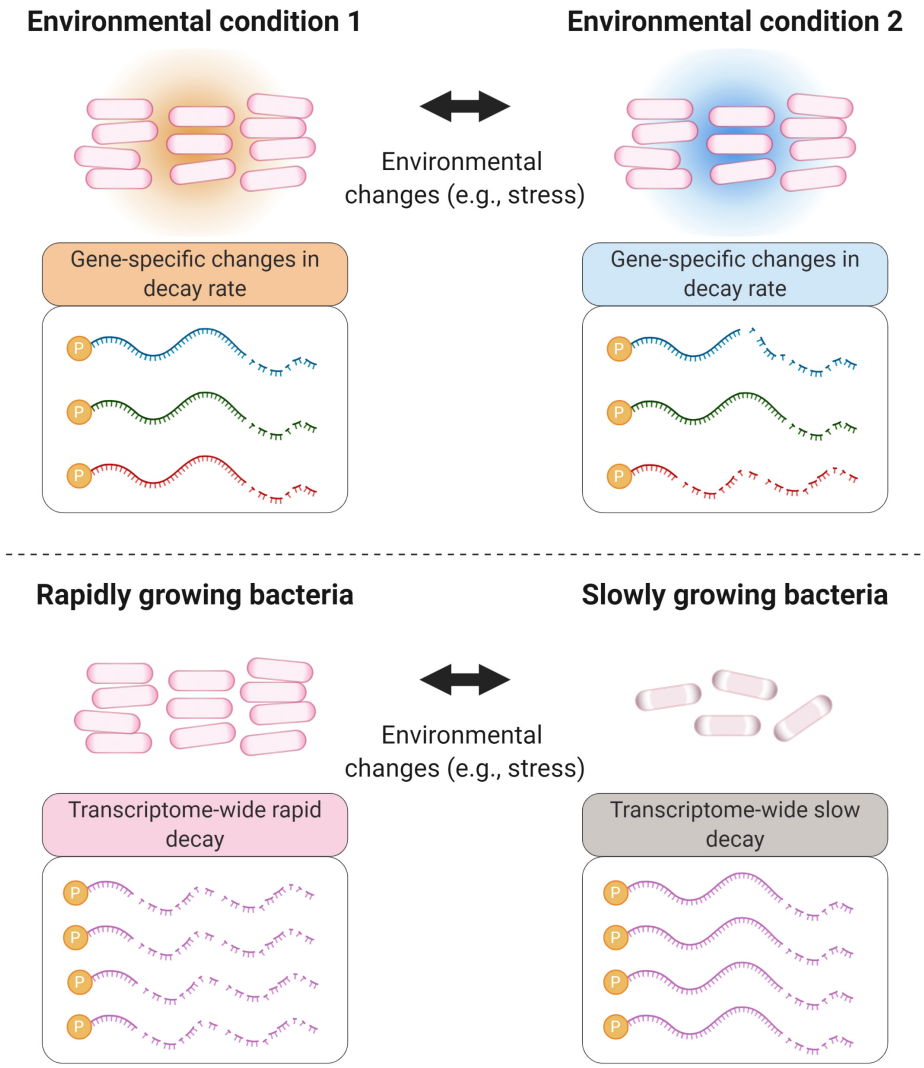


FIGURE 1 | Environmental changes cause mRNA degradation rates to change in both global and gene-specific ways. Bacterial adaptation to many stressors, and other changes in environment, involve modulation of degradation rates of specific transcripts encoding proteins relevant to the changing conditions (top panel). Some stressors, particularly those causing severe energy stress, trigger global stabilization of the mRNA pool (bottom panel). These scenarios are not mutually exclusive; stressors that cause global transcriptome stabilization typically also cause gene-specific changes in relative degradation rates.

mRNA STABILIZATION AS A RESPONSE TO STRESS

When bacteria are forced to slow or stop growth in response to stress, they must reduce their rates of protein synthesis. This can be done by direct modulation of translation or by regulation of transcription and transcript degradation rates. In recent decades, there have been many reports of mRNA stabilization as a response to different stressors, usually conditions that alter growth rate (see **Table 1**). In *E. coli*, the outer membrane protein A precursor transcript, *ompA*, is very stable in rapidly growing cells (Nilsson et al., 1984), but its half-life is significantly decreased in conditions of slow growth rate (Nilsson et al., 1984; Emory et al., 1992; Vytvytska et al., 2000). An inverse phenomenon was observed in stationary phase *E. coli* cells for *rpoS* and *rmf*, transcripts coding for the transcription factor σ38 and the ribosome modulation factor, respectively (Zgurskaya et al., 1997; Aiso et al., 2005). Research conducted in other organisms also showed regulation of degradation rates of specific mRNAs according to growth rate: *sdh*, coding for succinate dehydrogenase in *B. subtilis*, and *rpoS* in *Salmonella dublin* had mRNA half-lives negatively correlated with growth rate (Melin et al., 1989; Paesold and Krause, 1999). Furthermore, cell growth studies using chemostats revealed that most transcripts in *E. coli* stabilize at low growth rates (Esquerre et al., 2014), with those belonging to the COGs "Coenzyme transport and metabolism" and "Intracellular trafficking, secretion and vesicular transport" being enriched among the most highly stabilized transcripts. On the other hand, genes in "Cell motility" and "Secondary metabolites biosynthesis, transport and catabolism" had shorter

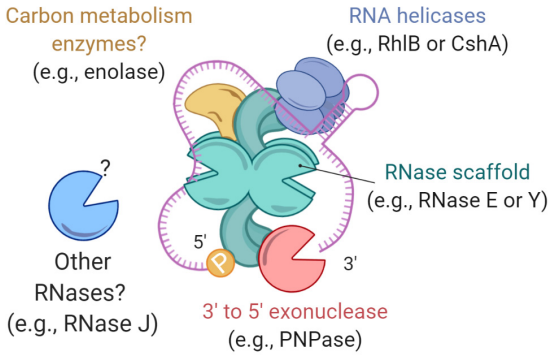


FIGURE 2 | Bacterial degradosomes. The bacterial degradosome is scaffolded by an RNase such as RNase E in *E. coli* and RNase Y in *B. subtilis*. The RNase scaffolds have catalytic domains and natively disordered scaffold domains that bind other degradosome proteins. Typical degradosome components in both gram-positive and gram-negative bacteria are RNA helicases, carbon metabolism enzymes, and other RNases.

half-lives than the transcript population mean (Esquerre et al., 2015). This reinforces the ideas that transcript half-lives may be linked to gene function and can be regulated as conditions require. For example, in *E. coli,* genes from the COGs "Carbohydrate transport and metabolism" and "Nucleotide transport and metabolism" are amongst the most stable at normal growth rates (Esquerre et al., 2014, 2015, 2016). Although these findings propose a link between growth rate and mRNA stability, it is possible that metabolic status rather than growth rate *per se* is the key determinant of global mRNA stability. In *M. smegmatis*, a drug-induced increase in metabolic activity resulted in accelerated mRNA decay and vice versa, even though growth was halted in both conditions (Vargas-Blanco et al., 2019). Another study supported these findings, showing that mRNA stabilization upon changes in nutrient availability could be dissociated from changes in growth rate (Morin et al., 2020).

Growth rate is altered as a consequence of metabolic changes as bacteria adapt to different environments. Because the ultimate goal of an organism is to survive and multiply, we can assume that in stress conditions – such as low-nutrient environments – bacteria trigger mechanisms that regulate energy usage and preserve energetically expensive macromolecules, such as mRNA. Thus, transcript stabilization is a logical response to various forms of energy stress. Indeed, *E. coli* stabilizes most of its transcriptome in anaerobic conditions (Georgellis et al., 1993) as well as in carbon starvation and stationary phase (Esquerre et al., 2014; Chen et al., 2015; Morin et al., 2020). Studies on *Rhizobium leguminosarum*, *Vibrio* sp. *S14,* and *Lactococcus lactis* also showed increased transcriptome half-lives when the bacteria are subjected to nutrient starvation (Albertson et al., 1990; Thorne and Williams, 1997; Redon et al., 2005a,b). *S. aureus* induces global mRNA stabilization in response to low and high temperatures, as well as during the stringent response (Anderson et al., 2006). Under hypoxic conditions, the median mRNA half-life in *M. tuberculosis* increases from ∼9.5 min to more than 30 min, and cells shifted from 37°C to room temperature stabilized their transcriptomes so dramatically that half-lives could not be measured (Rustad et al., 2013). Similarly, transcript stabilization occurs in *M. smegmatis* in response to carbon starvation and hypoxia (Smeulders et al., 1999; Vargas-Blanco et al., 2019). Intriguingly, transcript destabilization can be resumed within seconds upon re-oxygenation of hypoxic *M. smegmatis* cultures, suggesting a highly sensitive mechanism regulating mRNA degradation in response to stress and energy status (Vargas-Blanco et al., 2019).

This response seems to be conserved even in some eukaryotes such as *Saccharomyces cerevisiae*, where the mRNA turnover rate is slower under stress than in log phase (Jona et al., 2000), and in plants as part of their immune response (Yu et al., 2019). However, the adaptive mechanism(s) underlying global mRNA stabilization as a stress response remain unknown. In the following sections we will discuss in more detail diverse bacterial strategies that contribute to global and gene-specific regulation of RNA stability. Our intent is to highlight recent findings on regulation of RNA degradation, to serve as a base for development of experiments to uncover how mRNA stabilization occurs as a response to stress.

Regulation of RNA Degradation Proteins

In this section we will discuss factors that have been shown to regulate the abundance and activity of endo- and exonucleases. We invite the reader to consult some excellent reviews (Condon, 2003; Arraiano et al., 2010; Bechhofer and Deutscher, 2019) for additional information on the roles and activities of RNases.

As we described in a previous section, RNases have preferred cleavage sequences. These patterns can be either masked or exposed by alternative RNA folding configurations as a result of intracellular changes, allowing modulation of specific cleavage events, e.g., the *lysC* riboswitch which is sensitive to lysine concentration (Caron et al., 2012). However, this regulatory paradigm tends to be used to control specific messages rather than the overall transcriptome stability. Hence, a major open question is: Are there elements that control RNase abundance or RNase activity that regulate transcriptome stability globally?

Abundance of key RNases that catalyze rate-limiting steps in mRNA degradation can affect bulk mRNA decay. For example, depletion or mutation of RNase E caused bulk mRNA stabilization in *E. coli* (Lopez et al., 1999; Sousa et al., 2001); depletion or mutation of RNase Y caused bulk mRNA stabilization in *B. subtilis* and *S. pyogenes* (Shahbabian et al., 2009; Chen et al., 2013); depletion of RNase J caused bulk mRNA stabilization in *Helicobacter pylori* (Redko et al., 2016); and deletion of RNases J1 and J2 caused mRNA stabilization in *B. subtilis* (Even et al., 2005). Mechanisms for regulation of RNase abundance have been reported in some bacteria. In *E. coli*, RNase III autoregulates its abundance by cleaving its own operon to induce its degradation when RNase III protein levels are high (Bardwell et al., 1989; Matsunaga et al., 1996, 1997; Xu et al., 2008). Similarly, in *E. coli* a stem-loop located in the 5′ UTR of *rne* responds to changes in RNase E levels, allowing this enzyme to autoregulate its own production (Diwa et al., 2000; Diwa and Belasco, 2002). There is evidence that in some cases, stability of other mRNAs can be regulated by changes in RNase

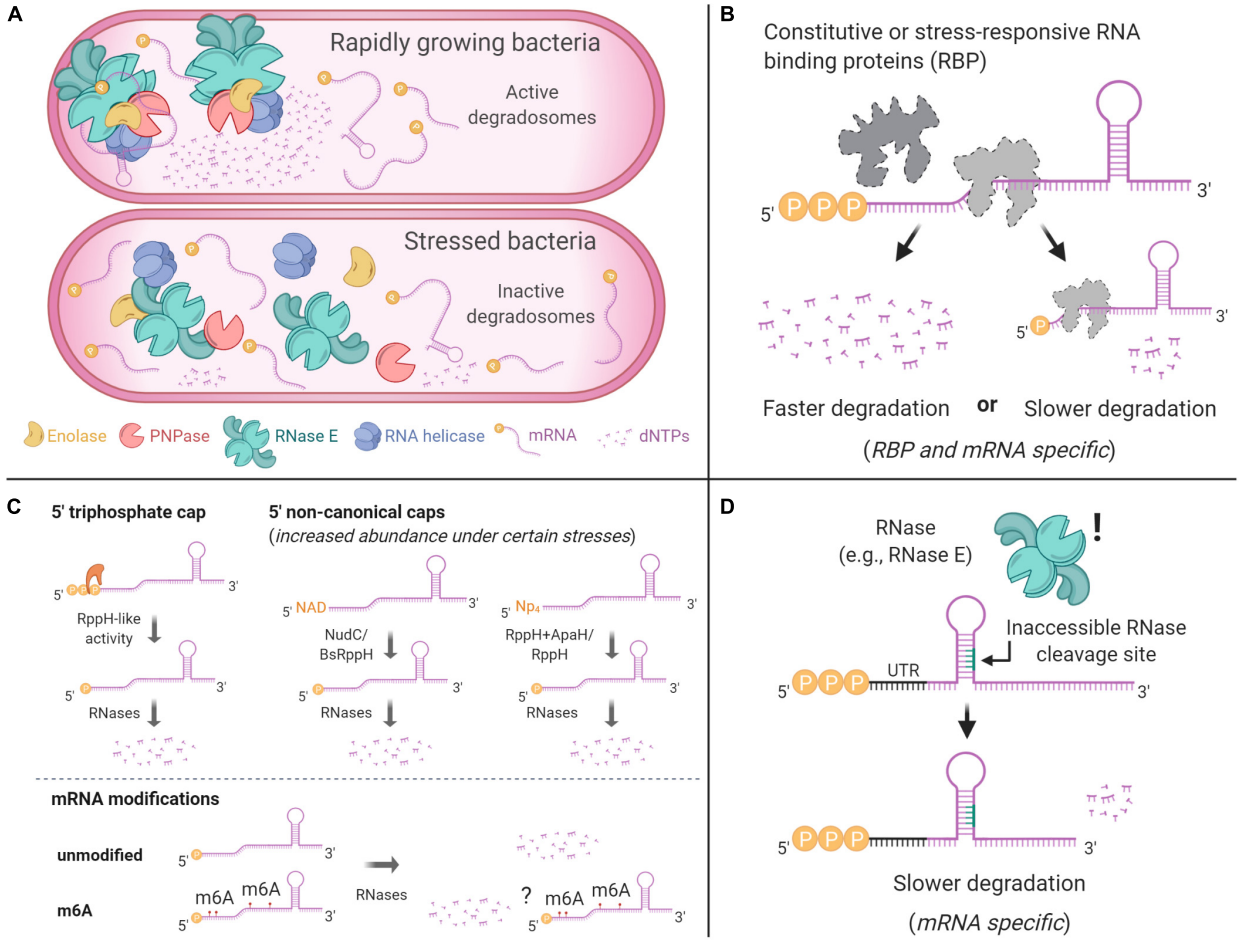


FIGURE 3 | Common mechanisms that can protect mRNAs from degradation. **(A)** Degradosome localization can influence its RNA degradation activity. In *E. coli*, the degradosome is anchored to the cytoplasmic membrane via RNase E's N-terminal domain, where it displays higher RNA processing activity in degradation foci. A cytoplasmic RNase E is less efficient in degradosome assembly and RNA processing. In *B. subtilis,* RNase Y is associated with the membrane and is more active when in smaller foci and less active when in larger foci. **(B)** RNA binding proteins can modulate mRNA degradation. Some of them, such as CsrA in γ-Proteobacteria, have regulatory roles as a response to environmental changes. **(C)** The chemical nature of mRNA 5′ ends can protect transcripts from degradation. These caps may vary depending on stress conditions. Nucleotide modifications in the bodies of transcripts have also been reported, but they have not been shown to alter mRNA stability. **(D)** RNA degradation depends on RNase accessibility to cleavage sites. Secondary structures that block cleavage sites can result in slower RNA degradation.

abundance. In *E. coli*, the *betT* and *proP* transcripts, encoding osmoregulators, showed increased abundance and stability when cells were subject to osmotic stress, apparently as a consequence of lower RNase III concentrations (Sim et al., 2014). However, there is not yet evidence that global stress-induced mRNA stabilization can be attributed to reduced RNase abundance. In *M. tuberculosis,* a quantitative proteomics study comparing exponentially growing and hypoxic cultures showed no alteration in levels of RNase E, RNase J, RNase III, PNPase, or the helicase HelY even after 20 days under hypoxia (Schubert et al., 2015). Only one RNA helicase, RhlE, had reduced levels in hypoxia (Schubert et al., 2015). Similarly, a study of *M. smegmatis* showed no variation in levels of RNase E, PNPase, or the predicted RNA helicase msmeg_1930 under hypoxia, re-aeration, or exponential growth (Vargas-Blanco et al., 2019). Because mycobacterial transcriptomes are rapidly stabilized upon encountering hypoxia and other stress conditions (Rustad et al., 2013; Vargas-Blanco et al., 2019), it is unlikely that alteration of RNase abundance is part of the early RNA stabilization responses in these organisms.

It is possible that the activity of existing RNA degradation enzymes is regulated. RNA helicases are ATP-dependent, and ATP levels decrease in some bacteria in severe energy stress (Rao et al., 2008; Vargas-Blanco et al., 2019). This raises the possibility that RNA degradation could be directly modulated by ATP levels. However, when this hypothesis was tested in *M. smegmatis,* mRNA stabilization was found to occur prior to a decrease in intracellular ATP levels upon exposure to hypoxic conditions (Vargas-Blanco et al., 2019). While these findings suggest that nucleotide sensing – particularly changes in ATP concentrations – does not influence the initial global stabilization response in mycobacteria, it is possible that ATP concentrations or ATP/ADP ratios could be responsible for further stabilization

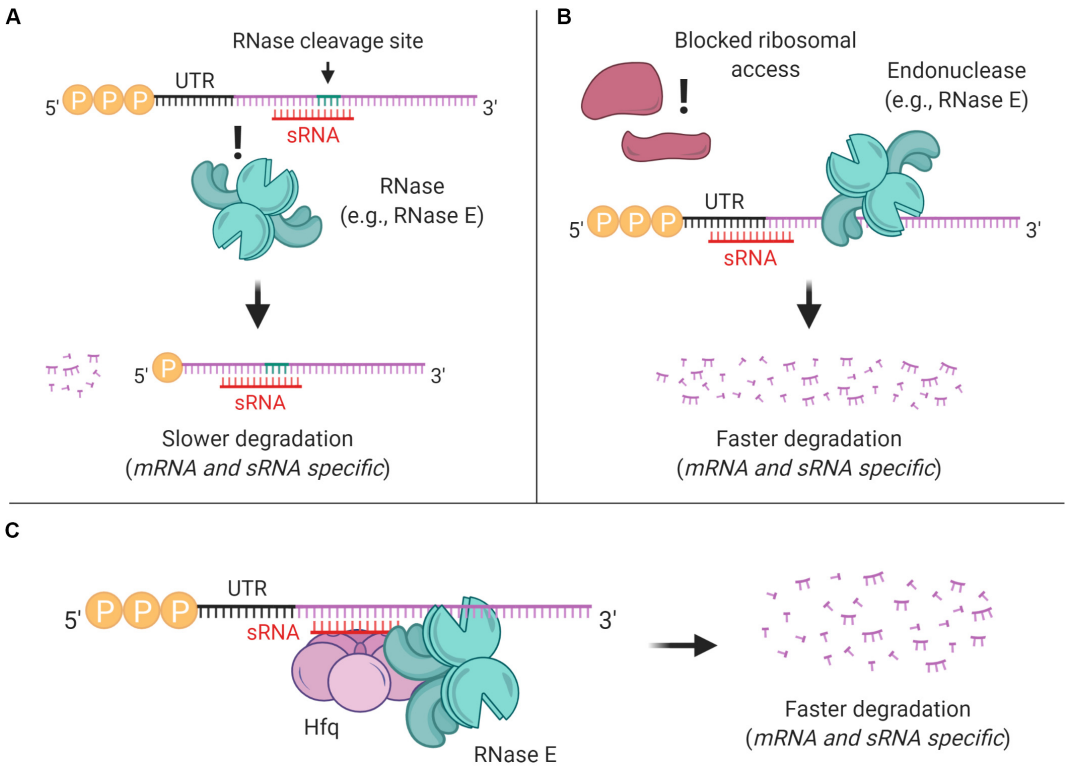


FIGURE 4 | sRNAs can affect mRNA stability through multiple mechanisms. **(A)** sRNA binding can mask preferred RNase cleavage sites, thereby stabilizing transcripts. **(B)** sRNA binding can block ribosome access to Shine-Dalgarno sites, reducing translation and typically destabilizing transcripts. **(C)** In *E. coli* and some other gram-negative bacteria, sRNA-mRNA pairing is often mediated by Hfq, which typically leads to mRNA degradation.

in later stages of dormancy, and/or that ATP levels contribute to global mRNA stabilization in other bacteria. The roles of nucleotides associated with the stringent response are discussed separately below.

In *E. coli,* inhibition of RNase E activity by RraA and RraB (Regulator of ribonuclease activity A and B) result in increased bulk mRNA half-life (Lee et al., 2003). However, in the case of RraA, the effect was observed after a significant overexpression of the inhibitor (Lee et al., 2003), something not observed under stress. Alternatively, inhibition of RNase activity by other factors may regulate transcript degradation. RNase E was recently shown to have a 5′ linear scanning function, and its cleavage activity is impaired upon encountering obstacles, such as sRNAs or ribosomes (Richards and Belasco, 2019). Furthermore, in *E. coli*, the activity of RNase E has been shown to depend on its anchorage to the inner membrane (**Figure 3A**). YFP-tagged RNase E forms small foci localized at the inner membrane (Strahl et al., 2015) which are dependent on metabolic activity; in anaerobic conditions RNase E rapidly dissociates from the membrane and diffuses in the cytoplasm, a response apparently dependent on enolase (Murashko and Lin-Chao, 2017). A cytoplasmic version of RNase E was unstable, and led to increased mRNA half-lives (Hadjeras et al., 2019). Interestingly, the cytoplasmic RNase E was able to assemble a degradosome and had a comparable *in vitro* activity to wild type RNase E, supporting the role of membrane attachment and cellular localization in RNase E activity (Moffitt et al., 2016; Hadjeras et al., 2019). Conversely, in *Caulobacter crescentus*, RNase E is cytoplasmic and forms bacterial ribonucleoprotein (BR) bodies, which dynamically assemble and disassemble in the presence of mRNA (Al-Husini et al., 2018). BR body formation was dependent on the RNase E scaffold domains and the presence of mRNA, while disassembly of the bodies required mRNA cleavage (Al-Husini et al., 2018). Intriguingly, the formation of BR-bodies increased under some stress conditions but was unaffected by others, suggesting they play an as-yet undefined role in stress response (Al-Husini et al., 2018). Further work is needed to understand the extent to which RNase localization contributes to regulation of mRNA degradation rates in various species.

In *B. subtilis*, the activity of RNase Y appears to be regulated by both subcellular localization and association with proteins termed the Y-complex (YaaT, YlbF, and YmcA). The Y-complex affects expression of genes involved in biofilm formation, sporulation, and competence, and in some cases, this was shown to be a direct consequence of altered mRNA degradation rates for the relevant genes (Tortosa et al., 2000; Carabetta et al., 2013; DeLoughery et al., 2016; Dubnau et al., 2016). The Y complex has been viewed as a specificity factor for RNase Y, required in particular for processing of polycistronic transcripts (DeLoughery et al., 2018). RNase Y also localizes in the cell membrane, where it can form RNase Y foci (Hunt et al., 2006; Lehnik-Habrink et al., 2011; Hamouche et al., 2020). These foci

seem to represent a less active form of the enzyme, as they increased in size in absence of RNA or in Y-complex mutants (Hamouche et al., 2020).

The Stringent Response and mRNA Degradation

The stringent response is perhaps one of the most well-studied mechanisms of prokaryotic stress adaptation. This response is modulated by guanosine-3′,5′-bisphosphate (ppGpp) and/or guanosine-3′-diphosphate-5′-triphosphate (pppGpp), alarmones collectively referred to as (p)ppGpp. In gram-negative bacteria, (p)ppGpp is synthesized by RelA in response to uncharged-tRNAs binding ribosomes, or by SpoT, a (p)ppGpp synthase/hydrolase, during fatty acid starvation (Seyfzadeh et al., 1993; Battesti and Bouveret, 2009). In some gram-positive bacteria, (p)ppGpp is synthesized by a dual RelA/SpoT homolog (Atkinson et al., 2011; Frederix and Downie, 2011; Corrigan et al., 2016). Once produced, (p)ppGpp halts the synthesis of stable RNA (tRNAs and ribosomes) while upregulating stress-associated genes and downregulating those associated with cell growth (Gentry et al., 1993; Chakraburtty and Bibb, 1997; Martinez-Costa et al., 1998; Avarbock et al., 2000; Artsimovitch et al., 2004; Corrigan et al., 2016). Intriguingly, (p)ppGpp was reported to inhibit PNPase in the actinomycetes *Nonomuraea* sp. and *Streptomyces coelicolor* but not in *E. coli* (Gatewood and Jones, 2010; Siculella et al., 2010), suggesting the stringent response may have a previously overlooked role in directly regulating mRNA degradation in some groups of bacteria. However, a recent study on the stringent response in *M. smegmatis* showed that (p)ppGpp was not required for mRNA stabilization in response to carbon starvation or hypoxia (Vargas-Blanco et al., 2019).

In the pathogen *Borrelia burgdorferi*, a connection between the stringent response and the expression of 241 sRNAs was recently stablished, 187 of which were upregulated during nutrient stress (Drecktrah et al., 2018). The authors of the aforementioned study described potential mechanisms of regulation by Rel_{Bbu} on transcription and fate of some transcripts, such as destabilization of the glycerol uptake facilitator transcript, *glpF*. The SR0546 sRNA is among the sRNAs induced by nutrient starvation; the upregulation of its target, *bosR*, encoding a transcriptional regulator, may suggest a regulatory role of (p)ppGpp on specific mRNA stabilization. However, the effects of these stringent response-induced sRNAs on mRNA stability have not yet been directly tested.

A surprising role of RelZ (initially called MS_RHII-RSD), a dual (p)ppGpp synthase and RNase HII, was reported for *M. smegmatis* (Murdeshwar and Chatterji, 2012). R-loops (RNA/DNA hybrids) are harmful structures that cause replication stress and can be removed by the RNase H domain of RelZ, while stalled ribosome removal is attributed to their alarmone synthase domain. RelZ was shown to be upregulated under short UV exposure in *M. smegmatis* (Krishnan et al., 2016), and while its role is suspected to increase cell viability under stress conditions (Petchiappan et al., 2020), the stringent response seems to not intervene in transcriptome stability regulation. This pathway leads to degradation of transcripts involved in R-loops, but given the low frequency of R-loop formation, the effects on mRNA pools are likely to be minimal.

Overall, there is much evidence that the stringent response regulates expression of specific transcripts in various bacteria. However, the extent to which control of mRNA stability contributes to these effects is mostly untested. The stringent response also plays important roles in mediating global responses to starvation and other forms of energy stress, but there is not yet evidence that it contributes to global mRNA stabilization, which is a consistent component of these stress responses. This suggests that the stringent response may not be the mediator of global mRNA stabilization in response to stress, or that its involvement in this process is species-specific.

Transcript Modifications as Regulators of mRNA Decay

Bacterial mRNA is primarily transcribed using nucleoside triphosphates as initiating nucleotides, making mRNAs triphosphorylated at their 5′ ends. In *S. aureus,* RNase J1 exhibits strong *in vitro* exo- and endonucleolytic activities on 5′ triphosphorylated transcripts (Hausmann et al., 2017). However, in most other organisms studied to date, RNases E, J, and Y more efficiently cleave mRNAs with 5′ monophosphates (**Figure 3C**). RNase E is an endoribonuclease, but has a binding pocket for monophosphorylated 5′ ends (Callaghan et al., 2005) that strongly stimulates its activity in organisms including *E. coli* and *M. tuberculosis* (Mackie, 1998; Zeller et al., 2007). Similarly, in *B. subtilis*, RNase J1, and to a lesser extent J2, show a strong preference toward 5′ monophosphorylated substrates (Even et al., 2005). RNase Y also shows preference toward monophosphorylated 5′ substrates, but to a lesser extent (Shahbabian et al., 2009). These findings contributed to the discovery of RppH, an RNA pyrophosphohydrolase. Similar enzymes were later found in other bacteria, such as *Bdellovibrio bacteriovorus* (Messing et al., 2009) and *B. subtilis* (Richards et al., 2011). However, while the role of 5′ triphosphate pyrophosphohydrolysis was initially attributed to RppH (Celesnik et al., 2007; Deana et al., 2008), recent findings have shown that the primary substrate of RppH in *E. coli* is 5′ diphosphorylated RNAs, and that 5′ diphosphorylated RNAs are abundant in the transcriptome (Luciano et al., 2017). As RppH cannot convert 5′ triphosphates to diphosphates, this suggests the existence of an unknown 5′ triphosphate to diphosphate phosphorylase. Given that 5′ monophosphates make transcripts more susceptible to degradation in multiple organisms, one could envision regulation of 5′ triphosphate pyrophosphohydrolysis as a potential mechanism for regulation of mRNA stability. However, to our knowledge there are not yet reports of if and how pyrophosphohydrolysis or γ-phosphate removal are regulated.

The presence of non-canonical mRNA 5′ ends has recently been reported for subsets of mRNAs in several bacterial species, suggesting another possible mechanism for regulation of mRNA stability (**Figure 3C**). Examples include NADH and NAD+ (Chen et al., 2009; Cahova et al., 2015), and less commonly, dephospho-CoA, succinyl-CoA, acetyl-CoA, and

methylmalonyl-CoA (Kowtoniuk et al., 2009). We will refer to these as 5′ caps, with the understanding that they are structurally and functionally distinct from eukaryotic mRNA caps. Other studies have shown additional types of 5′ capping, as well as potential mechanisms behind it (Bird et al., 2016; Zhang et al., 2016; Julius and Yuzenkova, 2017). In most cases, bacterial caps are incorporated directly into mRNAs during transcription initiation. RNA polymerase can initiate transcription with non-canonical nucleotides such as NAD in *E. coli* (Bird et al., 2016; Vvedenskaya et al., 2018) and *B. subtilis* (Frindert et al., 2018). Furthermore, *E. coli* RNA polymerase seems to initiate with dinucleoside tetraphosphates (Np_4N), Np_4A in particular, with an efficiency almost 60 times higher than for NAD (Luciano and Belasco, 2020). Alternative, posttranscriptional mechanisms may also contribute to Np_4 capping formation, as *in vitro* experiments using LysU (lysyl-tRNA synthetase) from *E. coli* suggest (Luciano et al., 2019).

The intracellular concentration of Np_4As were shown to be affected by overproduction of aminoacyl-tRNA synthetases (Brevet et al., 1989). Interestingly, some stress conditions also induce higher levels of Np_4Ns, for example heat shock (Lee et al., 1983), oxidative stress (Bochner et al., 1984), cadmium stress (Coste et al., 1987; Luciano et al., 2019) and disulfide stress (Bochner et al., 1984; Luciano et al., 2019). 5′ mRNA decapping was shown to require Nudix enzymes, such as NudC and BsRppH, to hydrolyze NAD-RNA substrates (Hofer et al., 2016; Frindert et al., 2018). On the other hand, hydrolysis of Np_4As requires RppH and ApaH, the latter carrying out the hydrolysis of Np_4As into two NDPs (Farr et al., 1989); in this context ApaH generates a diphosphorylated 5′ end that can be readily converted to monophosphate 5′ end by RppH (**Figure 3C**). Non-canonical mRNA 5′ ends also occur when transcription initiates with short RNA degradation products, resulting in mRNAs with 5′ hydroxyls (Druzhinin et al., 2015). Such transcripts have been found in *E. coli* and *Vibrio cholerae* and are present at increased abundance in stationary phase (Vvedenskaya et al., 2012; Druzhinin et al., 2015). However, the effects of these alternate 5′ ends on transcript stability have not been reported.

Some mRNA caps have been shown to stabilize mRNAs in *E. coli* (Bird et al., 2016; Luciano et al., 2019) and in *B. subtilis* (Frindert et al., 2018). For example, after increasing the cellular concentration of Np_4Ns in cadmium-stressed cells and in Δ*apaH* mutants, RNA stability was increased, suggesting that Np_4 caps have a stabilizing role (Luciano et al., 2019). Additionally, in this study Np_4 caps were suggested to be more abundant than NAD caps. Similarly, in the *E. coli* Δ*nudC* mutant strain there is an increase of up to fourfold in RNA stability for transcripts with non-canonical 5′ caps (Bird et al., 2016). Furthermore, NAD 5′ caps were almost twofold more abundant for cells in stationary phase when compared to exponential phase (Bird et al., 2016). Together, these findings present a potential mechanism for stabilization of mRNA under stress conditions. An interesting regulatory mechanism behind Np_4 decapping in *E. coli* was recently linked to methylation in m^7Gp_4Gm and m^6Ap_3A 5′ caps, which protects them from RppH cleavage but not from AppH (Hudecek et al., 2020). Methylated Np_nN caps were shown to be more abundant in stationary phase than exponential phase (Hudecek et al., 2020), consistent with the idea that these caps protect mRNA from degradation. Interestingly, the Np_nN caps found in that study did not include Ap_4N (Hudecek et al., 2020), presumably due to different stress conditions and detection techniques than those in Luciano et al. (2019). Since capped mRNAs appear to be generally more stable than canonical mRNAs, it is logical to infer that when stress conditions cause growth to slow or stop and transcription to slow or stop concomitantly, the proportion of capped mRNAs will increase as a result of their inherently longer half-lives. One could therefore speculate that the global mRNA stabilization observed in non-growing bacteria is due in part to an mRNA pool that is largely protected by 5′ caps. This is plausible assuming capping frequency remains constant or increases under stress. But, a recent study argues against this idea. Rapid transcript destabilization occurred in hypoxic *M. smegmatis* cultures after re-exposure to oxygen, even when transcription was blocked prior to re-aeration (Vargas-Blanco et al., 2019). Thus, mRNA capping does not explain the transcript stabilization observed in these conditions (early-stage hypoxia) – at least in *M. smegmatis* – but could be involved in mRNA stabilization in other conditions and/or other bacteria.

Another possible mechanism of mRNA stabilization involves posttranscriptional nucleotide modifications (**Figure 3C**). N^6-methyladenosine (m^6A) is a common base modification in mice and humans (Meyer et al., 2012; Linder et al., 2015). This methylation is enriched near stop codons and in 3′ UTRs (Yue et al., 2018), and is dependent on the consensus motif DRACH (Linder et al., 2015). Recent studies revealed m^6A to be an important part of a transcript stability regulatory mechanism, as it facilitates mRNA degradation in association with RBP in mice, zebra fish, and human cells (Schwartz et al., 2014; Wang et al., 2014; Zhao et al., 2017). Moreover, the levels of m^6A methylation are responsive to stress conditions, as shown for human cancer cells under hypoxic conditions (Panneerdoss et al., 2018), suggesting a posttranscriptional regulatory role. In *E. coli* and *Pseudomonas aeruginosa*, m^6A is present at similar levels, ∼0.2–0.3% of adenines (Deng W. et al., 2015), to those reported for yeast and other eukaryotes (Wei et al., 1975; Bodi et al., 2010). However, in contrast to mammals, m^6A appears distributed throughout the gene, with modest enrichments near the 5′ ends and centers of transcripts, and with a similar m^6A motif for *E. coli* and *P. aeruginosa* (UGCCAG and GGYCAG, respectively) (Deng X. et al., 2015). Contrary to eukaryotes, m^6A methylation has not been shown to have a global role in mRNA degradation in bacterial stress responses. A deep analysis in *E. coli* and *P. aeruginosa* revealed no difference in the m^6A levels for cells growing in LB when compared to other (unspecified) growth media, or oxidative stress; interestingly, increasing the temperature from 37 to 45°C lowered m^6A methylation levels, but only for *P. aeruginosa* (Deng X. et al., 2015). Furthermore, the m^6A levels were lower in other bacteria (∼0.02–0.08%, for *S. aureus*, *B. subtilis*, *Anabaena* sp., and *Synechocystis* sp.) (Deng X. et al., 2015), suggesting that this particular base modification may not be conserved across bacteria. In *E. coli*, codon modifications of the *ermCL* mRNA with m^6A blocked translation, though it had no impact on mRNA degradation rates

(Hoernes et al., 2016). While it is conceivable that m^6A has a role in the regulation of bacterial translation, current evidence does not suggest it regulates mRNA fate.

5-methylcytosine (m^5C) has also been found in mRNA. In eukaryotes, m^5C has been shown to increase transcript stability (Arango et al., 2018; Chen X. et al., 2019; Yang et al., 2019; Schumann et al., 2020), while reports on translation regulation are controversial (Huang et al., 2019; Yang et al., 2019; Schumann et al., 2020). m^5C modifications have been found in mRNA and 23S rRNA in the archaeon *Solfolobus solfataricus* (Edelheit et al., 2013). However, there is no defined role of m^5C in *S. solfataricus*, and evidence of m^5C in bacteria or regulatory roles in RNA degradation have not been reported.

Another modification, and perhaps the most abundant in RNA, is pseudouridine (Ψ) (Rozenski et al., 1999). Ψ is present at the position U55 in all *E. coli* tRNAs (Gutgsell et al., 2000), and is widespread across kingdoms (Nishikura and De Robertis, 1981; Becker et al., 1997; Ishida et al., 2011). In *E. coli,* deletion of *truB*, encoding a tRNA Ψ 55 synthase (Nurse et al., 1995), was shown reduce viability after a temperature shock (37–50°C); however, no viability changes were observed during exponential growth at 37°C (Kinghorn et al., 2002). In *Thermus thermophilus*, a Δ*truB* mutant showed a growth defect when cultured at 50°C (Ishida et al., 2011). Thus, it is possible that the presence of tRNA modifications under stress conditions contributes to survival in other bacteria. Other tRNA modifications have been also reported in bacteria and yeast during stress, contributing to a translational bias with implications for translation regulation (Chan et al., 2010, 2012; Laxman et al., 2013; Deng W. et al., 2015; Chionh et al., 2016). However, while stress may alter tRNA modifications, ultimately these changes lead to translational regulation without clear evidence, at least in bacteria, of effects on mRNAs. On the other hand, Ψ modifications on mRNA have been shown to increase mRNA stability in yeast and human cells (Carlile et al., 2014) and in *Toxoplasma gondii* (Nakamoto et al., 2017). A broad study involving *E. coli* and human cells found that even a single replacement of U with Ψ in mRNA can interfere with translation (Eyler et al., 2019). Whether these modifications ultimately regulate mRNA stability in bacteria as a response to stress is an open question. Based on evidence aforementioned for *M. smegmatis* regarding the rapidity of transcript destabilization after stress alleviation (Vargas-Blanco et al., 2019), we speculate that base modifications are unlikely to be the primary mechanism of mRNA stabilization in hypoxic mycobacteria, although it could play roles in other organisms or conditions.

Roles of Ribosomes, Translation, sRNAs, and RNA-Binding Proteins in Regulation of mRNA Decay

Experiments conducted by Bechhofer and others in *B. subtilis* showed that ribosome stalling can increase *ermC* half-life. In this scenario, ribosomes acted as obstacles at the 5′ ends of transcripts, resulting in protection from endonucleolytic cleavage downstream (Shivakumar et al., 1980; Bechhofer and Dubnau, 1987; Bechhofer and Zen, 1989). These findings would become early evidence of a 5′–3′ polarity for endonucleolytic activity, dependent upon or enhanced by (1) interaction with a 5′ monophosphate, and (2) RNase linear scanning mechanisms, as it would be later reported by others (Bouvet and Belasco, 1992; Jourdan and McDowall, 2008; Kime et al., 2010; Richards and Belasco, 2016, 2019). In *E. coli*, the use of puromycin or kasugamycin – translation inhibitors that cause ribosomes to dissociate from transcripts – caused faster mRNA decay in the absence of new transcription (Varmus et al., 1971; Pato et al., 1973; Schneider et al., 1978). On the other hand, the use of chloramphenicol, fusidic acid or tetracycline – elongation inhibitors that cause ribosomes to stall on transcripts – resulted in transcript stabilization (Varmus et al., 1971; Fry et al., 1972; Pato et al., 1973; Schneider et al., 1978), findings also later shown in *M. smegmatis* (Vargas-Blanco et al., 2019). These results are consistent with ribosome binding having a protective effect on mRNAs (**Figure 5**). In experiments where transcription was not blocked, it is possible that the mRNA stabilization seen in response to elongation inhibitors may also be conferred in part by the sudden increase in rRNA synthesis that these drugs cause, which increases the abundance of potential RNase substrates and could therefore titrate the activity of RNases such as PNPase and RNase E (Lopez et al., 1998). However, the increase in rRNA synthesis cannot fully explain these effects.

In *B. subtilis,* the stability of *gsiB,* encoding general stress protein, and *ermC,* encoding erythromycin resistance leader peptide, are associated with ribosome binding (Sandler and Weisblum, 1989; Hambraeus et al., 2000). Mutations to the RBS sites of *gsiB*, *aprE* (coding for subtilisin), and SP82 phage mRNA resulted in reductions of their mRNA half-lives (Hue et al., 1995; Jurgen et al., 1998; Hambraeus et al., 2002). Transcript stability conferred by ribosomes does not always require productive translation, at least for *ermC* (Hambraeus et al., 2002) and *ompA* (Emory and Belasco, 1990), where transcripts were stable in the absence of start codons as long as strong Shine-Dalgarno (SD) sequences were present (Arnold et al., 1998). A later study also in *E. coli* reported that ribosome protection is independent of translation for another transcript (Wagner et al., 1994). Transcript stabilization in a translation-independent manner was also shown for *B. subtilis*, with the insertion of an alternative SD (not involved in translation) to the gene reporter *cryIII* (Agaisse and Lereclus, 1996). These findings suggest that binding of a 30S subunit to a transcript, regardless of translation, may suffice to impair RNase degradation.

However, other studies did find a correlation between translation itself and stability. In *E. coli*, codon composition can influence translation rate and mRNA stability; codon-optimized transcripts were more stable than their corresponding non-modified, inefficiently-translated versions (Boel et al., 2016). Similar results were shown for *S. cerevisiae* (Presnyak et al., 2015). A transcriptome-wide analysis in *E. coli* also identified a positive correlation between mRNA stability and codon content optimality, for bacteria growing at different rates (Esquerre et al., 2015). This directly contradicted a previous report that codon optimality and half-life were inversely correlated (Lenz et al., 2011), possibly due to use of different codon optimality metrics. In *B. subtilis*, translation initiation is necessary to prevent

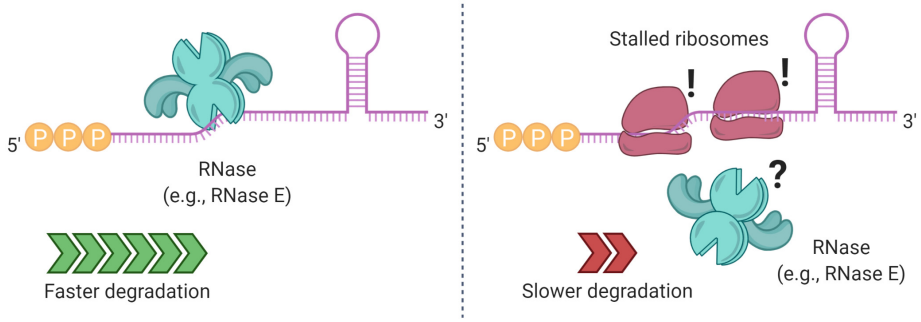


FIGURE 5 | Ribosome binding and stalling can alter mRNA degradation. In some cases, ribosome stalling can mask RNase cleavage sites, increasing the half-life of a transcript. Elements that prevent ribosome binding, such as translation initiation inhibitors, lead to shorter mRNA half-lives.

swift degradation of the *hbs* transcript, which encodes the DNA binding protein HBsu (Daou-Chabo et al., 2009; Braun et al., 2017). In *M. smegmatis* and *M. tuberculosis*, RNase E cleaves the *furA-katG* operon, producing an unstable *furA* message that is rapidly degraded while the *katG* transcript is stabilized as it becomes readily accessible for translation (Sala et al., 2008). Overall, regulation of mRNA stability by translation initiation and SD strength seems to be gene-specific.

While it is generally accepted in *E. coli* that occlusion of RNase cleavage sites by ribosome occupancy may protect a transcript from degradation (Joyce and Dreyfus, 1998), ribosome association with mRNA has not been shown to regulate mRNA stability globally in response to stress. However, data from *B. subtilis* suggest an interesting mechanism by which RNase activity could affect translation and therefore mRNA degradation on a transcriptome-wide scale (Bruscella et al., 2011). The *infC-rpmI-rplT* operon, which encodes translation initiation factor 3 (IF-3) along with two ribosomal proteins, is expressed from two promoters. The resulting transcripts have different sensitivities to RNase Y, and the RNase Y-sensitive transcript is not competent for translation of IF-3. As a result, inhibition of RNase Y expression alters the relative abundance of the two transcript and causes reduced translation of IF-3. If this were to cause globally reduced translation due to IF-3 deficiency, mRNA decay could be globally increased as a result, although this effect would presumably be counteracted by the globally reduced RNase Y activity. Complex interplays between RNase levels and translation may therefore have the potential to globally impact mRNA decay in *B. subtilis*.

RNA-binding proteins (RBPs), stalled ribosomes, and SD-like sequences in close proximity to transcript 5′ ends can also alter mRNA fate (Sharp and Bechhofer, 2005). In *B. subtilis,* interaction of the RBP Glp with the 5′ UTR of *glpD*, encoding glycerol-3-phosphate dehydrogenase, increases the transcript's stability (Glatz et al., 1996). Other RBPs can modulate the stability of target genes during stress conditions (**Figure 3B**). For example, H-NS, a histone-like protein, regulates the RNA stability of *rpoS* in *E. coli* and *V. cholerae* in stressful environments (Brescia et al., 2004; Silva et al., 2008; Wang et al., 2012). The carbon storage regulator CsrA is an RBP that regulates gene expression posttranscriptionally in *E. coli* and other γ-Proteobacteria in response to environmental changes, described in Timmermans and Van Melderen (2010) and Romeo and Babitzke (2018). CsrA regulatory roles are best studied in *E. coli*. The *glgCAP* transcript, encoding genes implicated in the biosynthesis of glycogen, is destabilized when bound by CsrA (Liu et al., 1995). This response is halted when *E. coli* enters stationary phase, where CsrA is sequestered by the sRNA CsrB in a ribonucleoprotein complex (Liu et al., 1997). Conversely, CsrA was shown to stabilize some transcripts. CsrA directly binds the *pgaA* transcript, increasing its half-life along with the rest of the *pgaABC* polycistron, encoding genes associated to biofilm formation (Wang et al., 2005). Similarly, CsrA stabilizes the *flhDC* transcript, encoding the flagellar activation genes $FlhD_2C_2$ (Wei et al., 2001). More recently, a transcriptome-wide study together with bioinformatics predictions showed a major role for CsrA as an mRNA stabilization factor in *E. coli* (M9 minimal media, doubling time of 6.9 h) for more than a thousand transcripts, of which many were predicted to have at least one putative CsrA binding site (Esquerre et al., 2016). CsrA could directly bind transcripts and protect them from RNases, or could affect mRNA stability indirectly by modulating expression or activity of other post-transcriptional regulators, e.g., the RNA chaperone Hfq, encoded by *hfq*. In *E. coli*, CsrA can bind the *hfq* mRNA at a single binding site that overlaps its SD region, preventing ribosome access and decreasing its half-life, however, in stationary phase CsrA is sequestered, allowing higher expression of Hfq (Baker et al., 2007). Regulatory roles for CsrA in gram-positive bacteria have only recently been reported. In *B. subtilis*, CsrA mediates the interaction of the sRNA SR1 and the *ahrC* mRNA, encoding a transcription regulator of arginine metabolism, to regulate the expression of the arginine catabolic operons (Muller et al., 2019). However, CsrA-SR1 only mildly increased *ahrC* half-life, and it had no impact on SR1 degradation, indicating that the regulation was primarily at the level of protein synthesis (Muller et al., 2019).

The homohexameric Hfq, highly studied in *E. coli* and present in a large number of bacteria (Sun et al., 2002), is an important regulator of mRNA-sRNA pairing. The multiple roles of Hfq include modulation of sRNA-mediated translation blockage or promotion, and regulation of transcript degradation as a direct consequence of altered translation or through

translation-independent mechanisms. For example, guiding a cognate sRNA to the 5′ region of mRNAs can result either in translation disruption by preventing the 30S subunit from binding (**Figure 4B**), or the opposite outcome by disruption of stem-loops that inhibit its binding (Wassarman et al., 2001; Arluison et al., 2002; Moller et al., 2002; Schumacher et al., 2002; Zhang et al., 2003; Afonyushkin et al., 2005; Sittka et al., 2008). Hfq can also allow RNase E access to specific mRNAs, or modulate the synthesis of Poly(A) tails, assisting PNPase in 3′–5′ degradation, as it will be discussed shortly. The physical properties, sequence specificity, protein interaction partners, sRNAs/mRNAs binding kinetics, and other important aspects of Hfq function will not be described here, as they are well described elsewhere; we refer the reader to the following detailed reviews (Vogel and Luisi, 2011; Updegrove et al., 2016; Kavita et al., 2018; Santiago-Frangos and Woodson, 2018).

A common outcome of Hfq sRNA/mRNA interactions is specific regulation of mRNA half-life (**Figure 4C**). For example, the destabilization of *ptsG*, encoding a glucose permease, in *E. coli* is mediated by the sRNA SgrS as a response to phosphosugar accumulation (Vanderpool and Gottesman, 2004). Similarly, degradation of *ompA* was also shown to be impacted by the specific binding of the sRNA MicA to its translational start site, blocking binding of the 30S ribosomal subunit and recruiting Hfq to promote RNase E cleavage (Lundberg et al., 1990; Vytvytska et al., 2000; Udekwu et al., 2005). While the regulatory roles of Hfq are widely accepted for other gram-negative bacteria as well (Sonnleitner et al., 2006; Cui et al., 2013), in gram-positive bacteria Hfq is less well characterized. Hfq rescue experiments in *E. coli* and *S. enterica* serovar Typhimurium using Hfq from *B. subtilis* and *S. aureus*, respectively, failed at rescuing the phenotypes (Vecerek et al., 2008; Rochat et al., 2012). These findings suggest important structural and/or functional differences in Hfq across evolutionarily divergent groups of bacteria. A study in *B. subtilis* found that the absence of Hfq does not impair growth under almost 2000 conditions including different carbon, nitrogen, phosphorus and sulfur sources, osmolarity or pH changes in a large phenotypic analysis (Rochat et al., 2015). Similar findings were shown for *S. aureus* (Bohn et al., 2007). However, Hfq became necessary for survival in stationary phase (Hammerle et al., 2014; Rochat et al., 2015). Surprisingly, the absence of Hfq in rich media conditions did not alter the transcriptome of *B. subtilis* (Rochat et al., 2015), while in minimal media, 68 mRNAs and a single sRNA were affected (Hammerle et al., 2014). Both of these studies reported transcriptome changes in the absence of Hfq for *B. subtilis* in stationary phase, particularly for sporulation and TA systems. Nevertheless, these changes do not necessarily confer fitness or increased survival (Rochat et al., 2015). Overall, while Hfq was shown to impact the *B. subtilis* transcriptome under certain stress conditions, its role as a regulator of transcript stability seems to greatly vary across species. In another gram-positive, the pathogen *Listeria monocytogenes*, Hfq interacts with the sRNA LhrA, increasing its stability and controlling the fate of its target mRNAs. But, ∼50 other sRNA seem to function in an Hfq-independent manner (Christiansen et al., 2006; Nielsen et al., 2010; Nielsen et al., 2011). Unexpectedly, hypoxia, stationary phase and low temperature (30°C) did not affect sRNA levels in a Δ*hfq* strain (Toledo-Arana et al., 2009). Hence, it seems that Hfq may have a smaller role in control of mRNA stability, and an overall restricted role in sRNA/mRNA regulation in gram-positive bacteria; and it appears to not be required at all in some bacteria, such as mycobacteria, that lack identified Hfq orthologs (Sun et al., 2002).

mRNA Folding Alters mRNA Decay

mRNA secondary structures can modulate translation and transcript stability (**Figure 3D**). Previously, we have discussed how specific 5′ UTR folding prevents RNase and ribosome accessibility to the *lysC* transcript (Caron et al., 2012). In other transcripts, secondary structures can also prevent RNase E from carrying out the first endonucleolytic cleavage, delaying subsequent steps in the decay pathways. In *Rhodobacter capsulatus*, formation of multiple hairpins can prevent endonucleolytic cleavage of the *puf* operon (Klug and Cohen, 1990). A stem-loop at the 5′ UTR confers stability to *recA*, coding for the nucleoprotein filament RecA in *Acinetobacter baumannii* (Ching et al., 2017), as well as *vacA*, coding for vacuolating cytotoxin A in *Helicobacter pylori* (Amilon et al., 2015). In the case of *vacA*, the stem-loop is also essential for transcript stabilization in acidic and osmotic stress (Amilon et al., 2015). The distance between the start codon and secondary structures can also affect mRNA half-life, as was shown for the ΔermC mRNA in *B. subtilis*, where placing a stem-loop too close to the SD decreased transcript stability (Sharp and Bechhofer, 2005). Secondary structure at transcript 3′ ends also affects stability. The mRNA 3′ end hairpins formed by Rho-independent transcriptional terminators typically stabilize transcripts, as 3′–5′ RNases have difficulty initiating decay without a single-stranded substrate (Adhya et al., 1979; Farnham and Platt, 1981; Abe and Aiba, 1996). In *E. coli*, the poly(A) polymerase (PAP I) is an enzyme responsible for synthesizing poly(A) tails in mRNA (Li et al., 1998). The addition of poly(A) tails to bacterial mRNAs facilitates degradation of transcripts with 3′ hairpins, allowing PNPase – an enzyme that also has a minor polyadenylation role – and other enzymes to carry out exonucleolytic activity (Donovan and Kushner, 1986; Blum et al., 1999; **Figure 6**).

Thus, it is possible for poly(A) tails to act as regulators of mRNA stability, making PAP I a promising candidate for posttranscriptional regulation. However, while this enzyme has been characterized in *E. coli*, PAP I homologs in *B. subtilis* have not yet been identified (Campos-Guillen et al., 2005). An interesting role of Hfq in *E. coli* was reported for transcripts carrying long poly(A) tails, as binding to the tail prevents the access of PNPase, thereby increasing mRNA stability (Hajnsdorf and Regnier, 2000; Folichon et al., 2005). However, on shorter poly(A) tails (<10 nt), Hfq has poor accessibility, making the transcripts susceptible to the activity of PNPase and RNase II (Regnier and Hajnsdorf, 2013). Interestingly, in *E. coli*, the absence of PAP I disrupts the regulatory role of some sRNAs, leading to an unexpected destabilization of some sRNAs and transcripts, e.g., RyhB and MicA (Sinha et al., 2018). This appears to result from accumulation of transcripts that are normally degraded in a PAP I-dependent fashion. The accumulated

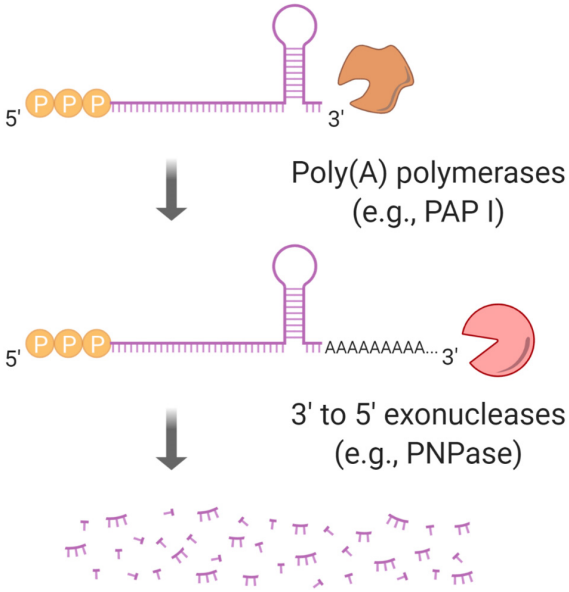


FIGURE 6 | Polyadenylation regulates mRNA half-life. Stem-loops at mRNA 3′ ends block 3′–5′ exoribonucleases such as PNPase. PAP I, a poly(A) polymerase, can facilitate an exoribonuclease "grip" by synthesizing a poly(A) tail.

transcripts participate in non-specific interactions with sRNAs, leading to degradation of the sRNA-mRNA pairs. Thus, it is suggested that many PAP I targets are transcripts that do not normally interact with sRNAs (Cameron et al., 2019).

Regulation of PNPase abundance has been shown for *E. coli*, as its transcript *pnp* is post-transcriptionally regulated by its own product and RNase III. This mechanism can be disrupted by transcript association with the ribosomal protein S1 (Briani et al., 2008; Carzaniga et al., 2015). Moreover, an increase of the pool of polyadenylated transcripts increases *pnp* half-life, an effect attributed to PNPase titration (Mohanty and Kushner, 2000, 2002). Regardless of this autoregulatory characteristic, changes in PNPase abundance were not detected as a response to hypoxic stress in *M. smegmatis* (Vargas-Blanco et al., 2019), despite increased mRNA stability. While these findings suggest that regulation by mRNA polyadenylation via PNPase abundance is not a mechanism of transcriptome stabilization in mycobacteria, it is possible that polyadenylation activity by other enzymes, such as PcnA and PcnB, Adilakshmi et al. (2000) might have a role in regulation of mRNA turnover in stress. Further research is needed to investigate this possibility.

The Relationship Between mRNA Abundance and mRNA Decay Rates

In bacteria, the steady-state mRNA concentration is a function of transcription rates and transcript degradation rates, and to a lesser extent, of mRNA dilution. The contribution of mRNA dilution occurring during cell growth is usually ignored, given that doubling times are significantly longer than the median mRNA half-life. For example, in *L. lactis* mRNA half-lives complied with this assumption for 85% of the measured transcripts, at multiple growth rates (Dressaire et al., 2013). In stress conditions, bacterial growth is generally impaired, making the impact of mRNA dilution even smaller and reinforcing the roles of transcription and RNA turnover as the major determinants of mRNA abundance. Also under stress conditions, transcript abundance per cell is typically lower than in conditions of rapid growth. For example, low transcript abundance was observed for *S. aureus* in cold shock, heat shock, and stringent response when compared to unstressed exponential phase (Anderson et al., 2006). The per-cell mRNA concentration decreased in *L. lactis* during progressive adaptation to carbon starvation (Redon et al., 2005a) or isoleucine starvation (Dressaire et al., 2013). The mRNA concentration was three times higher for *E. coli* growing in LB when compared to growth in in minimal media (Bartholomaus et al., 2016). For *M. smegmatis* in early hypoxic stress, the levels of *atpB*, *atpE*, *rnj*, *rraA*, and *sigA* ranged between ∼5 and 75% of those in cells growing in aerobic conditions, and after extended periods of hypoxic or carbon starvation stress, mRNA levels dropped to under 5% of those in log phase (Vargas-Blanco et al., 2019). Given the generally longer half-lives of mRNAs in stressed bacteria, the observation of reduced mRNA concentrations in these conditions may seem counter-intuitive. However, these observations can be reconciled if transcription is also greatly reduced. It is possible that maintaining lower overall mRNA abundance in stress conditions is an adaptive mechanism to favor translation of genes needed for survival of that particular stressor. For example, in a transcriptome-wide study in *E. coli*, mRNA abundance decreased in response to osmotic stress (from ∼2,400 to 1,600 transcripts per cell), a change that may allow specific transcripts – associated with stress response – to be more accessible to ribosomes and translated (Bartholomaus et al., 2016). Interestingly, transcripts with higher copy numbers per cell in normal conditions (>2 copies/cell) were downregulated the most in osmotic stress (Bartholomaus et al., 2016).

The question has arisen if lower mRNA concentrations can actually cause their degradation to be slowed. This idea is suggested by an observation made by several groups, in several species, that in log phase growth, mRNA half-lives are inversely correlated with steady-state abundance (**Figure 7**). For example, a weak negative correlation was shown between mRNA concentration and mRNA half-life for *E. coli* cells in exponential phase (Bernstein et al., 2002). Stronger negative correlations were reported in *L. lactis* (Redon et al., 2005b), and in *M. tuberculosis* (Rustad et al., 2013), both in exponentially growing bacteria. Moreover, in the latter study the overexpression of genes in the DosR regulon resulted in transcripts with shorter half-lives. Other reports in *E. coli* and *L. lactis* showed that cells growing at different growth rates also show a negative correlation between these parameters (Dressaire et al., 2013; Esquerre et al., 2015). For example, changes in growth rate from 0.1 to 0.63 h^{-1} – using chemostats – resulted in increased mRNA levels and a decreased median mRNA half-life from 4.2 to 2.8 min, respectively (Esquerre et al., 2014, 2015). Transcription modulation using five constructs with distinct 5′ UTRs in *lacLM* mRNA also depicted a similar trend in *L. lactis* in exponential phase, and a

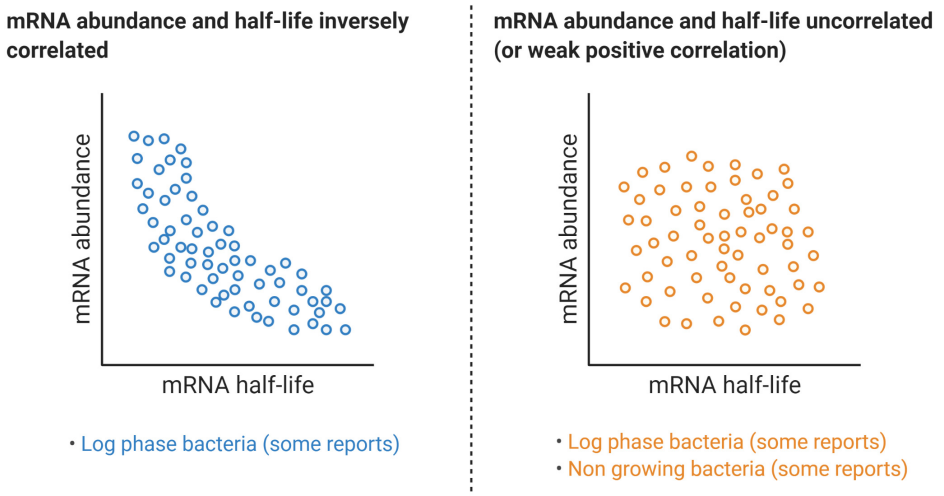


FIGURE 7 | Relationships between mRNA abundance and mRNA decay rates. While some reports have shown a clear negative correlation between a transcript half-life and its abundance, a similar number of reports have found no correlation at all or a modest positive correlation, even for the same organism. **Table 1** compiles transcriptome-wide analyses of mRNA decay in different organisms, techniques used, and information on the reported relationships between mRNA abundance and mRNA half-life.

similar outcome was obtained for *lacZ* in *E. coli*, using P_{BAD}-mediated transcription regulation (Nouaille et al., 2017). Two of the studies described here (Rustad et al., 2013; Nouaille et al., 2017) reported inverse relationships between mRNA abundance and half-life in defined systems where expression was modulated by inducible promotors and growth rate was not affected. This strongly suggested that transcription rate can directly influence degradation rate. However, contradictory findings have been reported.

An *E. coli* transcriptome-wide mRNA half-life study by a different group reported that the rate of mRNA degradation had a very weak positive correlation with mRNA abundance for both exponential phase (R^2 = 0.07) and stationary phase (R^2 = 0.19) (Chen et al., 2015), in contrast to other *E. coli* studies (Bernstein et al., 2002; Esquerre et al., 2014, 2015). In *Bacillus cereus*, mRNA half-life had a positive correlation with expression level (Kristoffersen et al., 2012), while in *Stenotrophomonas maltophilia* and *Chlamydia trachomatis* trachoma and lymphogranuloma venereum biovars no correlations were found (Bernardini and Martinez, 2017; Ferreira et al., 2017). In *M. smegmatis*, induced overexpression of *dCas9* (in the absence of a gene-targeting sgRNA) did not alter its half-life in log phase (Vargas-Blanco et al., 2019). Surprisingly, overexpressing *dCas9* under hypoxic stress increased its mRNA stability by approximately twofold (Vargas-Blanco et al., 2019). Moreover, re-exposure of hypoxic *M. smegmatis* cultures to oxygen caused half-lives of several tested genes to immediately return to log-phase like levels, despite transcription being blocked by rifampicin and transcript levels therefore remaining low (Vargas-Blanco et al., 2019). Other reports have indicated that the relationship between mRNA abundance and half-life differs in various stress conditions. In carbon-starved *L. lactis* there was a positive correlation between mRNA degradation and abundance (Redon et al., 2005b), while the opposite was observed during isoleucine starvation (Dressaire et al., 2013). Work in eukaryotes suggests complexities that could conceivably occur in bacteria as well. In *S. cerevisiae*, under DNA damaging conditions, upregulated genes are usually stabilized and repressed genes are prone to degradation (Shalem et al., 2008). Conversely, under oxidative stress upregulated genes are destabilized, with the opposite scenario for repressed genes (Shalem et al., 2008). Furthermore, an in-depth analysis in that work revealed a trend between these two stress conditions: Genes with a rapid transcriptional regulation show a negative correlation between mRNA abundance and mRNA degradation. On the other hand, genes subject to a slow transcriptional response follow a positive correlation between mRNA abundance and degradation (Shalem et al., 2008).

Clearly, further work is needed to reconcile contradictory findings in bacteria with respect to the relationships between mRNA abundance and stability. Some reported differences may be attributable to differences between species, while others may result from differences in methodology for measuring half-life. Most studies measure half-life by measuring decreases in mRNA abundance following transcription blockage by rifampicin. Variability may arise from the time-points chosen to assay abundance following transcriptional block, given that we and others have reported multiphasic decay kinetics (Hambraeus et al., 2003; Selinger et al., 2003; Chen et al., 2015; Nguyen et al., 2020). Methodology for normalization and for calculating half-lives also vary (see **Table 1**).

THE IMPORTANCE OF RNA DECAY IN CLINICALLY IMPORTANT SPECIES

Pathogenic bacteria have developed mechanisms that allow them to survive often-hostile host environments by sensing cues

and mounting specific responses at both transcriptional and posttranscriptional levels. These pathogens exhibit highly specific responses to some stressors, as well as broader responses to conditions such as energy stress, where resources are preserved by global modulation of processes including translation, protein degradation, transcription, and RNA stabilization (Bohne et al., 1994; Sherman et al., 2001; Park et al., 2003; Christiansen et al., 2004; Wood et al., 2005; Papenfort et al., 2006; Liu et al., 2010; Fritsch et al., 2011; Galagan et al., 2013; Guo et al., 2014; Sievers et al., 2015; Quereda et al., 2018; Ignatov et al., 2020).

In *L. monocytogenes*, PrfA serves as a transcriptional regulator of multiple virulence factors, such as phospholipases PlcA and PlcB, and the toxin listeriolysin O (Leimeister-Wachter et al., 1990, 1991; Quereda et al., 2018). Expression of PrfA itself is regulated by several mechanisms at the translational and transcriptional level. For example, PrfA translation is temperature-regulated by a stem-loop in its transcript, *prfA*, that prevents ribosome access to the SD sequence at 30°C but not at 37°C (Johansson et al., 2002). *prfA* is also regulated by an S-adenosylmethionine riboswitch and its product, the sRNA SreA, that blocks translation after binding the 5′ UTR (Loh et al., 2009). Additionally, while the stem-loop increases *prfA* stability (Loh et al., 2012), the binding of SreA to *prfA* triggers transcript degradation (Loh et al., 2009). Also in *L. monocytogenes*, posttranscriptional regulation of Tcsa, the T cell-stimulating antigen encoded by *tcsA*, was recently reported to be under the control of the sRNA LhrC in a translation-independent manner, by recruiting an undefined RNase (Ross et al., 2019). In *S. aureus* SarA, a histone-like protein, influences mRNA turnover of virulence factors, such as protein A (*spa*) and the collagen adhesion protein (*cna*) during exponential growth (Roberts et al., 2006; Morrison et al., 2012). Also in *S. aureus*, the multifunctional RNAIII binds other RNAs, recruiting RNase III to initiate transcript degradation. Some of RNAIII's targets are *spa*, *coa* (encoding coagulase), *sbi* (encoding the IgG-binding protein Sbi), and *SA1000* (encoding the fibrinogen-binding protein SA1000) (Huntzinger et al., 2005; Boisset et al., 2007; Chevalier et al., 2010), playing an important role in *S. aureus* virulence and response to stress. In *S. enterica*, under low Mg^{2+} conditions synthesis of the antisense AmgR RNA leads to interaction and destabilization of the *mgtC* transcript (encoding the virulence protein MgtC), in an RNase E-dependent manner (Lee and Groisman, 2010). Hence, regulation of the stabilities of specific mRNAs has a major role in the survival and virulence responses of pathogens.

Recent reports have suggested unexpected relationships between RNases and drug resistance. Nonsense and INDEL mutations in *Rv2752c*, encoding RNase J, were associated with drug resistance in a GWAS study that identified resistance-associated mutations in whole-genome sequences of hundreds of *M. tuberculosis* clinical isolates (Hicks et al., 2018), as well as an earlier study performing similar analyses on a smaller set of clinical isolates (Zhang et al., 2013). Another study, reporting whole-genome sequences of 154 *M. leprae* clinical isolates from 25 countries, found a disproportionately high number of polymorphisms in *ML1040c*, encoding RNase D, and *ML1512c*, encoding RNase J (Benjak et al., 2018). These mutations were not directly associated with drug resistance, but appeared to be under positive selection (Benjak et al., 2018).

Global mRNA stabilization is another feature associated with bacterial stress response and non-growing conditions (see **Table 1**). Cells in quiescent states contain relatively low levels of mRNA, with greatly reduced transcriptional and translational activity (Betts et al., 2002; Wood et al., 2005; Kumar et al., 2012; Rittershaus et al., 2013). In some cases, these states share similarities with *B. subtilis* spores, in which the bacteria have dramatically reduced mRNA turnover (Segev et al., 2012). This can be interpreted as a concerted cellular effort to downregulate global gene expression and preserve cellular resources, until encountering a suitable environment to resume growth. At the same time, having paused translational machinery may permit allocation of resources toward specific responses needed to survive a given condition, such as those described in the previous paragraph. Importantly, stress responses that establish and maintain non-growing states not only allow pathogens to survive these stressors, but also induce broad antibiotic tolerance, since most antibiotics are relatively ineffective at killing non-growing cells (for example, Rao et al., 2008). This relationship between growth arrest and antibiotic tolerance may be one of the reasons why months of multidrug therapy are required to prevent relapse in tuberculosis patients, where large numbers of bacteria are likely semi-dormant in hypoxic granulomas (Garton et al., 2008). The apparent universality of mRNA stabilization as a response to energy stress and other stressors that inhibit growth, compared to gene-specific mRNA regulation, brings up fascinating possibilities as a prospective target for therapeutic development. There has been a surge in antimicrobial resistance in recent decades, prompting collaborative efforts between academia and industry to develop new antimicrobials (Ventola, 2015a,b; World Health Organization [WHO], 2019). As we approach an understanding of the mechanisms behind mRNA turnover – and strive to unveil how transcript fate is regulated under stress conditions – we would like to emphasize the essentiality of mRNA degradation in bacteria, and the roles of RNases in the virulence and survival responses of pathogens. Many clinically important antibiotics target transcription and translation, highlighting the potential of targeting these central dogma processes from the opposite angle. In early steps in this direction, a protein degradation inhibitor was found to have strong activity against mycobacteria (Gavrish et al., 2014) and inhibitors of RNase E have been reported (Kime et al., 2015).

CONCLUSION

Transcriptome stabilization as a stress response is widespread across the bacterial domain. This globally concerted response is implicated in gene regulation and survival, as well as pathogenesis in bacteria. We have described and discussed various mechanisms of mRNA degradation and stabilization, many of which have established roles in regulation of specific genes, but have not yet been able to explain transcriptome-wide half-life alterations.

We hope that the information presented here helps to inspire further study that will uncover the mechanism(s) behind global transcriptome stabilization in stress, which so far remains elusive. Finally, we hope to inspire the reader to find these mysteries as scientifically stimulating as we do.

AUTHOR CONTRIBUTIONS

DV-B and SS wrote the manuscript. Both authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the NSF CAREER award 1652756 to SS. DV-B was partially supported by the Fulbright Foreign Student Program.

ACKNOWLEDGMENTS

We thank all members of the Shell lab for technical assistance and helpful discussions. The figures were created with BioRender.com.

REFERENCES

Abe, H., and Aiba, H. (1996). Differential contributions of two elements of rho-independent terminator to transcription termination and mRNA stabilization. *Biochimie* 78, 1035–1042. doi: 10.1016/s0300-9084(97)86727-2

Adhya, S., Sarkar, P., Valenzuela, D., and Maitra, U. (1979). Termination of transcription by *Escherichia coli* RNA polymerase: influence of secondary structure of RNA transcripts on rho-independent and rho-dependent termination. *Proc. Natl. Acad. Sci. U.S.A.* 76, 1613–1617. doi: 10.1073/pnas.76.4.1613

Adilakshmi, T., Ayling, P. D., and Ratledge, C. (2000). Polyadenylylation in mycobacteria: evidence for oligo(dT)-primed cDNA synthesis. *Microbiology* 146(Pt 3), 633–638. doi: 10.1099/00221287-146-3-633

Afonyushkin, T., Vecerek, B., Moll, I., Blasi, U., and Kaberdin, V. R. (2005). Both RNase E and RNase III control the stability of sodB mRNA upon translational inhibition by the small regulatory RNA RyhB. *Nucleic Acids Res.* 33, 1678–1689. doi: 10.1093/nar/gki313

Agaisse, H., and Lereclus, D. (1996). STAB-SD: a Shine-Dalgarno sequence in the 5′ untranslated region is a determinant of mRNA stability. *Mol. Microbiol.* 20, 633–643. doi: 10.1046/j.1365-2958.1996.5401046.x

Aiso, T., Yoshida, H., Wada, A., and Ohki, R. (2005). Modulation of mRNA stability participates in stationary-phase-specific expression of ribosome modulation factor. *J. Bacteriol.* 187, 1951–1958. doi: 10.1128/JB.187.6.1951-1958.2005

Ait-Bara, S., and Carpousis, A. J. (2015). RNA degradosomes in bacteria and chloroplasts: classification, distribution and evolution of RNase E homologs. *Mol. Microbiol.* 97, 1021–1135. doi: 10.1111/mmi.13095

Albertson, M. H., Nyström, T., and Kjelleberg, S. (1990). Functional mRNA half-lives in the marine *Vibrio* sp. S14 during starvation and recovery. *Microbiology* 136, 2195–2199. doi: 10.1099/00221287-136-11-2195

Al-Husini, N., Tomares, D. T., Bitar, O., Childers, W. S., and Schrader, J. M. (2018). alpha-proteobacterial RNA degradosomes assemble liquid-liquid phase-separated RNP bodies. *Mol. Cell* 71, 1027–1039.e14. doi: 10.1016/j.molcel.2018.08.003

Amilon, K. R., Letley, D. P., Winter, J. A., Robinson, K., and Atherton, J. C. (2015). Expression of the *Helicobacter pylori* virulence factor vacuolating cytotoxin A (vacA) is influenced by a potential stem-loop structure in the 5′ untranslated region of the transcript. *Mol. Microbiol.* 98, 831–846. doi: 10.1111/mmi.13160

Anderson, K. L., Roberts, C., Disz, T., Vonstein, V., Hwang, K., Overbeek, R., et al. (2006). Characterization of the *Staphylococcus aureus* heat shock, cold shock, stringent, and SOS responses and their effects on log-phase mRNA turnover. *J. Bacteriol.* 188, 6739–6756. doi: 10.1128/JB.00609-06

Arango, D., Sturgill, D., Alhusaini, N., Dillman, A. A., Sweet, T. J., Hanson, G., et al. (2018). Acetylation of cytidine in mRNA promotes translation efficiency. *Cell* 175, 1872–1886.e24. doi: 10.1016/j.cell.2018.10.030

Arluison, V., Derreumaux, P., Allemand, F., Folichon, M., Hajnsdorf, E., and Regnier, P. (2002). Structural modelling of the m-like protein Hfq from *Escherichia coli*. *J. Mol. Biol.* 320, 705–712. doi: 10.1016/s0022-2836(02)00548-x

Arnold, T. E., Yu, J., and Belasco, J. G. (1998). mRNA stabilization by the ompA 5′ untranslated region: two protective elements hinder distinct pathways for mRNA degradation. *RNA* 4, 319–330.

Arraiano, C. M., Andrade, J. M., Domingues, S., Guinote, I. B., Malecki, M., Matos, R. G., et al. (2010). The critical role of RNA processing and degradation in the control of gene expression. *FEMS Microbiol. Rev.* 34, 883–923. doi: 10.1111/j.1574-6976.2010.00242.x

Artsimovitch, I., Patlan, V., Sekine, S., Vassylyeva, M. N., Hosaka, T., Ochi, K., et al. (2004). Structural basis for transcription regulation by alarmone ppGpp. *Cell* 117, 299–310. doi: 10.1016/s0092-8674(04)00401-5

Atkinson, G. C., Tenson, T., and Hauryliuk, V. (2011). The RelA/SpoT homolog (RSH) superfamily: distribution and functional evolution of ppGpp synthetases and hydrolases across the tree of life. *PLoS One* 6:e23479. doi: 10.1371/journal.pone.0023479

Avarbock, D., Avarbock, A., and Rubin, H. (2000). Differential regulation of opposing RelMtb activities by the aminoacylation state of a tRNA.ribosome.mRNA.RelMtb complex. *Biochemistry* 39, 11640–11648. doi: 10.1021/bi001256k

Baga, M., Goransson, M., Normark, S., and Uhlin, B. E. (1988). Processed mRNA with differential stability in the regulation of *E. coli* pilin gene expression. *Cell* 52, 197–206. doi: 10.1016/0092-8674(88)90508-9

Baker, C. S., Eory, L. A., Yakhnin, H., Mercante, J., Romeo, T., and Babitzke, P. (2007). CsrA inhibits translation initiation of *Escherichia coli* hfq by binding to a single site overlapping the Shine-Dalgarno sequence. *J. Bacteriol.* 189, 5472–5481. doi: 10.1128/JB.00529-07

Bandyra, K. J., Bouvier, M., Carpousis, A. J., and Luisi, B. F. (2013). The social fabric of the RNA degradosome. *Biochim. Biophys. Acta* 1829, 514–522. doi: 10.1016/j.bbagrm.2013.02.011

Bardwell, J. C., Regnier, P., Chen, S. M., Nakamura, Y., Grunberg-Manago, M., and Court, D. L. (1989). Autoregulation of RNase III operon by mRNA processing. *EMBO J.* 8, 3401–3407. doi: 10.1002/j.1460-2075.1989.tb08504.x

Bartholomaus, A., Fedyunin, I., Feist, P., Sin, C., Zhang, G., Valleriani, A., et al. (2016). Bacteria differently regulate mRNA abundance to specifically respond to various stresses. *Philos. Trans. A Math. Phys. Eng. Sci.* 374:20150069. doi: 10.1098/rsta.2015.0069

Battesti, A., and Bouveret, E. (2009). Bacteria possessing two RelA/SpoT-like proteins have evolved a specific stringent response involving the acyl carrier protein-SpoT interaction. *J. Bacteriol.* 191, 616–624. doi: 10.1128/JB.01195-08

Bechhofer, D. H., and Deutscher, M. P. (2019). Bacterial ribonucleases and their roles in RNA metabolism. *Crit. Rev. Biochem. Mol. Biol.* 54, 242–300. doi: 10.1080/10409238.2019.1651816

Bechhofer, D. H., and Dubnau, D. (1987). Induced mRNA stability in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U.S.A.* 84, 498–502. doi: 10.1073/pnas.84.2.498

Bechhofer, D. H., and Zen, K. H. (1989). Mechanism of erythromycin-induced ermC mRNA stability in *Bacillus subtilis*. *J. Bacteriol.* 171, 5803–5811. doi: 10.1128/jb.171.11.5803-5811.1989

Becker, H. F., Motorin, Y., Planta, R. J., and Grosjean, H. (1997). The yeast gene YNL292w encodes a pseudouridine synthase (Pus4) catalyzing the formation of psi55 in both mitochondrial and cytoplasmic tRNAs. *Nucleic Acids Res.* 25, 4493–4499. doi: 10.1093/nar/25.22.4493

Belasco, J. G., Beatty, J. T., Adams, C. W., von Gabain, A., and Cohen, S. N. (1985). Differential expression of photosynthesis genes in *R. capsulata* results from segmental differences in stability within the polycistronic rxcA transcript. *Cell* 40, 171–181. doi: 10.1016/0092-8674(85)90320-4

Benjak, A., Avanzi, C., Singh, P., Loiseau, C., Girma, S., Busso, P., et al. (2018). Phylogenomics and antimicrobial resistance of the leprosy bacillus *Mycobacterium leprae*. *Nat. Commun.* 9:352. doi: 10.1038/s41467-017-02576-z

Bernardini, A., and Martinez, J. L. (2017). Genome-wide analysis shows that RNase G plays a global role in the stability of mRNAs in *Stenotrophomonas maltophilia*. *Sci. Rep.* 7:16016. doi: 10.1038/s41598-017-16091-0

Bernstein, J. A., Khodursky, A. B., Lin, P. H., Lin-Chao, S., and Cohen, S. N. (2002). Global analysis of mRNA decay and abundance in *Escherichia coli* at single-gene resolution using two-color fluorescent DNA microarrays. *Proc. Natl. Acad. Sci. U.S.A.* 99, 9697–9702. doi: 10.1073/pnas.112318199

Bernstein, J. A., Lin, P. H., Cohen, S. N., and Lin-Chao, S. (2004). Global analysis of *Escherichia coli* RNA degradosome function using DNA microarrays. *Proc. Natl. Acad. Sci. U.S.A.* 101, 2758–2763. doi: 10.1073/pnas.0308747101

Betts, J. C., Lukey, P. T., Robb, L. C., McAdam, R. A., and Duncan, K. (2002). Evaluation of a nutrient starvation model of Mycobacterium tuberculosis persistence by gene and protein expression profiling. *Mol. Microbiol.* 43, 717–731. doi: 10.1046/j.1365-2958.2002.02779.x

Bird, J. G., Zhang, Y., Tian, Y., Panova, N., Barvik, I., Greene, L., et al. (2016). The mechanism of RNA 5′ capping with NAD+, NADH and desphospho-CoA. *Nature* 535, 444–447. doi: 10.1038/nature18622

Blum, E., Carpousis, A. J., and Higgins, C. F. (1999). Polyadenylation promotes degradation of 3′-structured RNA by the *Escherichia coli* mRNA degradosome *in vitro*. *J. Biol. Chem.* 274, 4009–4016. doi: 10.1074/jbc.274.7.4009

Bochner, B. R., Lee, P. C., Wilson, S. W., Cutler, C. W., and Ames, B. N. (1984). AppppA and related adenylylated nucleotides are synthesized as a consequence of oxidation stress. *Cell* 37, 225–232. doi: 10.1016/0092-8674(84)90318-0

Bodi, Z., Button, J. D., Grierson, D., and Fray, R. G. (2010). Yeast targets for mRNA methylation. *Nucleic Acids Res.* 38, 5327–5335. doi: 10.1093/nar/gkq266

Boel, G., Letso, R., Neely, H., Price, W. N., Wong, K. H., Su, M., et al. (2016). Codon influence on protein expression in *E. coli* correlates with mRNA levels. *Nature* 529, 358–363. doi: 10.1038/nature16509

Bohn, C., Rigoulay, C., and Bouloc, P. (2007). No detectable effect of RNA-binding protein Hfq absence in *Staphylococcus aureus*. *BMC Microbiol.* 7:10. doi: 10.1186/1471-2180-7-10

Bohne, J., Sokolovic, Z., and Goebel, W. (1994). Transcriptional regulation of prfA and PrfA-regulated virulence genes in *Listeria monocytogenes*. *Mol. Microbiol.* 11, 1141–1150. doi: 10.1111/j.1365-2958.1994.tb00390.x

Boisset, S., Geissmann, T., Huntzinger, E., Fechter, P., Bendridi, N., Possedko, M., et al. (2007). *Staphylococcus aureus* RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. *Genes Dev.* 21, 1353–1366. doi: 10.1101/gad.423507

Bouvet, P., and Belasco, J. G. (1992). Control of RNase E-mediated RNA degradation by 5′-terminal base pairing in *E. coli*. *Nature* 360, 488–491. doi: 10.1038/360488a0

Braun, F., Durand, S., and Condon, C. (2017). Initiating ribosomes and a 5′/3′-UTR interaction control ribonuclease action to tightly couple *B. subtilis* hbs mRNA stability with translation. *Nucleic Acids Res.* 45, 11386–11400. doi: 10.1093/nar/gkx793

Braun, F., Le Derout, J., and Regnier, P. (1998). Ribosomes inhibit an RNase E cleavage which induces the decay of the rpsO mRNA of *Escherichia coli*. *EMBO J.* 17, 4790–4797. doi: 10.1093/emboj/17.16.4790

Brescia, C. C., Kaw, M. K., and Sledjeski, D. D. (2004). The DNA binding protein H-NS binds to and alters the stability of RNA *in vitro* and *in vivo*. *J. Mol. Biol.* 339, 505–514. doi: 10.1016/j.jmb.2004.03.067

Brevet, A., Chen, J., Leveque, F., Plateau, P., and Blanquet, S. (1989). *In vivo* synthesis of adenylylated bis(5′-nucleosidyl) tetraphosphates (Ap4N) by *Escherichia coli* aminoacyl-tRNA synthetases. *Proc. Natl. Acad. Sci. U.S.A.* 86, 8275–8279. doi: 10.1073/pnas.86.21.8275

Briani, F., Curti, S., Rossi, F., Carzaniga, T., Mauri, P., and Deho, G. (2008). Polynucleotide phosphorylase hinders mRNA degradation upon ribosomal protein S1 overexpression in *Escherichia coli*. *RNA* 14, 2417–2429. doi: 10.1261/rna.1123908

Bruscella, P., Shahbabian, K., Laalami, S., and Putzer, H. (2011). RNase Y is responsible for uncoupling the expression of translation factor IF3 from that of the ribosomal proteins L35 and L20 in *Bacillus subtilis*. *Mol. Microbiol.* 81, 1526–1541. doi: 10.1111/j.1365-2958.2011.07793.x

Cahova, H., Winz, M. L., Hofer, K., Nubel, G., and Jaschke, A. (2015). NAD captureSeq indicates NAD as a bacterial cap for a subset of regulatory RNAs. *Nature* 519, 374–377. doi: 10.1038/nature14020

Callaghan, A. J., Marcaida, M. J., Stead, J. A., McDowall, K. J., Scott, W. G., and Luisi, B. F. (2005). Structure of *Escherichia coli* RNase E catalytic domain and implications for RNA turnover. *Nature* 437, 1187–1191. doi: 10.1038/nature04084

Cameron, T. A., Matz, L. M., Sinha, D., and De Lay, N. R. (2019). Polynucleotide phosphorylase promotes the stability and function of Hfq-binding sRNAs by degrading target mRNA-derived fragments. *Nucleic Acids Res.* 47, 8821–8837. doi: 10.1093/nar/gkz616

Campos-Guillen, J., Bralley, P., Jones, G. H., Bechhofer, D. H., and Olmedo-Alvarez, G. (2005). Addition of poly(A) and heteropolymeric 3′ ends in *Bacillus subtilis* wild-type and polynucleotide phosphorylase-deficient strains. *J. Bacteriol.* 187, 4698–4706. doi: 10.1128/JB.187.14.4698-4706.2005

Carabetta, V. J., Tanner, A. W., Greco, T. M., Defrancesco, M., Cristea, I. M., and Dubnau, D. (2013). A complex of YlbF, YmcA and YaaT regulates sporulation, competence and biofilm formation by accelerating the phosphorylation of Spo0A. *Mol. Microbiol.* 88, 283–300. doi: 10.1111/mmi.12186

Carlile, T. M., Rojas-Duran, M. F., Zinshteyn, B., Shin, H., Bartoli, K. M., and Gilbert, W. V. (2014). Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. *Nature* 515, 143–146. doi: 10.1038/nature13802

Caron, M. P., Bastet, L., Lussier, A., Simoneau-Roy, M., Masse, E., and Lafontaine, D. A. (2012). Dual-acting riboswitch control of translation initiation and mRNA decay. *Proc. Natl. Acad. Sci. U.S.A.* 109, E3444–E3453. doi: 10.1073/pnas.1214024109

Carpousis, A. J. (2007). The RNA degradosome of *Escherichia coli*: an mRNA-degrading machine assembled on RNase E. *Annu. Rev. Microbiol.* 61, 71–87. doi: 10.1146/annurev.micro.61.080706.093440

Carpousis, A. J., Van Houwe, G., Ehretsmann, C., and Krisch, H. M. (1994). Copurification of *E. coli* RNAase E and PNPase: evidence for a specific association between two enzymes important in RNA processing and degradation. *Cell* 76, 889–900. doi: 10.1016/0092-8674(94)90363-8

Carzaniga, T., Deho, G., and Briani, F. (2015). RNase III-independent autogenous regulation of *Escherichia coli* polynucleotide phosphorylase via translational repression. *J. Bacteriol.* 197, 1931–1938. doi: 10.1128/JB.00105-15

Celesnik, H., Deana, A., and Belasco, J. G. (2007). Initiation of RNA decay in *Escherichia coli* by 5′ pyrophosphate removal. *Mol. Cell* 27, 79–90. doi: 10.1016/j.molcel.2007.05.038

Chakraburtty, R., and Bibb, M. (1997). The ppGpp synthetase gene (relA) of *Streptomyces coelicolor* A3(2) plays a conditional role in antibiotic production and morphological differentiation. *J. Bacteriol.* 179, 5854–5861. doi: 10.1128/jb.179.18.5854-5861.1997

Chan, C. T., Dyavaiah, M., DeMott, M. S., Taghizadeh, K., Dedon, P. C., and Begley, T. J. (2010). A quantitative systems approach reveals dynamic control of tRNA modifications during cellular stress. *PLoS Genet.* 6:e1001247. doi: 10.1371/journal.pgen.1001247

Chan, C. T., Pang, Y. L., Deng, W., Babu, I. R., Dyavaiah, M., Begley, T. J., et al. (2012). Reprogramming of tRNA modifications controls the oxidative stress response by codon-biased translation of proteins. *Nat. Commun.* 3:937. doi: 10.1038/ncomms1938

Chandran, V., and Luisi, B. F. (2006). Recognition of enolase in the *Escherichia coli* RNA degradosome. *J. Mol. Biol.* 358, 8–15. doi: 10.1016/j.jmb.2006.02.012

Chao, Y., Li, L., Girodat, D., Forstner, K. U., Said, N., Corcoran, C., et al. (2017). *In vivo* cleavage map illuminates the central role of RNase E in coding and non-coding RNA pathways. *Mol. Cell* 65, 39–51. doi: 10.1016/j.molcel.2016.11.002

Chen, H., Previero, A., and Deutscher, M. P. (2019). A novel mechanism of ribonuclease regulation: GcvB and Hfq stabilize the mRNA that encodes RNase BN/Z during exponential phase. *J. Biol. Chem.* 294, 19997–20008. doi: 10.1074/jbc.RA119.011367

Chen, H., Shiroguchi, K., Ge, H., and Xie, X. S. (2015). Genome-wide study of mRNA degradation and transcript elongation in *Escherichia coli*. *Mol. Syst. Biol.* 11:781. doi: 10.15252/msb.20145794

Chen, L. H., Emory, S. A., Bricker, A. L., Bouvet, P., and Belasco, J. G. (1991). Structure and function of a bacterial mRNA stabilizer: analysis of the 5′

untranslated region of ompA mRNA. *J. Bacteriol.* 173, 4578–4586. doi: 10.1128/jb.173.15.4578-4586.1991

Chen, X., Li, A., Sun, B. F., Yang, Y., Han, Y. N., Yuan, X., et al. (2019). 5-methylcytosine promotes pathogenesis of bladder cancer through stabilizing mRNAs. *Nat. Cell Biol.* 21, 978–990. doi: 10.1038/s41556-019-0361-y

Chen, Y. G., Kowtoniuk, W. E., Agarwal, I., Shen, Y., and Liu, D. R. (2009). LC/MS analysis of cellular RNA reveals NAD-linked RNA. *Nat. Chem. Biol.* 5, 879–881. doi: 10.1038/nchembio.235

Chen, Z., Itzek, A., Malke, H., Ferretti, J. J., and Kreth, J. (2013). Multiple roles of RNase Y in *Streptococcus pyogenes* mRNA processing and degradation. *J. Bacteriol.* 195, 2585–2594. doi: 10.1128/jb.00097-13

Chevalier, C., Boisset, S., Romilly, C., Masquida, B., Fechter, P., Geissmann, T., et al. (2010). *Staphylococcus aureus* RNAIII binds to two distant regions of coa mRNA to arrest translation and promote mRNA degradation. *PLoS Pathog.* 6:e1000809. doi: 10.1371/journal.ppat.1000809

Ching, C., Gozzi, K., Heinemann, B., Chai, Y., and Godoy, V. G. (2017). RNA-Mediated cis Regulation in *Acinetobacter baumannii* Modulates Stress-Induced Phenotypic Variation. *J. Bacteriol.* 199:e00799-16. doi: 10.1128/JB.00799-16

Chionh, Y. H., McBee, M., Babu, I. R., Hia, F., Lin, W., Zhao, W., et al. (2016). tRNA-mediated codon-biased translation in mycobacterial hypoxic persistence. *Nat. Commun.* 7:13302. doi: 10.1038/ncomms13302

Cho, K. H. (2017). The structure and function of the gram-positive bacterial RNA degradosome. *Front. Microbiol.* 8:154. doi: 10.3389/fmicb.2017.00154

Christiansen, J. K., Larsen, M. H., Ingmer, H., Sogaard-Andersen, L., and Kallipolitis, B. H. (2004). The RNA-binding protein Hfq of *Listeria monocytogenes*: role in stress tolerance and virulence. *J. Bacteriol.* 186, 3355–3362. doi: 10.1128/JB.186.11.3355-3362.2004

Christiansen, J. K., Nielsen, J. S., Ebersbach, T., Valentin-Hansen, P., Sogaard-Andersen, L., and Kallipolitis, B. H. (2006). Identification of small Hfq-binding RNAs in *Listeria monocytogenes*. *RNA* 12, 1383–1396. doi: 10.1261/rna.49706

Commichau, F. M., Rothe, F. M., Herzberg, C., Wagner, E., Hellwig, D., Lehnik-Habrink, M., et al. (2009). Novel activities of glycolytic enzymes in *Bacillus subtilis*: interactions with essential proteins involved in mRNA processing. *Mol. Cell. Proteomics* 8, 1350–1360. doi: 10.1074/mcp.M800546-MCP200

Condon, C. (2003). RNA processing and degradation in *Bacillus subtilis*. *Microbiol. Mol. Biol. Rev.* 67, 157–174. doi: 10.1128/mmbr.67.2.157-174.2003

Condon, C., Putzer, H., and Grunberg-Manago, M. (1996). Processing of the leader mRNA plays a major role in the induction of thrS expression following threonine starvation in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U.S.A.* 93, 6992–6997. doi: 10.1073/pnas.93.14.6992

Corrigan, R. M., Bellows, L. E., Wood, A., and Grundling, A. (2016). ppGpp negatively impacts ribosome assembly affecting growth and antimicrobial tolerance in Gram-positive bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 113, E1710–E1719. doi: 10.1073/pnas.1522179113

Coste, H., Brevet, A., Plateau, P., and Blanquet, S. (1987). Non-adenylylated bis(5′-nucleosidyl) tetraphosphates occur in Saccharomyces cerevisiae and in *Escherichia coli* and accumulate upon temperature shift or exposure to cadmium. *J. Biol. Chem.* 262, 12096–12103.

Courbet, A., Endy, D., Renard, E., Molina, F., and Bonnet, J. (2015). Detection of pathological biomarkers in human clinical samples via amplifying genetic switches and logic gates. *Sci. Transl. Med.* 7:289ra283. doi: 10.1126/scitranslmed.aaa3601

Cui, M., Wang, T., Xu, J., Ke, Y., Du, X., Yuan, X., et al. (2013). Impact of Hfq on global gene expression and intracellular survival in *Brucella melitensis*. *PLoS One* 8:e71933. doi: 10.1371/journal.pone.0071933

Daeffler, K. N., Galley, J. D., Sheth, R. U., Ortiz-Velez, L. C., Bibb, C. O., Shroyer, N. F., et al. (2017). Engineering bacterial thiosulfate and tetrathionate sensors for detecting gut inflammation. *Mol. Syst. Biol.* 13:923. doi: 10.15252/msb.20167416

Daou-Chabo, R., Mathy, N., Benard, L., and Condon, C. (2009). Ribosomes initiating translation of the hbs mRNA protect it from 5′-to-3′ exoribonucleolytic degradation by RNase J1. *Mol. Microbiol.* 71, 1538–1550. doi: 10.1111/j.1365-2958.2009.06620.x

Deana, A., Celesnik, H., and Belasco, J. G. (2008). The bacterial enzyme RppH triggers messenger RNA degradation by 5′ pyrophosphate removal. *Nature* 451, 355–358. doi: 10.1038/nature06475

DeLoughery, A., Dengler, V., Chai, Y., and Losick, R. (2016). Biofilm formation by *Bacillus subtilis* requires an endoribonuclease-containing multisubunit complex that controls mRNA levels for the matrix gene repressor SinR. *Mol. Microbiol.* 99, 425–437. doi: 10.1111/mmi.13240

DeLoughery, A., Jean-Benoît, L., Losick, R., and Gene-Wei, L. (2018). Maturation of polycistronic mRNAs by the endoribonuclease RNase Y and its associated Y-complex in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U.S.A.* 115, E5585–E5594. doi: 10.1073/pnas.1803283115

Deng, W., Babu, I. R., Su, D., Yin, S., Begley, T. J., and Dedon, P. C. (2015). Trm9-catalyzed tRNA modifications regulate global protein expression by codon-biased translation. *PLoS Genet.* 11:e1005706. doi: 10.1371/journal.pgen.1005706

Deng, X., Chen, K., Luo, G. Z., Weng, X., Ji, Q., Zhou, T., et al. (2015). Widespread occurrence of N6-methyladenosine in bacterial mRNA. *Nucleic Acids Res.* 43, 6557–6567. doi: 10.1093/nar/gkv596

Deng, Z., Liu, Z., Bi, Y., Wang, X., Zhou, D., Yang, R., et al. (2014). Rapid degradation of Hfq-free RyhB in *Yersinia pestis* by PNPase independent of putative ribonucleolytic complexes. *Biomed Res. Int.* 2014:798918. doi: 10.1155/2014/798918

Deutscher, M. P. (2006). Degradation of RNA in bacteria: comparison of mRNA and stable RNA. *Nucleic Acids Res.* 34, 659–666. doi: 10.1093/nar/gkj472

Diwa, A., Bricker, A. L., Jain, C., and Belasco, J. G. (2000). An evolutionarily conserved RNA stem-loop functions as a sensor that directs feedback regulation of RNase E gene expression. *Genes Dev.* 14, 1249–1260.

Diwa, A. A., and Belasco, J. G. (2002). Critical features of a conserved RNA stem-loop important for feedback regulation of RNase E synthesis. *J. Biol. Chem.* 277, 20415–20422. doi: 10.1074/jbc.M202313200

Donovan, W. P., and Kushner, S. R. (1986). Polynucleotide phosphorylase and ribonuclease II are required for cell viability and mRNA turnover in *Escherichia coli* K-12. *Proc. Natl. Acad. Sci. U.S.A.* 83, 120–124. doi: 10.1073/pnas.83.1.120

Drecktrah, D., Hall, L. S., Rescheneder, P., Lybecker, M., and Samuels, D. S. (2018). The stringent response-regulated sRNA transcriptome of *Borrelia burgdorferi*. *Front. Cell. Infect. Microbiol.* 8:231. doi: 10.3389/fcimb.2018.00231

Dressaire, C., Picard, F., Redon, E., Loubiere, P., Queinnec, I., Girbal, L., et al. (2013). Role of mRNA stability during bacterial adaptation. *PLoS One* 8:e59059. doi: 10.1371/journal.pone.0059059

Druzhinin, S. Y., Tran, N. T., Skalenko, K. S., Goldman, S. R., Knoblauch, J. G., Dove, S. L., et al. (2015). A conserved pattern of primer-dependent transcription initiation in *Escherichia coli* and *Vibrio cholerae* revealed by 5′, RNA-seq. *PLoS Genet.* 11:e1005348. doi: 10.1371/journal.pgen.1005348

Dubnau, E. J., Carabetta, V. J., Tanner, A. W., Miras, M., Diethmaier, C., and Dubnau, D. (2016). A protein complex supports the production of Spo0A-P and plays additional roles for biofilms and the K-state in *Bacillus subtilis*. *Mol. Microbiol.* 101, 606–624. doi: 10.1111/mmi.13411

Durand, S., Gilet, L., Bessieres, P., Nicolas, P., and Condon, C. (2012). Three essential ribonucleases-RNase Y, J1, and III-control the abundance of a majority of *Bacillus subtilis* mRNAs. *PLoS Genet.* 8:e1002520. doi: 10.1371/journal.pgen.1002520

Edelheit, S., Schwartz, S., Mumbach, M. R., Wurtzel, O., and Sorek, R. (2013). Transcriptome-wide mapping of 5-methylcytidine RNA modifications in bacteria, archaea, and yeast reveals m5C within archaeal mRNAs. *PLoS Genet.* 9:e1003602. doi: 10.1371/journal.pgen.1003602

Emory, S. A., and Belasco, J. G. (1990). The ompA 5′ untranslated RNA segment functions in *Escherichia coli* as a growth-rate-regulated mRNA stabilizer whose activity is unrelated to translational efficiency. *J. Bacteriol.* 172, 4472–4481. doi: 10.1128/jb.172.8.4472-4481.1990

Emory, S. A., Bouvet, P., and Belasco, J. G. (1992). A 5′-terminal stem-loop structure can stabilize mRNA in *Escherichia coli*. *Genes Dev.* 6, 135–148. doi: 10.1101/gad.6.1.135

Esquerre, T., Bouvier, M., Turlan, C., Carpousis, A. J., Girbal, L., and Cocaign-Bousquet, M. (2016). The Csr system regulates genome-wide mRNA stability and transcription and thus gene expression in *Escherichia coli*. *Sci. Rep.* 6:25057. doi: 10.1038/srep25057

Esquerre, T., Laguerre, S., Turlan, C., Carpousis, A. J., Girbal, L., and Cocaign-Bousquet, M. (2014). Dual role of transcription and transcript stability in the regulation of gene expression in *Escherichia coli* cells cultured on glucose at different growth rates. *Nucleic Acids Res.* 42, 2460–2472. doi: 10.1093/nar/gkt1150

Esquerre, T., Moisan, A., Chiapello, H., Arike, L., Vilu, R., Gaspin, C., et al. (2015). Genome-wide investigation of mRNA lifetime determinants in *Escherichia coli* cells cultured at different growth rates. *BMC Genomics* 16:275.

Even, S., Pellegrini, O., Zig, L., Labas, V., Vinh, J., Brechemmier-Baey, D., et al. (2005). Ribonucleases J1 and J2: two novel endoribonucleases in *B. subtilis* with functional homology to *E.coli* RNase E. *Nucleic Acids Res.* 33, 2141–2152. doi: 10.1093/nar/gki505

Eyler, D. E., Franco, M. K., Batool, Z., Wu, M. Z., Dubuke, M. L., Dobosz-Bartoszek, M., et al. (2019). Pseudouridinylation of mRNA coding sequences alters translation. *Proc. Natl. Acad. Sci. U.S.A.* 116, 23068–23074. doi: 10.1073/pnas.1821754116

Faner, M. A., and Feig, A. L. (2013). Identifying and characterizing Hfq-RNA interactions. *Methods* 63, 144–159. doi: 10.1016/j.ymeth.2013.04.023

Farnham, P. J., and Platt, T. (1981). Rho-independent termination: dyad symmetry in DNA causes RNA polymerase to pause during transcription *in vitro*. *Nucleic Acids Res.* 9, 563–577. doi: 10.1093/nar/9.3.563

Farr, S. B., Arnosti, D. N., Chamberlin, M. J., and Ames, B. N. (1989). An apaH mutation causes AppppA to accumulate and affects motility and catabolite repression in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 86, 5010–5014. doi: 10.1073/pnas.86.13.5010

Ferreira, R., Borges, V., Borrego, M. J., and Gomes, J. P. (2017). Global survey of mRNA levels and decay rates of *Chlamydia trachomatis* trachoma and lymphogranuloma venereum biovars. *Heliyon* 3:e00364. doi: 10.1016/j.heliyon.2017.e00364

Folichon, M., Allemand, F., Regnier, P., and Hajnsdorf, E. (2005). Stimulation of poly(A) synthesis by *Escherichia coli* poly(A)polymerase I is correlated with Hfq binding to poly(A) tails. *FEBS J.* 272, 454–463. doi: 10.1111/j.1742-4658.2004.04485.x

Frederix, M., and Downie, A. J. (2011). Quorum sensing: regulating the regulators. *Adv. Microb. Physiol.* 58, 23–80. doi: 10.1016/B978-0-12-381043-4.00002-7

Frindert, J., Zhang, Y., Nubel, G., Kahloon, M., Kolmar, L., Hotz-Wagenblatt, A., et al. (2018). Identification, biosynthesis, and decapping of NAD-capped RNAs in *B. subtilis*. *Cell Rep.* 24, 1890–1901.e8. doi: 10.1016/j.celrep.2018.07.047

Fritsch, F., Mauder, N., Williams, T., Weiser, J., Oberle, M., and Beier, D. (2011). The cell envelope stress response mediated by the LiaFSRLm three-component system of *Listeria monocytogenes* is controlled via the phosphatase activity of the bifunctional histidine kinase LiaSLm. *Microbiology* 157(Pt 2), 373–386. doi: 10.1099/mic.0.044776-0

Fry, M., Israeli-Reches, M., and Artman, M. (1972). Stabilization and breakdown of *Escherichia coli* messenger ribonucleic acid in the presence of chloramphenicol. *Biochemistry* 11, 3054–3059. doi: 10.1021/bi00766a017

Galagan, J. E., Minch, K., Peterson, M., Lyubetskaya, A., Azizi, E., Sweet, L., et al. (2013). The *Mycobacterium tuberculosis* regulatory network and hypoxia. *Nature* 499, 178–183. doi: 10.1038/nature12337

Garton, N. J., Waddell, S. J., Sherratt, A. L., Lee, S. M., Smith, R. J., Senner, C., et al. (2008). Cytological and transcript analyses reveal fat and lazy persister-like bacilli in tuberculous sputum. *PLoS Med.* 5:e75. doi: 10.1371/journal.pmed.0050075

Gatewood, M. L., and Jones, G. H. (2010). (p)ppGpp inhibits polynucleotide phosphorylase from streptomyces but not from *Escherichia coli* and increases the stability of bulk mRNA in *Streptomyces coelicolor*. *J. Bacteriol.* 192, 4275–4280. doi: 10.1128/JB.00367-10

Gavrish, E., Sit, C. S., Cao, S., Kandror, O., Spoering, A., Peoples, A., et al. (2014). Lassomycin, a ribosomally synthesized cyclic peptide, kills *Mycobacterium tuberculosis* by targeting the ATP-dependent protease ClpC1P1P2. *Chem. Biol.* 21, 509–518. doi: 10.1016/j.chembiol.2014.01.014

Gentry, D. R., Hernandez, V. J., Nguyen, L. H., Jensen, D. B., and Cashel, M. (1993). Synthesis of the stationary-phase sigma factor sigma s is positively regulated by ppGpp. *J. Bacteriol.* 175, 7982–7989. doi: 10.1128/jb.175.24.7982-7989.1993

Georgellis, D., Barlow, T., Arvidson, S., and von Gabain, A. (1993). Retarded RNA turnover in *Escherichia coli*: a means of maintaining gene expression during anaerobiosis. *Mol. Microbiol.* 9, 375–381. doi: 10.1111/j.1365-2958.1993.tb01698.x

Glatz, E., Nilsson, R. P., Rutberg, L., and Rutberg, B. (1996). A dual role for the *Bacillus subtilis* glpD leader and the GlpP protein in the regulated expression of glpD: antitermination and control of mRNA stability. *Mol. Microbiol.* 19, 319–328. doi: 10.1046/j.1365-2958.1996.376903.x

Gualerzi, C. O., Giuliodori, A. M., and Pon, C. L. (2003). Transcriptional and post-transcriptional control of cold-shock genes. *J. Mol. Biol.* 331, 527–539. doi: 10.1016/s0022-2836(03)00732-0

Guo, M. S., Updegrove, T. B., Gogol, E. B., Shabalina, S. A., Gross, C. A., and Storz, G. (2014). MicL, a new sigmaE-dependent sRNA, combats envelope stress by repressing synthesis of Lpp, the major outer membrane lipoprotein. *Genes Dev.* 28, 1620–1634. doi: 10.1101/gad.243485.114

Gutgsell, N., Englund, N., Niu, L., Kaya, Y., Lane, B. G., and Ofengand, J. (2000). Deletion of the *Escherichia coli* pseudouridine synthase gene truB blocks formation of pseudouridine 55 in tRNA *in vivo*, does not affect exponential growth, but confers a strong selective disadvantage in competition with wild-type cells. *RNA* 6, 1870–1881. doi: 10.1017/s1355838200001588

Hadjeras, L., Poljak, L., Bouvier, M., Morin-Ogier, Q., Canal, I., Cocaign-Bousquet, M., et al. (2019). Detachment of the RNA degradosome from the inner membrane of *Escherichia coli* results in a global slowdown of mRNA degradation, proteolysis of RNase E and increased turnover of ribosome-free transcripts. *Mol. Microbiol.* 111, 1715–1731. doi: 10.1111/mmi.14248

Hajnsdorf, E., and Regnier, P. (2000). Host factor Hfq of *Escherichia coli* stimulates elongation of poly(A) tails by poly(A) polymerase I. *Proc. Natl. Acad. Sci. U.S.A.* 97, 1501–1505. doi: 10.1073/pnas.040549897

Hambraeus, G., Karhumaa, K., and Rutberg, B. (2002). A 5′ stem-loop and ribosome binding but not translation are important for the stability of Bacillus subtilis aprE leader mRNA. *Microbiology* 148(Pt 6), 1795–1803. doi: 10.1099/00221287-148-6-1795

Hambraeus, G., Persson, M., and Rutberg, B. (2000). The aprE leader is a determinant of extreme mRNA stability in *Bacillus subtilis*. *Microbiology* 146(Pt 12), 3051–3059. doi: 10.1099/00221287-146-12-3051

Hambraeus, G., von Wachenfeldt, C., and Hederstedt, L. (2003). Genome-wide survey of mRNA half-lives in *Bacillus subtilis* identifies extremely stable mRNAs. *Mol. Genet. Genomics* 269, 706–714. doi: 10.1007/s00438-003-0883-6

Hammerle, H., Amman, F., Vecerek, B., Stulke, J., Hofacker, I., and Blasi, U. (2014). Impact of Hfq on the *Bacillus subtilis* transcriptome. *PLoS One* 9:e98661. doi: 10.1371/journal.pone.0098661

Hamouche, L., Billaudeau, C., Rocca, A., Chastanet, A., Ngo, S., Laalami, S., et al. (2020). Dynamic membrane localization of RNase Y in *Bacillus subtilis*. *mBio* 11:e03337-19. doi: 10.1128/mBio.03337-19

Hausmann, S., Guimaraes, V. A., Garcin, D., Baumann, N., Linder, P., and Redder, P. (2017). Both exo- and endo-nucleolytic activities of RNase J1 from *Staphylococcus aureus* are manganese dependent and active on triphosphorylated 5′-ends. *RNA Biol.* 14, 1431–1443. doi: 10.1080/15476286.2017.1300223

Hicks, N. D., Yang, J., Zhang, X., Zhao, B., Grad, Y. H., Liu, L., et al. (2018). Clinically prevalent mutations in *Mycobacterium tuberculosis* alter propionate metabolism and mediate multidrug tolerance. *Nat. Microbiol.* 3, 1032–1042. doi: 10.1038/s41564-018-0218-3

Hoernes, T. P., Clementi, N., Faserl, K., Glasner, H., Breuker, K., Lindner, H., et al. (2016). Nucleotide modifications within bacterial messenger RNAs regulate their translation and are able to rewire the genetic code. *Nucleic Acids Res.* 44, 852–862. doi: 10.1093/nar/gkv1182

Hofer, K., Li, S., Abele, F., Frindert, J., Schlotthauer, J., Grawenhoff, J., et al. (2016). Structure and function of the bacterial decapping enzyme NudC. *Nat. Chem. Biol.* 12, 730–734. doi: 10.1038/nchembio.2132

Huang, T., Chen, W., Liu, J., Gu, N., and Zhang, R. (2019). Genome-wide identification of mRNA 5-methylcytosine in mammals. *Nat. Struct. Mol. Biol.* 26, 380–388. doi: 10.1038/s41594-019-0218-x

Hudecek, O., Benoni, R., Reyes-Gutierrez, P. E., Culka, M., Sanderova, H., Hubalek, M., et al. (2020). Dinucleoside polyphosphates act as 5′-RNA caps in bacteria. *Nat. Commun.* 11:1052. doi: 10.1038/s41467-020-14896-8

Hue, K. K., Cohen, S. D., and Bechhofer, D. H. (1995). A polypurine sequence that acts as a 5′ mRNA stabilizer in *Bacillus subtilis*. *J. Bacteriol.* 177, 3465–3471. doi: 10.1128/jb.177.12.3465-3471.1995

Hunt, A., Rawlins, J. P., Thomaides, H. B., and Errington, J. (2006). Functional analysis of 11 putative essential genes in *Bacillus subtilis*. *Microbiology* 152(Pt 10), 2895–2907. doi: 10.1099/mic.0.29152-0

Huntzinger, E., Boisset, S., Saveanu, C., Benito, Y., Geissmann, T., Namane, A., et al. (2005). *Staphylococcus aureus* RNAIII and the endoribonuclease III coordinately regulate spa gene expression. *EMBO J.* 24, 824–835. doi: 10.1038/sj.emboj.7600572

Ignatov, D., Vaitkevicius, K., Durand, S., Cahoon, L., Sandberg, S. S., Liu, X., et al. (2020). An mRNA-mRNA interaction couples expression of a virulence

factor and its chaperone in *Listeria monocytogenes*. *Cell Rep.* 30, 4027–4040.e7. doi: 10.1016/j.celrep.2020.03.006

Ishida, K., Kunibayashi, T., Tomikawa, C., Ochi, A., Kanai, T., Hirata, A., et al. (2011). Pseudouridine at position 55 in tRNA controls the contents of other modified nucleotides for low-temperature adaptation in the extreme-thermophilic eubacterium *Thermus thermophilus*. *Nucleic Acids Res.* 39, 2304–2318. doi: 10.1093/nar/gkq1180

Johansson, J., Mandin, P., Renzoni, A., Chiaruttini, C., Springer, M., and Cossart, P. (2002). An RNA thermosensor controls expression of virulence genes in *Listeria monocytogenes*. *Cell* 110, 551–561. doi: 10.1016/s0092-8674(02)00905-4

Jona, G., Choder, M., and Gileadi, O. (2000). Glucose starvation induces a drastic reduction in the rates of both transcription and degradation of mRNA in yeast. *Biochim. Biophys. Acta* 1491, 37–48. doi: 10.1016/s0167-4781(00)00016-6

Jourdan, S. S., and McDowall, K. J. (2008). Sensing of 5′ monophosphate by *Escherichia coli* RNase G can significantly enhance association with RNA and stimulate the decay of functional mRNA transcripts *in vivo*. *Mol. Microbiol.* 67, 102–115. doi: 10.1111/j.1365-2958.2007.06028.x

Joyce, S. A., and Dreyfus, M. (1998). In the absence of translation, RNase E can bypass 5′ mRNA stabilizers in *Escherichia coli*. *J. Mol. Biol.* 282, 241–254. doi: 10.1006/jmbi.1998.2027

Julius, C., and Yuzenkova, Y. (2017). Bacterial RNA polymerase caps RNA with various cofactors and cell wall precursors. *Nucleic Acids Res.* 45, 8282–8290. doi: 10.1093/nar/gkx452

Jurgen, B., Schweder, T., and Hecker, M. (1998). The stability of mRNA from the gsiB gene of *Bacillus subtilis* is dependent on the presence of a strong ribosome binding site. *Mol. Gen. Genet.* 258, 538–545. doi: 10.1007/s004380050765

Kavita, K., de Mets, F., and 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

Khemici, V., Poljak, L., Luisi, B. F., and Carpousis, A. J. (2008). The RNase E of *Escherichia coli* is a membrane-binding protein. *Mol. Microbiol.* 70, 799–813. doi: 10.1111/j.1365-2958.2008.06454.x

Khemici, V., Prados, J., Linder, P., and Redder, P. (2015). Decay-initiating endoribonucleolytic cleavage by RNase y is kept under tight control via sequence preference and sub-cellular localisation. *PLoS Genet.* 11:e1005577. doi: 10.1371/journal.pgen.1005577

Kido, M., Yamanaka, K., Mitani, T., Niki, H., Ogura, T., and Hiraga, S. (1996). RNase E polypeptides lacking a carboxyl-terminal half suppress a mukB mutation in *Escherichia coli*. *J. Bacteriol.* 178, 3917–3925. doi: 10.1128/jb.178.13.3917-3925.1996

Kime, L., Jourdan, S. S., Stead, J. A., Hidalgo-Sastre, A., and McDowall, K. J. (2010). Rapid cleavage of RNA by RNase E in the absence of 5′ monophosphate stimulation. *Mol. Microbiol.* 76, 590–604. doi: 10.1111/j.1365-2958.2009.06935.x

Kime, L., Vincent, H. A., Gendoo, D. M., Jourdan, S. S., Fishwick, C. W., Callaghan, A. J., et al. (2015). The first small-molecule inhibitors of members of the ribonuclease E family. *Sci. Rep.* 5:8028. doi: 10.1038/srep08028

Kinghorn, S. M., O'Byrne, C. P., Booth, I. R., and Stansfield, I. (2002). Physiological analysis of the role of truB in *Escherichia coli*: a role for tRNA modification in extreme temperature resistance. *Microbiology* 148(Pt 11), 3511–3520. doi: 10.1099/00221287-148-11-3511

Klug, G., and Cohen, S. N. (1990). Combined actions of multiple hairpin loop structures and sites of rate-limiting endonucleolytic cleavage determine differential degradation rates of individual segments within polycistronic puf operon mRNA. *J. Bacteriol.* 172, 5140–5146. doi: 10.1128/jb.172.9.5140-5146.1990

Kowtoniuk, W. E., Shen, Y., Heemstra, J. M., Agarwal, I., and Liu, D. R. (2009). A chemical screen for biological small molecule-RNA conjugates reveals CoA-linked RNA. *Proc. Natl. Acad. Sci. U.S.A.* 106, 7768–7773. doi: 10.1073/pnas.0900528106

Krishnan, S., Petchiappan, A., Singh, A., Bhatt, A., and Chatterji, D. (2016). R-loop induced stress response by second (p)ppGpp synthetase in *Mycobacterium smegmatis*: functional and domain interdependence. *Mol. Microbiol.* 102, 168–182. doi: 10.1111/mmi.13453

Kristoffersen, S. M., Haase, C., Weil, M. R., Passalacqua, K. D., Niazi, F., Hutchison, S. K., et al. (2012). Global mRNA decay analysis at single nucleotide resolution reveals segmental and positional degradation patterns in a Gram-positive bacterium. *Genome Biol.* 13:R30. doi: 10.1186/gb-2012-13-4-r30

Kumar, A., Majid, M., Kunisch, R., Rani, P. S., Qureshi, I. A., Lewin, A., et al. (2012). Mycobacterium tuberculosis DosR regulon gene Rv0079 encodes a putative, 'dormancy associated translation inhibitor (DATIN)'. *PLoS One* 7:e38709. doi: 10.1371/journal.pone.0038709

Laxman, S., Sutter, B. M., Wu, X., Kumar, S., Guo, X., Trudgian, D. C., et al. (2013). Sulfur amino acids regulate translational capacity and metabolic homeostasis through modulation of tRNA thiolation. *Cell* 154, 416–429. doi: 10.1016/j.cell.2013.06.043

Lee, E. J., and Groisman, E. A. (2010). An antisense RNA that governs the expression kinetics of a multifunctional virulence gene. *Mol. Microbiol.* 76, 1020–1033. doi: 10.1111/j.1365-2958.2010.07161.x

Lee, K., Zhan, X., Gao, J., Qiu, J., Feng, Y., Meganathan, R., et al. (2003). RraA. a protein inhibitor of RNase E activity that globally modulates RNA abundance in *E. coli*. *Cell* 114, 623–634.

Lee, P. C., Bochner, B. R., and Ames, B. N. (1983). AppppA, heat-shock stress, and cell oxidation. *Proc. Natl. Acad. Sci. U.S.A.* 80, 7496–7500. doi: 10.1073/pnas.80.24.7496

Lehnik-Habrink, M., Newman, J., Rothe, F. M., Solovyova, A. S., Rodrigues, C., Herzberg, C., et al. (2011). RNase Y in Bacillus subtilis: a Natively disordered protein that is the functional equivalent of RNase E from *Escherichia coli*. *J. Bacteriol.* 193, 5431–5441. doi: 10.1128/JB.05500-11

Lehnik-Habrink, M., Pfortner, H., Rempeters, L., Pietack, N., Herzberg, C., and Stulke, J. (2010). The RNA degradosome in *Bacillus subtilis*: identification of CshA as the major RNA helicase in the multiprotein complex. *Mol. Microbiol.* 77, 958–971. doi: 10.1111/j.1365-2958.2010.07264.x

Leimeister-Wachter, M., Domann, E., and Chakraborty, T. (1991). Detection of a gene encoding a phosphatidylinositol-specific phospholipase C that is co-ordinately expressed with listeriolysin in *Listeria monocytogenes*. *Mol. Microbiol.* 5, 361–366. doi: 10.1111/j.1365-2958.1991.tb02117.x

Leimeister-Wachter, M., Haffner, C., Domann, E., Goebel, W., and Chakraborty, T. (1990). Identification of a gene that positively regulates expression of listeriolysin, the major virulence factor of *Listeria monocytogenes*. *Proc. Natl. Acad. Sci. U.S.A.* 87, 8336–8340. doi: 10.1073/pnas.87.21.8336

Lenz, G., Doron-Faigenboim, A., Ron, E. Z., Tuller, T., and Gophna, U. (2011). Sequence features of *E. coli* mRNAs affect their degradation. *PLoS One* 6:e28544. doi: 10.1371/journal.pone.0028544

Li, Z., Pandit, S., and Deutscher, M. P. (1998). Polyadenylation of stable RNA precursors *in vivo*. *Proc. Natl. Acad. Sci. U.S.A.* 95, 12158–12162. doi: 10.1073/pnas.95.21.12158

Liang, W., and Deutscher, M. P. (2013). Ribosomes regulate the stability and action of the exoribonuclease RNase R. *J. Biol. Chem.* 288, 34791–34798. doi: 10.1074/jbc.M113.519553

Linder, B., Grozhik, A. V., Olarerin-George, A. O., Meydan, C., Mason, C. E., and Jaffrey, S. R. (2015). Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat. Methods* 12, 767–772. doi: 10.1038/nmeth.3453

Liu, M. Y., Gui, G., Wei, B., Preston, J. F. III, Oakford, L., Yuksel, U., et al. (1997). The RNA molecule CsrB binds to the global regulatory protein CsrA and antagonizes its activity in *Escherichia coli*. *J. Biol. Chem.* 272, 17502–17510. doi: 10.1074/jbc.272.28.17502

Liu, M. Y., Yang, H., and Romeo, T. (1995). The product of the pleiotropic *Escherichia coli* gene csrA modulates glycogen biosynthesis via effects on mRNA stability. *J. Bacteriol.* 177, 2663–2672. doi: 10.1128/jb.177.10.2663-2672.1995

Liu, Y., Wu, N., Dong, J., Gao, Y., Zhang, X., Mu, C., et al. (2010). Hfq is a global regulator that controls the pathogenicity of *Staphylococcus aureus*. *PLoS One* 5:e13069. doi: 10.1371/journal.pone.0013069

Loh, E., Dussurget, O., Gripenland, J., Vaitkevicius, K., Tiensuu, T., Mandin, P., et al. (2009). A trans-acting riboswitch controls expression of the virulence regulator PrfA in *Listeria monocytogenes*. *Cell* 139, 770–779. doi: 10.1016/j.cell.2009.08.046

Loh, E., Memarpour, F., Vaitkevicius, K., Kallipolitis, B. H., Johansson, J., and Sonden, B. (2012). An unstructured 5′-coding region of the prfA mRNA is required for efficient translation. *Nucleic Acids Res.* 40, 1818–1827. doi: 10.1093/nar/gkr850

Lopez, P. J., Marchand, I., Joyce, S. A., and Dreyfus, M. (1999). The C-terminal half of RNase E, which organizes the *Escherichia coli* degradosome, participates

in mRNA degradation but not rRNA processing *in vivo*. *Mol. Microbiol.* 33, 188–199. doi: 10.1046/j.1365-2958.1999.01465.x

Lopez, P. J., Marchand, I., Yarchuk, O., and Dreyfus, M. (1998). Translation inhibitors stabilize *Escherichia coli* mRNAs independently of ribosome protection. *Proc. Natl. Acad. Sci. U.S.A.* 95, 6067–6072. doi: 10.1073/pnas.95.11.6067

Luciano, D. J., and Belasco, J. G. (2020). Np4A alarmones function in bacteria as precursors to RNA caps. *Proc. Natl. Acad. Sci. U.S.A.* 117, 3560–3567. doi: 10.1073/pnas.1914229117

Luciano, D. J., Levenson-Palmer, R., and Belasco, J. G. (2019). Stresses that raise Np4A levels induce protective nucleoside tetraphosphate capping of bacterial RNA. *Mol. Cell* 75, 957–966.e8. doi: 10.1016/j.molcel.2019.05.031

Luciano, D. J., Vasilyev, N., Richards, J., Serganov, A., and Belasco, J. G. (2017). A novel RNA phosphorylation state enables 5′ end-dependent degradation in *Escherichia coli*. *Mol. Cell* 67, 44–54.e6. doi: 10.1016/j.molcel.2017.05.035

Ludwig, H., Homuth, G., Schmalisch, M., Dyka, F. M., Hecker, M., and Stulke, J. (2001). Transcription of glycolytic genes and operons in *Bacillus subtilis*: evidence for the presence of multiple levels of control of the gapA operon. *Mol. Microbiol.* 41, 409–422. doi: 10.1046/j.1365-2958.2001.02523.x

Lundberg, U., von Gabain, A., and Melefors, O. (1990). Cleavages in the 5′ region of the ompA and bla mRNA control stability: studies with an *E. coli* mutant altering mRNA stability and a novel endoribonuclease. *EMBO J.* 9, 2731–2741. doi: 10.1002/j.1460-2075.1990.tb07460.x

Mackie, G. A. (1992). Secondary structure of the mRNA for ribosomal protein S20. Implications for cleavage by ribonuclease E. *J. Biol. Chem.* 267, 1054–1061.

Mackie, G. A. (1998). Ribonuclease E is a 5′-end-dependent endonuclease. *Nature* 395, 720–723. doi: 10.1038/27246

Maggi, N., Pasqualucci, C. R., Ballotta, R., and Sensi, P. (1966). Rifampicin: a new orally active rifamycin. *Chemotherapy* 11, 285–292. doi: 10.1159/000220462

Marcaida, M. J., DePristo, M. A., Chandran, V., Carpousis, A. J., and Luisi, B. F. (2006). The RNA degradosome: life in the fast lane of adaptive molecular evolution. *Trends Biochem. Sci.* 31, 359–365. doi: 10.1016/j.tibs.2006.05.005

Martinez, I., El-Said Mohamed, M., Santos, V. E., Garcia, J. L., Garcia-Ochoa, F., and Diaz, E. (2017). Metabolic and process engineering for biodesulfurization in Gram-negative bacteria. *J. Biotechnol.* 262, 47–55. doi: 10.1016/j.jbiotec.2017.09.004

Martinez-Costa, O. H., Fernandez-Moreno, M. A., and Malpartida, F. (1998). The relA/spoT-homologous gene in *Streptomyces coelicolor* encodes both ribosome-dependent (p)ppGpp-synthesizing and -degrading activities. *J. Bacteriol.* 180, 4123–4132. doi: 10.1128/jb.180.16.4123-4132.1998

Martini, M. C., Zhou, Y., Sun, H., and Shell, S. S. (2019). Defining the transcriptional and post-transcriptional landscapes of *Mycobacterium smegmatis* in aerobic growth and hypoxia. *Front. Microbiol.* 10:591. doi: 10.3389/fmicb.2019.00591

Mathy, N., Hebert, A., Mervelet, P., Benard, L., Dorleans, A., Li de la Sierra-Gallay, I., et al. (2010). *Bacillus subtilis* ribonucleases J1 and J2 form a complex with altered enzyme behaviour. *Mol. Microbiol.* 75, 489–498. doi: 10.1111/j.1365-2958.2009.07004.x

Matsunaga, J., Simons, E. L., and Simons, R. W. (1996). RNase III autoregulation: structure and function of rncO, the posttranscriptional "operator". *RNA* 2, 1228–1240.

Matsunaga, J., Simons, E. L., and Simons, R. W. (1997). *Escherichia coli* RNase III (rnc) autoregulation occurs independently of rnc gene translation. *Mol. Microbiol.* 26, 1125–1135. doi: 10.1046/j.1365-2958.1997.6652007.x

McDowall, K. J., Kaberdin, V. R., Wu, S. W., Cohen, S. N., and Lin-Chao, S. (1995). Site-specific RNase E cleavage of oligonucleotides and inhibition by stem-loops. *Nature* 374, 287–290. doi: 10.1038/374287a0

McDowall, K. J., Lin-Chao, S., and Cohen, S. N. (1994). A+U content rather than a particular nucleotide order determines the specificity of RNase E cleavage. *J. Biol. Chem.* 269, 10790–10796.

McLaren, R. S., Newbury, S. F., Dance, G. S., Causton, H. C., and Higgins, C. F. (1991). mRNA degradation by processive 3′-5′ exoribonucleases *in vitro* and the implications for prokaryotic mRNA decay *in vivo*. *J. Mol. Biol.* 221, 81–95. doi: 10.1016/0022-2836(91)90806-h

Melin, L., Rutberg, L., and von Gabain, A. (1989). Transcriptional and posttranscriptional control of the *Bacillus subtilis* succinate dehydrogenase operon. *J. Bacteriol.* 171, 2110–2115. doi: 10.1128/jb.171.4.2110-2115.1989

Messing, S. A., Gabelli, S. B., Liu, Q., Celesnik, H., Belasco, J. G., Pineiro, S. A., et al. (2009). Structure and biological function of the RNA pyrophosphohydrolase BdRppH from *Bdellovibrio bacteriovorus*. *Structure* 17, 472–481. doi: 10.1016/j.str.2008.12.022

Meyer, K. D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C. E., and Jaffrey, S. R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3′ UTRs and near stop codons. *Cell* 149, 1635–1646. doi: 10.1016/j.cell.2012.05.003

Moffitt, J. R., Pandey, S., Boettiger, A. N., Wang, S., and Zhuang, X. (2016). Spatial organization shapes the turnover of a bacterial transcriptome. *eLife* 5:e13065. doi: 10.7554/eLife.13065

Mohanty, B. K., and Kushner, S. R. (2000). Polynucleotide phosphorylase functions both as a 3′ right-arrow 5′ exonuclease and a poly(A) polymerase in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 97, 11966–11971. doi: 10.1073/pnas.220295997

Mohanty, B. K., and Kushner, S. R. (2002). Polyadenylation of *Escherichia coli* transcripts plays an integral role in regulating intracellular levels of polynucleotide phosphorylase and RNase E. *Mol. Microbiol.* 45, 1315–1324. doi: 10.1046/j.1365-2958.2002.03097.x

Moll, I., Afonyushkin, T., Vytvytska, O., Kaberdin, V. R., and Blasi, U. (2003). Coincident Hfq binding and RNase E cleavage sites on mRNA and small regulatory RNAs. *RNA* 9, 1308–1314. doi: 10.1261/rna.5850703

Moller, T., Franch, T., Hojrup, P., Keene, D. R., Bachinger, H. P., Brennan, R. G., et al. (2002). Hfq: a bacterial Sm-like protein that mediates RNA-RNA interaction. *Mol. Cell* 9, 23–30. doi: 10.1016/s1097-2765(01)00436-1

Montero Llopis, P., Jackson, A. F., Sliusarenko, O., Surovtsev, I., Heinritz, J., Emonet, T., et al. (2010). Spatial organization of the flow of genetic information in bacteria. *Nature* 466, 77–81. doi: 10.1038/nature09152

Morin, M., Enjalbert, B., Ropers, D., Girbal, L., and Cocaign-Bousquet, M. (2020). Genomewide Stabilization of mRNA during a "Feast-to-Famine" Growth Transition in *Escherichia coli*. *mSphere* 5:e00276-20. doi: 10.1128/mSphere.00276-20

Morita, T., and Aiba, H. (2011). RNase E action at a distance: degradation of target mRNAs mediated by an Hfq-binding small RNA in bacteria. *Genes Dev.* 25, 294–298. doi: 10.1101/gad.2030311

Morita, T., Kawamoto, H., Mizota, T., Inada, T., and Aiba, H. (2004). Enolase in the RNA degradosome plays a crucial role in the rapid decay of glucose transporter mRNA in the response to phosphosugar stress in *Escherichia coli*. *Mol. Microbiol.* 54, 1063–1075. doi: 10.1111/j.1365-2958.2004.04329.x

Morrison, J. M., Anderson, K. L., Beenken, K. E., Smeltzer, M. S., and Dunman, P. M. (2012). The staphylococcal accessory regulator, SarA, is an RNA-binding protein that modulates the mRNA turnover properties of late-exponential and stationary phase *Staphylococcus aureus* cells. *Front. Cell. Infect. Microbiol.* 2:26. doi: 10.3389/fcimb.2012.00026

Muller, P., Gimpel, M., Wildenhain, T., and Brantl, S. (2019). A new role for CsrA: promotion of complex formation between an sRNA and its mRNA target in *Bacillus subtilis*. *RNA Biol.* 16, 972–987. doi: 10.1080/15476286.2019.1605811

Murashko, O. N., Kaberdin, V. R., and Lin-Chao, S. (2012). Membrane binding of *Escherichia coli* RNase E catalytic domain stabilizes protein structure and increases RNA substrate affinity. *Proc. Natl. Acad. Sci. U.S.A.* 109, 7019–7024. doi: 10.1073/pnas.1120181109

Murashko, O. N., and Lin-Chao, S. (2017). *Escherichia coli* responds to environmental changes using enolasic degradosomes and stabilized DicF sRNA to alter cellular morphology. *Proc. Natl. Acad. Sci. U.S.A.* 114, E8025–E8034. doi: 10.1073/pnas.1703731114

Murdeshwar, M. S., and Chatterji, D. (2012). MS_RHII-RSD, a dual-function RNase HII-(p)ppGpp synthetase from *Mycobacterium smegmatis*. *J. Bacteriol.* 194, 4003–4014. doi: 10.1128/JB.00258-12

Nakamoto, M. A., Lovejoy, A. F., Cygan, A. M., and Boothroyd, J. C. (2017). mRNA pseudouridylation affects RNA metabolism in the parasite *Toxoplasma gondii*. *RNA* 23, 1834–1849. doi: 10.1261/rna.062794.117

Newbury, S. F., Smith, N. H., Robinson, E. C., Hiles, I. D., and Higgins, C. F. (1987). Stabilization of translationally active mRNA by prokaryotic REP sequences. *Cell* 48, 297–310. doi: 10.1016/0092-8674(87)90433-8

Nguyen, T. G., Vargas-Blanco, D. A., Roberts, L. A., and Shell, S. S. (2020). The impact of leadered and leaderless gene structures on translation efficiency, transcript stability, and predicted transcription rates in *Mycobacterium smegmatis*. *J. Bacteriol.* 202:e00746-19.

Nielsen, J. S., Larsen, M. H., Lillebaek, E. M., Bergholz, T. M., Christiansen, M. H., Boor, K. J., et al. (2011). A small RNA controls expression of the chitinase ChiA in *Listeria monocytogenes*. *PLoS One* 6:e19019. doi: 10.1371/journal.pone.0019019

Nielsen, J. S., Lei, L. K., Ebersbach, T., Olsen, A. S., Klitgaard, J. K., Valentin-Hansen, P., et al. (2010). Defining a role for Hfq in Gram-positive bacteria: evidence for Hfq-dependent antisense regulation in *Listeria monocytogenes*. *Nucleic Acids Res.* 38, 907–919. doi: 10.1093/nar/gkp1081

Nilsson, G., Belasco, J. G., Cohen, S. N., and von Gabain, A. (1984). Growth-rate dependent regulation of mRNA stability in *Escherichia coli*. *Nature* 312, 75–77. doi: 10.1038/312075a0

Nilsson, P., Naureckiene, S., and Uhlin, B. E. (1996). Mutations affecting mRNA processing and fimbrial biogenesis in the *Escherichia coli* pap operon. *J. Bacteriol.* 178, 683–690. doi: 10.1128/jb.178.3.683-690.1996

Nilsson, P., and Uhlin, B. E. (1991). Differential decay of a polycistronic *Escherichia coli* transcript is initiated by RNaseE-dependent endonucleolytic processing. *Mol. Microbiol.* 5, 1791–1799. doi: 10.1111/j.1365-2958.1991.tb01928.x

Nishikura, K., and De Robertis, E. M. (1981). RNA processing in microinjected *Xenopus* oocytes. Sequential addition of base modifications in the spliced transfer RNA. *J. Mol. Biol.* 145, 405–420. doi: 10.1016/0022-2836(81)90212-6

Nouaille, S., Mondeil, S., Finoux, A. L., Moulis, C., Girbal, L., and Cocaign-Bousquet, M. (2017). The stability of an mRNA is influenced by its concentration: a potential physical mechanism to regulate gene expression. *Nucleic Acids Res.* 45, 11711–11724. doi: 10.1093/nar/gkx781

Nurse, K., Wrzesinski, J., Bakin, A., Lane, B. G., and Ofengand, J. (1995). Purification, cloning, and properties of the tRNA psi 55 synthase from *Escherichia coli*. *RNA* 1, 102–112.

Olson, P. D., Kuechenmeister, L. J., Anderson, K. L., Daily, S., Beenken, K. E., Roux, C. M., et al. (2011). Small molecule inhibitors of *Staphylococcus aureus* RnpA alter cellular mRNA turnover, exhibit antimicrobial activity, and attenuate pathogenesis. *PLoS Pathog.* 7:e1001287. doi: 10.1371/journal.ppat.1001287

Ow, M. C., Liu, Q., and Kushner, S. R. (2000). Analysis of mRNA decay and rRNA processing in *Escherichia coli* in the absence of RNase E-based degradosome assembly. *Mol. Microbiol.* 38, 854–866. doi: 10.1046/j.1365-2958.2000.02186.x

Paesold, G., and Krause, M. (1999). Analysis of rpoS mRNA in *Salmonella dublin*: identification of multiple transcripts with growth-phase-dependent variation in transcript stability. *J. Bacteriol.* 181, 1264–1268. doi: 10.1128/jb.181.4.1264-1268.1999

Panneerdoss, S., Eedunuri, V. K., Yadav, P., Timilsina, S., Rajamanickam, S., Viswanadhapalli, S., et al. (2018). Cross-talk among writers, readers, and erasers of m(6)A regulates cancer growth and progression. *Sci. Adv.* 4:eaar8263. doi: 10.1126/sciadv.aar8263

Papenfort, K., Pfeiffer, V., Mika, F., Lucchini, S., Hinton, J. C., and Vogel, J. (2006). SigmaE-dependent small RNAs of *Salmonella* respond to membrane stress by accelerating global omp mRNA decay. *Mol. Microbiol.* 62, 1674–1688. doi: 10.1111/j.1365-2958.2006.05524.x

Park, H. D., Guinn, K. M., Harrell, M. I., Liao, R., Voskuil, M. I., Tompa, M., et al. (2003). Rv3133c/dosR is a transcription factor that mediates the hypoxic response of *Mycobacterium tuberculosis*. *Mol. Microbiol.* 48, 833–843. doi: 10.1046/j.1365-2958.2003.03474.x

Pato, M. L., Bennett, P. M., and von Meyenburg, K. (1973). Messenger ribonucleic acid synthesis and degradation in *Escherichia coli* during inhibition of translation. *J. Bacteriol.* 116, 710–718. doi: 10.1128/jb.116.2.710-718.1973

Pertzev, A. V., and Nicholson, A. W. (2006). Characterization of RNA sequence determinants and antideterminants of processing reactivity for a minimal substrate of *Escherichia coli* ribonuclease III. *Nucleic Acids Res.* 34, 3708–3721. doi: 10.1093/nar/gkl459

Petchiappan, A., Naik, S. Y., and Chatterji, D. (2020). RelZ-mediated stress response in *Mycobacterium smegmatis*: pGpp synthesis and its regulation. *J. Bacteriol.* 202:e00444-19. doi: 10.1128/JB.00444-19

Plocinski, P., Macios, M., Houghton, J., Niemiec, E., Plocinska, R., Brzostek, A., et al. (2019). Proteomic and transcriptomic experiments reveal an essential role of RNA degradosome complexes in shaping the transcriptome of *Mycobacterium tuberculosis*. *Nucleic Acids Res.* 47, 5892–5905. doi: 10.1093/nar/gkz251

Presnyak, V., Alhusaini, N., Chen, Y. H., Martin, S., Morris, N., Kline, N., et al. (2015). Codon optimality is a major determinant of mRNA stability. *Cell* 160, 1111–1124. doi: 10.1016/j.cell.2015.02.029

Py, B., Causton, H., Mudd, E. A., and Higgins, C. F. (1994). A protein complex mediating mRNA degradation in *Escherichia coli*. *Mol. Microbiol.* 14, 717–729. doi: 10.1111/j.1365-2958.1994.tb01309.x

Py, B., Higgins, C. F., Krisch, H. M., and Carpousis, A. J. (1996). A DEAD-box RNA helicase in the *Escherichia coli* RNA degradosome. *Nature* 381, 169–172. doi: 10.1038/381169a0

Quereda, J. J., Andersson, C., Cossart, P., Johansson, J., and Pizarro-Cerda, J. (2018). Role in virulence of phospholipases, listeriolysin O and listeriolysin S from epidemic *Listeria monocytogenes* using the chicken embryo infection model. *Vet. Res.* 49:13. doi: 10.1186/s13567-017-0496-4

Ramirez-Pena, E., Trevino, J., Liu, Z., Perez, N., and Sumby, P. (2010). The group A *Streptococcus* small regulatory RNA FasX enhances streptokinase activity by increasing the stability of the ska mRNA transcript. *Mol. Microbiol.* 78, 1332–1347. doi: 10.1111/j.1365-2958.2010.07427.x

Rao, S. P., Alonso, S., Rand, L., Dick, T., and Pethe, K. (2008). The protonmotive force is required for maintaining ATP homeostasis and viability of hypoxic, nonreplicating *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U.S.A.* 105, 11945–11950. doi: 10.1073/pnas.0711697105

Redko, Y., Galtier, E., Arnion, H., Darfeuille, F., Sismeiro, O., Coppee, J. Y., et al. (2016). RNase J depletion leads to massive changes in mRNA abundance in *Helicobacter pylori*. *RNA Biol.* 13, 243–253. doi: 10.1080/15476286.2015.1132141

Redon, E., Loubiere, P., and Cocaign-Bousquet, M. (2005a). Transcriptome analysis of the progressive adaptation of *Lactococcus lactis* to carbon starvation. *J. Bacteriol.* 187, 3589–3592. doi: 10.1128/JB.187.10.3589-3592.2005

Redon, E., Loubiere, P., and Cocaign-Bousquet, M. (2005b). Role of mRNA stability during genome-wide adaptation of *Lactococcus lactis* to carbon starvation. *J. Biol. Chem.* 280, 36380–36385. doi: 10.1074/jbc.M506006200

Regnier, P., and Hajnsdorf, E. (2013). 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, 602–609. doi: 10.4161/rna.23664

Richards, J., and Belasco, J. G. (2016). Distinct requirements for 5′-monophosphate-assisted RNA cleavage by *Escherichia coli* RNase E and RNase G. *J. Biol. Chem.* 291, 20825. doi: 10.1074/jbc.A115.702555

Richards, J., and Belasco, J. G. (2019). Obstacles to scanning by RNase E govern bacterial mRNA lifetimes by hindering access to distal cleavage sites. *Mol. Cell* 74, 284–295.e5. doi: 10.1016/j.molcel.2019.01.044

Richards, J., Liu, Q., Pellegrini, O., Celesnik, H., Yao, S., Bechhofer, D. H., et al. (2011). An RNA pyrophosphohydrolase triggers 5′-exonucleolytic degradation of mRNA in *Bacillus subtilis*. *Mol. Cell* 43, 940–949. doi: 10.1016/j.molcel.2011.07.023

Riglar, D. T., Giessen, T. W., Baym, M., Kerns, S. J., Niederhuber, M. J., Bronson, R. T., et al. (2017). Engineered bacteria can function in the mammalian gut long-term as live diagnostics of inflammation. *Nat. Biotechnol.* 35, 653–658. doi: 10.1038/nbt.3879

Rittershaus, E. S., Baek, S. H., and Sassetti, C. M. (2013). The normalcy of dormancy: common themes in microbial quiescence. *Cell Host Microbe* 13, 643–651. doi: 10.1016/j.chom.2013.05.012

Roberts, C., Anderson, K. L., Murphy, E., Projan, S. J., Mounts, W., Hurlburt, B., et al. (2006). Characterizing the effect of the *Staphylococcus aureus* virulence factor regulator, SarA, on log-phase mRNA half-lives. *J. Bacteriol.* 188, 2593–2603. doi: 10.1128/JB.188.7.2593-2603.2006

Rochat, T., Bouloc, P., Yang, Q., Bossi, L., and Figueroa-Bossi, N. (2012). Lack of interchangeability of Hfq-like proteins. *Biochimie* 94, 1554–1559. doi: 10.1016/j.biochi.2012.01.016

Rochat, T., Delumeau, O., Figueroa-Bossi, N., Noirot, P., Bossi, L., Dervyn, E., et al. (2015). Tracking the elusive function of *Bacillus subtilis* Hfq. *PLoS One* 10:e0124977. doi: 10.1371/journal.pone.0124977

Romeo, T., and Babitzke, P. (2018). Global regulation by CsrA and its RNA antagonists. *Microbiol. Spectr.* 6:10.1128/microbiolspec.RWR-0009-2017.

Ross, J. A., Thorsing, M., Lillebaek, E. M. S., Teixeira Dos Santos, P., and Kallipolitis, B. H. (2019). The LhrC sRNAs control expression of T cell-stimulating antigen TcsA in *Listeria monocytogenes* by decreasing tcsA mRNA stability. *RNA Biol.* 16, 270–281. doi: 10.1080/15476286.2019.1572423

Rozenski, J., Crain, P. F., and McCloskey, J. A. (1999). The RNA modification database: 1999 update. *Nucleic Acids Res.* 27, 196–197. doi: 10.1093/nar/27.1.196

Rustad, T. R., Minch, K. J., Brabant, W., Winkler, J. K., Reiss, D. J., Baliga, N. S., et al. (2013). Global analysis of mRNA stability in *Mycobacterium tuberculosis*. *Nucleic Acids Res.* 41, 509–517. doi: 10.1093/nar/gks1019

Sala, C., Forti, F., Magnoni, F., and Ghisotti, D. (2008). The katG mRNA of *Mycobacterium tuberculosis* and *Mycobacterium smegmatis* is processed at its 5′ end and is stabilized by both a polypurine sequence and translation initiation. *BMC Mol. Biol.* 9:33. doi: 10.1186/1471-2199-9-33

Sandler, P., and Weisblum, B. (1989). Erythromycin-induced ribosome stall in the ermA leader: a barricade to 5′-to-3′ nucleolytic cleavage of the ermA transcript. *J. Bacteriol.* 171, 6680–6688. doi: 10.1128/jb.171.12.6680-6688.1989

Santiago-Frangos, A., and Woodson, S. A. (2018). Hfq chaperone brings speed dating to bacterial sRNA. *Wiley Interdiscip. Rev. RNA* 9:e1475. doi: 10.1002/wrna.1475

Schneider, E., Blundell, M., and Kennell, D. (1978). Translation and mRNA decay. *Mol. Gen. Genet.* 160, 121–129.

Schubert, O. T., Ludwig, C., Kogadeeva, M., Zimmermann, M., Rosenberger, G., Gengenbacher, M., et al. (2015). Absolute proteome composition and dynamics during dormancy and resuscitation of *Mycobacterium tuberculosis*. *Cell Host Microbe* 18, 96–108. doi: 10.1016/j.chom.2015.06.001

Schumacher, M. A., Pearson, R. F., Moller, T., Valentin-Hansen, P., and Brennan, R. G. (2002). Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: a bacterial Sm-like protein. *EMBO J.* 21, 3546–3556. doi: 10.1093/emboj/cdf322

Schumann, U., Zhang, H. N., Sibbritt, T., Pan, A., Horvath, A., Gross, S., et al. (2020). Multiple links between 5-methylcytosine content of mRNA and translation. *BMC Biol.* 18:40. doi: 10.1186/s12915-020-00769-5

Schwartz, S., Mumbach, M. R., Jovanovic, M., Wang, T., Maciag, K., Bushkin, G. G., et al. (2014). Perturbation of m6A writers reveals two distinct classes of mRNA methylation at internal and 5′ sites. *Cell Rep.* 8, 284–296. doi: 10.1016/j.celrep.2014.05.048

Segev, E., Smith, Y., and Ben-Yehuda, S. (2012). RNA dynamics in aging bacterial spores. *Cell* 148, 139–149. doi: 10.1016/j.cell.2011.11.059

Selinger, D. W., Saxena, R. M., Cheung, K. J., Church, G. M., and Rosenow, C. (2003). Global RNA half-life analysis in *Escherichia coli* reveals positional patterns of transcript degradation. *Genome Res.* 13, 216–223. doi: 10.1101/gr.912603

Seyfzadeh, M., Keener, J., and Nomura, M. (1993). spoT-dependent accumulation of guanosine tetraphosphate in response to fatty acid starvation in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 90, 11004–11008. doi: 10.1073/pnas.90.23.11004

Shahbabian, K., Jamalli, A., Zig, L., and Putzer, H. (2009). RNase Y, a novel endoribonuclease, initiates riboswitch turnover in *Bacillus subtilis*. *EMBO J.* 28, 3523–3533. doi: 10.1038/emboj.2009.283

Shalem, O., Dahan, O., Levo, M., Martinez, M. R., Furman, I., Segal, E., et al. (2008). Transient transcriptional responses to stress are generated by opposing effects of mRNA production and degradation. *Mol. Syst. Biol.* 4:223. doi: 10.1038/msb.2008.59

Sharp, J. S., and Bechhofer, D. H. (2005). Effect of 5′-proximal elements on decay of a model mRNA in *Bacillus subtilis*. *Mol. Microbiol.* 57, 484–495. doi: 10.1111/j.1365-2958.2005.04683.x

Sherman, D. R., Voskuil, M., Schnappinger, D., Liao, R., Harrell, M. I., and Schoolnik, G. K. (2001). Regulation of the *Mycobacterium tuberculosis* hypoxic response gene encoding alpha -crystallin. *Proc. Natl. Acad. Sci. U.S.A.* 98, 7534–7539. doi: 10.1073/pnas.121172498

Shivakumar, A. G., Hahn, J., Grandi, G., Kozlov, Y., and Dubnau, D. (1980). Posttranscriptional regulation of an erythromycin resistance protein specified by plasmic pE194. *Proc. Natl. Acad. Sci. U.S.A.* 77, 3903–3907. doi: 10.1073/pnas.77.7.3903

Siculella, L., Damiano, F., di Summa, R., Tredici, S. M., Alduina, R., Gnoni, G. V., et al. (2010). Guanosine 5′-diphosphate 3′-diphosphate (ppGpp) as a negative modulator of polynucleotide phosphorylase activity in a 'rare' actinomycete. *Mol. Microbiol.* 77, 716–729. doi: 10.1111/j.1365-2958.2010.07240.x

Sievers, S., Lund, A., Menendez-Gil, P., Nielsen, A., Storm Mollerup, M., Lambert Nielsen, S., et al. (2015). The multicopy sRNA LhrC controls expression of the oligopeptide-binding protein OppA in *Listeria monocytogenes*. *RNA Biol.* 12, 985–997. doi: 10.1080/15476286.2015.1071011

Silva, A. J., Sultan, S. Z., Liang, W., and Benitez, J. A. (2008). Role of the histone-like nucleoid structuring protein in the regulation of rpoS and RpoS-dependent genes in *Vibrio cholerae*. *J. Bacteriol.* 190, 7335–7345. doi: 10.1128/JB.00360-08

Sim, M., Lim, B., Sim, S. H., Kim, D., Jung, E., Lee, Y., et al. (2014). Two tandem RNase III cleavage sites determine betT mRNA stability in response to osmotic stress in *Escherichia coli*. *PLoS One* 9:e100520. doi: 10.1371/journal.pone.0100520

Sinha, D., Matz, L. M., Cameron, T. A., and De Lay, N. R. (2018). Poly(A) polymerase is required for RyhB sRNA stability and function in *Escherichia coli*. *RNA* 24, 1496–1511. doi: 10.1261/rna.067181.118

Sittka, A., Lucchini, S., Papenfort, K., Sharma, C. M., Rolle, K., Binnewies, T. T., 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

Smeulders, M. J., Keer, J., Speight, R. A., and Williams, H. D. (1999). Adaptation of *Mycobacterium smegmatis* to stationary phase. *J. Bacteriol.* 181, 270–283. doi: 10.1128/jb.181.1.270-283.1999

Sonnleitner, E., Schuster, M., Sorger-Domenigg, T., Greenberg, E. P., and Blasi, U. (2006). Hfq-dependent alterations of the transcriptome profile and effects on quorum sensing in *Pseudomonas aeruginosa*. *Mol. Microbiol.* 59, 1542–1558. doi: 10.1111/j.1365-2958.2006.05032.x

Sousa, S., Marchand, I., and Dreyfus, M. (2001). Autoregulation allows *Escherichia coli* RNase E to adjust continuously its synthesis to that of its substrates. *Mol. Microbiol.* 42, 867–878. doi: 10.1046/j.1365-2958.2001.02687.x

Steglich, C., Lindell, D., Futschik, M., Rector, T., Steen, R., and Chisholm, S. W. (2010). Short RNA half-lives in the slow-growing marine cyanobacterium *Prochlorococcus*. *Genome Biol.* 11:R54. doi: 10.1186/gb-2010-11-5-r54

Strahl, H., Turlan, C., Khalid, S., Bond, P. J., Kebalo, J. M., Peyron, P., et al. (2015). Membrane recognition and dynamics of the RNA degradosome. *PLoS Genet.* 11:e1004961. doi: 10.1371/journal.pgen.1004961

Sun, X., Zhulin, I., and Wartell, R. M. (2002). Predicted structure and phyletic distribution of the RNA-binding protein Hfq. *Nucleic Acids Res.* 30, 3662–3671. doi: 10.1093/nar/gkf508

Tao, F., Liu, Y., Luo, Q., Su, F., Xu, Y., Li, F., et al. (2011). Novel organic solvent-responsive expression vectors for biocatalysis: application for development of an organic solvent-tolerant biodesulfurizing strain. *Bioresour. Technol.* 102, 9380–9387. doi: 10.1016/j.biortech.2011.08.015

Tejada-Arranz, A., de Crecy-Lagard, V., and de Reuse, H. (2020). Bacterial RNA Degradosomes: molecular machines under tight control. *Trends Biochem. Sci.* 45, 42–57. doi: 10.1016/j.tibs.2019.10.002

Thorne, S. H., and Williams, H. D. (1997). Adaptation to nutrient starvation in *Rhizobium leguminosarum* bv. phaseoli: analysis of survival, stress resistance, and changes in macromolecular synthesis during entry to and exit from stationary phase. *J. Bacteriol.* 179, 6894–6901. doi: 10.1128/jb.179.22.6894-6901.1997

Timmermans, J., and Van Melderen, L. (2010). Post-transcriptional global regulation by CsrA in bacteria. *Cell. Mol. Life Sci.* 67, 2897–2908. doi: 10.1007/s00018-010-0381-z

Toledo-Arana, A., Dussurget, O., Nikitas, G., Sesto, N., Guet-Revillet, H., Balestrino, D., et al. (2009). The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature* 459, 950–956. doi: 10.1038/nature08080

Tortosa, P., Albano, M., and Dubnau, D. (2000). Characterization of ylbF, a new gene involved in competence development and sporulation in *Bacillus subtilis*. *Mol. Microbiol.* 35, 1110–1119. doi: 10.1046/j.1365-2958.2000.01779.x

Udekwu, K. I., Darfeuille, F., Vogel, J., Reimegard, J., Holmqvist, E., and Wagner, E. G. (2005). Hfq-dependent regulation of OmpA synthesis is mediated by an antisense RNA. *Genes Dev.* 19, 2355–2366. doi: 10.1101/gad.354405

Unciuleac, M. C., and Shuman, S. (2013). Distinctive effects of domain deletions on the manganese-dependent DNA polymerase and DNA phosphorylase activities of *Mycobacterium smegmatis* polynucleotide phosphorylase. *Biochemistry* 52, 2967–2981. doi: 10.1021/bi400281w

Unniraman, S., Chatterji, M., and Nagaraja, V. (2002). A hairpin near the 5′ end stabilises the DNA gyrase mRNA in *Mycobacterium smegmatis*. *Nucleic Acids Res.* 30, 5376–5381. doi: 10.1093/nar/gkf697

Updegrove, T. B., Zhang, A., and Storz, G. (2016). Hfq: the flexible RNA matchmaker. *Curr. Opin. Microbiol.* 30, 133–138. doi: 10.1016/j.mib.2016.02.003

Vanderpool, C. K., and Gottesman, S. (2004). Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of

the glucose phosphoenolpyruvate phosphotransferase system. *Mol. Microbiol.* 54, 1076–1089. doi: 10.1111/j.1365-2958.2004.04348.x

Vanzo, N. F., Li, Y. S., Py, B., Blum, E., Higgins, C. F., Raynal, L. C., et al. (1998). Ribonuclease E organizes the protein interactions in the *Escherichia coli* RNA degradosome. *Genes Dev.* 12, 2770–2781. doi: 10.1101/gad.12.17.2770

Vargas-Blanco, D. A., Zhou, Y., Zamalloa, L. G., Antonelli, T., and Shell, S. S. (2019). mRNA degradation rates are coupled to metabolic status in *Mycobacterium smegmatis*. *mBio* 10:e00957-19. doi: 10.1128/mBio.00957-19

Varmus, H. E., Perlman, R. L., and Pastan, I. (1971). Regulation of lac transcription in antibiotic-treated *E. coli*. *Nat. New Biol.* 230, 41–44. doi: 10.1038/newbio230041a0

Vecerek, B., Rajkowitsch, L., Sonnleitner, E., Schroeder, R., and Blasi, U. (2008). The C-terminal domain of *Escherichia coli* Hfq is required for regulation. *Nucleic Acids Res.* 36, 133–143. doi: 10.1093/nar/gkm985

Ventola, C. L. (2015a). The antibiotic resistance crisis: part 1: causes and threats. *P T* 40, 277–283.

Ventola, C. L. (2015b). The antibiotic resistance crisis: part 2: management strategies and new agents. *P T* 40, 344–352.

Vogel, J., and Luisi, B. F. (2011). Hfq and its constellation of RNA. *Nat. Rev. Microbiol.* 9, 578–589. doi: 10.1038/nrmicro2615

Vvedenskaya, I. O., Bird, J. G., Zhang, Y., Zhang, Y., Jiao, X., Barvik, I., et al. (2018). CapZyme-seq comprehensively defines promoter-sequence determinants for rna 5′ capping with NAD. *Mol. Cell* 70, 553–564.e9. doi: 10.1016/j.molcel.2018.03.014

Vvedenskaya, I. O., Sharp, J. S., Goldman, S. R., Kanabar, P. N., Livny, J., Dove, S. L., et al. (2012). Growth phase-dependent control of transcription start site selection and gene expression by nanoRNAs. *Genes Dev.* 26, 1498–1507. doi: 10.1101/gad.192732.112

Vytvytska, O., Moll, I., Kaberdin, V. R., von Gabain, A., and Blasi, U. (2000). Hfq (HF1) stimulates ompA mRNA decay by interfering with ribosome binding. *Genes Dev.* 14, 1109–1118.

Wagner, L. A., Gesteland, R. F., Dayhuff, T. J., and Weiss, R. B. (1994). An efficient Shine-Dalgarno sequence but not translation is necessary for lacZ mRNA stability in *Escherichia coli*. *J. Bacteriol.* 176, 1683–1688. doi: 10.1128/jb.176.6.1683-1688.1994

Wang, H., Ayala, J. C., Benitez, J. A., and Silva, A. J. (2012). Interaction of the histone-like nucleoid structuring protein and the general stress response regulator RpoS at *Vibrio cholerae* promoters that regulate motility and hemagglutinin/protease expression. *J. Bacteriol.* 194, 1205–1215. doi: 10.1128/JB.05900-11

Wang, X., Dubey, A. K., Suzuki, K., Baker, C. S., Babitzke, P., and Romeo, T. (2005). CsrA post-transcriptionally represses pgaABCD, responsible for synthesis of a biofilm polysaccharide adhesin of *Escherichia coli*. *Mol. Microbiol.* 56, 1648–1663. doi: 10.1111/j.1365-2958.2005.04648.x

Wang, Y., Li, Y., Toth, J. I., Petroski, M. D., Zhang, Z., and Zhao, J. C. (2014). N6-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. *Nat. Cell Biol.* 16, 191–198. doi: 10.1038/ncb2902

Wassarman, K. M., Repoila, F., Rosenow, C., Storz, G., and Gottesman, S. (2001). Identification of novel small RNAs using comparative genomics and microarrays. *Genes Dev.* 15, 1637–1651. doi: 10.1101/gad.901001

Wei, B. L., Brun-Zinkernagel, A. M., Simecka, J. W., Pruss, B. M., Babitzke, P., and Romeo, T. (2001). Positive regulation of motility and flhDC expression by the RNA-binding protein CsrA of *Escherichia coli*. *Mol. Microbiol.* 40, 245–256. doi: 10.1046/j.1365-2958.2001.02380.x

Wei, C. M., Gershowitz, A., and Moss, B. (1975). Methylated nucleotides block 5′ terminus of HeLa cell messenger RNA. *Cell* 4, 379–386. doi: 10.1016/0092-8674(75)90158-0

Wolfe, A. D., and Hahn, F. E. (1965). Mode of action of chloramphenicol. Ix. Effects of chloramphenicol upon a ribosomal amino acid polymerization system and its binding to bacterial ribosome. *Biochim. Biophys. Acta* 95, 146–155. doi: 10.1016/0005-2787(65)90219-4

Wood, D. N., Chaussee, M. A., Chaussee, M. S., and Buttaro, B. A. (2005). Persistence of *Streptococcus pyogenes* in stationary-phase cultures. *J. Bacteriol.* 187, 3319–3328. doi: 10.1128/JB.187.10.3319-3328.2005

World Health Organization [WHO] (2019). *No Time to Wait: Securing the Future from Drug-Resistant Infections*. Available online at: https://www.who.int/antimicrobial-resistance/interagency-coordination-group/final-report/en/ (accessed April 17, 2020).

Xu, C., Huang, R., Teng, L., Jing, X., Hu, J., Cui, G., et al. (2015). Cellulosome stoichiometry in *Clostridium cellulolyticum* is regulated by selective RNA processing and stabilization. *Nat. Commun.* 6:6900. doi: 10.1038/ncomms7900

Xu, W., Huang, J., and Cohen, S. N. (2008). Autoregulation of AbsB (RNase III) expression in *Streptomyces coelicolor* by endoribonucleolytic cleavage of absB operon transcripts. *J. Bacteriol.* 190, 5526–5530. doi: 10.1128/JB.00558-08

Yang, Y., Wang, L., Han, X., Yang, W. L., Zhang, M., Ma, H. L., et al. (2019). RNA 5-methylcytosine facilitates the maternal-to-zygotic transition by preventing maternal mRNA Decay. *Mol. Cell* 75, 1188–1202.e11. doi: 10.1016/j.molcel.2019.06.033

Yarmolinsky, M. B., and Haba, G. L. (1959). Inhibition by puromycin of amino acid incorporation into protein. *Proc. Natl. Acad. Sci. U.S.A.* 45, 1721–1729. doi: 10.1073/pnas.45.12.1721

Yu, X., Li, B., Jang, G. J., Jiang, S., Jiang, D., Jang, J. C., et al. (2019). Orchestration of Processing body dynamics and mrna decay in *Arabidopsis* immunity. *Cell Rep.* 28, 2194–2205.e6. doi: 10.1016/j.celrep.2019.07.054

Yue, Y., Liu, J., Cui, X., Cao, J., Luo, G., Zhang, Z., et al. (2018). VIRMA mediates preferential m(6)A mRNA methylation in 3′UTR and near stop codon and associates with alternative polyadenylation. *Cell Discov.* 4:10. doi: 10.1038/s41421-018-0019-0

Zeller, M. E., Csanadi, A., Miczak, A., Rose, T., Bizebard, T., and Kaberdin, V. R. (2007). Quaternary structure and biochemical properties of mycobacterial RNase E/G. *Biochem. J.* 403, 207–215. doi: 10.1042/BJ20061530

Zgurskaya, H. I., Keyhan, M., and Matin, A. (1997). The sigma S level in starving *Escherichia coli* cells increases solely as a result of its increased stability, despite decreased synthesis. *Mol. Microbiol.* 24, 643–651. doi: 10.1046/j.1365-2958.1997.3961742.x

Zhang, A., Wassarman, K. M., Rosenow, C., Tjaden, B. C., Storz, G., and Gottesman, S. (2003). Global analysis of small RNA and mRNA targets of Hfq. *Mol. Microbiol.* 50, 1111–1124. doi: 10.1046/j.1365-2958.2003.03734.x

Zhang, D., Liu, Y., Wang, Q., Guan, Z., Wang, J., Liu, J., et al. (2016). Structural basis of prokaryotic NAD-RNA decapping by NudC. *Cell Res.* 26, 1062–1066. doi: 10.1038/cr.2016.98

Zhang, H., Li, D., Zhao, L., Fleming, J., Lin, N., Wang, T., et al. (2013). Genome sequencing of 161 *Mycobacterium tuberculosis* isolates from China identifies genes and intergenic regions associated with drug resistance. *Nat. Genet.* 45, 1255–1260. doi: 10.1038/ng.2735

Zhao, B. S., Wang, X., Beadell, A. V., Lu, Z., Shi, H., Kuuspalu, A., et al. (2017). m(6)A-dependent maternal mRNA clearance facilitates zebrafish maternal-to-zygotic transition. *Nature* 542, 475–478. doi: 10.1038/nature21355

Zhao, J. P., Zhu, H., Guo, X. P., and Sun, Y. C. (2018). AU-rich long 3′ untranslated region regulates gene expression in bacteria. *Front. Microbiol.* 9:3080. doi: 10.3389/fmicb.2018.03080

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 14 October 2020
doi: 10.3389/fmicb.2020.570536

Differential Regulation of CsrC and CsrB by CRP-cAMP in *Salmonella enterica*

Youssef El Mouali†, Guillem Esteva-Martínez, David García-Pedemonte and Carlos Balsalobre*

Department of Genetics, Microbiology and Statistics, School of Biology, Universitat de Barcelona, Barcelona, Spain

Post-transcriptional regulation mediated by regulatory small RNAs (sRNAs) has risen as a key player in fine-tuning gene expression in response to environmental stimuli. Here, we show that, in *Salmonella enterica*, the central metabolic regulator CRP-cAMP differentially regulates the sRNAs CsrB and CsrC in a growth phase-dependent manner. While CsrB expression remains unchanged during growth, CsrC displays a growth phase-dependent expression profile, being weakly expressed at the logarithmic growth phase and induced upon entry into stationary phase. We show that CRP-cAMP contributes to the expression pattern of CsrC by repressing its expression during the logarithmic growth phase. The CRP-cAMP mediated repression of CsrC is independent of SirA, a known transcriptional CsrB/CsrC activator. We further show that the sRNA Spot 42, which is derepressed in a Δ*crp* strain, upregulates CsrC during logarithmic growth. We propose a model where the growth-dependent regulation of CsrC is sustained by the CRP-cAMP-mediated repression of Spot 42. Together, our data point toward a differential regulation of the sRNAs CsrB and CsrC in response to environmental stimuli, leading to fine-tuning of gene expression via the sequestration of the RNA-binding protein CsrA.

Keywords: CsrC, post-transcriptional regulation, sRNA, CRP-cAMP, growth phase regulation, Spot 42

OPEN ACCESS

Edited by:
Omar Orellana, University of Chile, Chile

Reviewed by:
Teppei Morita, Keio University, Japan
Iván Calderón, Andres Bello University, Chile

****Correspondence:***
Youssef El Mouali
youssef.elmoualibenomar@helmholtz-hiri.de

†Present address:
Youssef El Mouali, Helmholtz Institute for RNA-based Infection Research (HIRI), Würzburg, Germany

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *08 June 2020*
Accepted: *17 September 2020*
Published: *14 October 2020*

Citation:
El Mouali Y, Esteva-Martínez G, García-Pedemonte D and Balsalobre C (2020) Differential Regulation of CsrC and CsrB by CRP-cAMP in Salmonella enterica. Front. Microbiol. 11:570536. doi: 10.3389/fmicb.2020.570536

INTRODUCTION

Bacteria need to adapt rapidly to changing environmental conditions, which is particularly crucial for pathogenic bacteria during the process of infection. While transcription plays a major role in the regulation of gene expression, bacteria display a plethora of post-transcriptional mechanisms that allow the fine-tuning of gene expression in response to environmental cues. *Salmonella enterica* has been investigated extensively with respect to gene regulation and has become a model for the study of post-transcriptional RNA-mediated regulation.

Among the described post-transcriptional regulation mechanisms, a prominent role has been attributed to sRNAs that modulate gene expression primarily via binding to target mRNAs (Hör et al., 2020). However, sRNAs can also regulate gene expression at the post-translational level, such as the two sRNAs CsrB and CsrC. These sRNAs contain several stem loop structures with GGA motifs in their loop regions, enabling them to bind CsrA. CsrA is a widely conserved RNA-binding protein that inhibits translation by binding to GGA motifs around the ribosome-binding site of target mRNAs (Romeo and Babitzke, 2018). Binding of CsrB and CsrC to CsrA titrates CsrA away

from its target mRNAs, thereby counteracting its inhibitory activity (Liu et al., 1997; Weilbacher et al., 2003).

CRP is a transcription factor that acts as a metabolic sensor and becomes active upon binding to the intracellular second messenger cAMP (cyclic adenosine monophosphate) (Görke and Stülke, 2008). Interaction of CRP-cAMP with DNA leads to activation or repression of its target genes. In *E. coli*, it has been shown that CRP-cAMP represses the expression of both CsrB and CsrC (Pannuri et al., 2016). While repression of CsrB occurs through an indirect mechanism, the repression of CsrC occurs by direct binding of CRP-cAMP to the promoter region of *csrC*, where it competes with the *csrC* activator UvrY (Pannuri et al., 2016). In *Salmonella*, the homolog of the two-component system BarA-UvrY is BarA-SirA (Johnston et al., 1996), which, as in *E. coli*, it also positively regulates the expression of CsrB and CsrC (Teplitski et al., 2003; Fortune et al., 2006; Martínez et al., 2011). Expression studies using *S. enterica* cultures on solid media indicates that CRP-cAMP, in contrast to its role described in *E. coli*, positively regulate the expression of CsrB and CsrC via upregulation of *sirA* (Teplitski et al., 2006). In *S. enterica* other sRNAs, such as CyaR and Spot 42, are also regulated by CRP-cAMP, acting as an activator for CyaR and as a repressor for Spot 42 (Papenfort et al., 2008; El Mouali et al., 2018).

In this study we investigated the role of CRP-cAMP in the expression of the sRNAs CsrB and CsrC in the model organism *S. enterica* using liquid cultures. We describe that CsrC, but not CsrB, display a growth-dependent expression pattern. CsrC expression is silenced during logarithmic growth and highly expressed upon entry into stationary phase, while CsrB expression seems to be constitutive through the growth curve. CRP-cAMP plays a relevant role in this growth-dependent regulatory network. CRP-cAMP differentially regulates the levels of CsrB and CsrC as it does not affect CsrB expression but represses CsrC expression during logarithmic growth. Remarkably, while SirA is required for full expression of CsrC, the CRP-cAMP-mediated repression during logarithmic growth seems to be independent of SirA-mediated regulation. Our data further indicates that Spot 42 contributes to the CRP-cAMP-induced repression of CsrC, suggesting that CRP-cAMP and Spot 42 converge into the differential regulation of CsrB and CsrC in *Salmonella*.

RESULTS

CRP-cAMP Represses CsrC Levels at the Logarithmic Growth Phase

In order to study the CRP-cAMP mediated regulation of CsrB and CsrC in *Salmonella*, the expression of both sRNAs was monitored using transcriptional reporter fusions. The regulatory elements controlling CsrB and CsrC expression in *Salmonella* have been identified *in silico* and characterized experimentally (Martínez et al., 2014). Accordingly, the described regulatory regions of CsrB (−403, +18) and CsrC (−347, +60) were cloned as *lacZ* transcriptional fusions in the pQF50 vector (Farinha and Kropinski, 1989), allowing to monitor the levels of CsrB and CsrC during growth. The contribution of CRP-cAMP was assessed by determining the expression level of CsrB and CsrC in wild-type (WT) and in a Δ*crp* mutant strain, lacking the transcriptional factor CRP. In rich media, no differences in growth rate could be observed in WT and Δ*crp* strains carrying either *csrB-lacZ* or *csrC-lacZ* (**Figure 1A**). Transcriptional expression was monitored during logarithmic growth ($OD_{600\,nm}$ 0.4) and upon entry into stationary phase ($OD_{600\,nm}$ 2.0). In the WT, CsrB expression was apparently identical in the two growth phases (**Figure 1B**). By contrast, CsrC displayed a growth-dependent expression pattern, being less expressed during logarithmic growth and being induced after entering stationary phase (**Figure 1B**). Remarkably, CRP-cAMP contributes to the CsrC growth-dependent expression pattern. In Δ*crp*, the expression of CsrC is induced when compared to WT at the logarithmic phase and to a lesser extent at the stationary phase, indicating that CRP-cAMP represses CsrC expression (**Figure 1B**). By contrast, no effect on CsrB regulation by CRP-cAMP was observed. These results were unexpected since it was previously described that CRP-cAMP acts as an activator of both CsrB and CsrC expression when *Salmonella* is grown on LB agar media (Teplitski et al., 2006). To discern if the discrepancy could be consequence of differences in the genetic constructs used to monitor gene expression, a similar experiment as in Teplitski et al. (2006) was performed using our strains. The transcriptional expression was monitored after growth on LB agar media. Consistently with the previous report, the transcriptional expression of CsrB and CsrC was strongly diminished in a Δ*crp* derivative strain compared to WT when *Salmonella* cells were grown on LB agar media as noted by the white colony phenotype of Δ*crp* when compared to the blue colony phenotype of WT (**Supplementary Figure S1**; Teplitski et al., 2006). Altogether, CRP-cAMP seem to be required for activation of the expression of CsrB and CsrC in solid media, while it acts as a repressor of CsrC particularly at logarithmic growth phase. This indicates that CRP-cAMP-mediated regulation of CsrB and CsrC is dependent on the growth conditions.

The differential regulation of CsrC and CsrB by CRP-cAMP at the logarithmic growth phase was further corroborated by the direct RNA detection of CsrB and CsrC. In the WT, CsrC transcript was barely detected, indicating that CsrC expression is tightly silenced. In the Δ*crp* strain, high levels of CsrC were detected indicating that CRP-cAMP is involved in the CsrC silencing during logarithmic growth (**Figure 1C**). Contrasting the CRP-cAMP-dependent changes in CsrC levels, CsrB was not influenced by knockout of Δ*crp* (**Figure 1C**), thereby agreeing with the results of our transcriptional fusions (**Figure 1B**). In addition, deletion of Δ*csrB* cause a mild upregulation of CsrC, presumably by affecting the positive feed forward loop that free CsrA protein exerts on its repressors CsrB and CsrC (Romeo and Babitzke, 2018).

CRP becomes active upon binding to cAMP, which is produced by the adenylate cyclase Cya. Therefore, absence of *crp* or *cya* should display similar expression profiles. Accordingly, the expression of CsrC in logarithmic growth phase is induced in the Δ*cya* derivative strain when compared to WT while no change was observed in the expression of CsrB (**Figure 1D**).

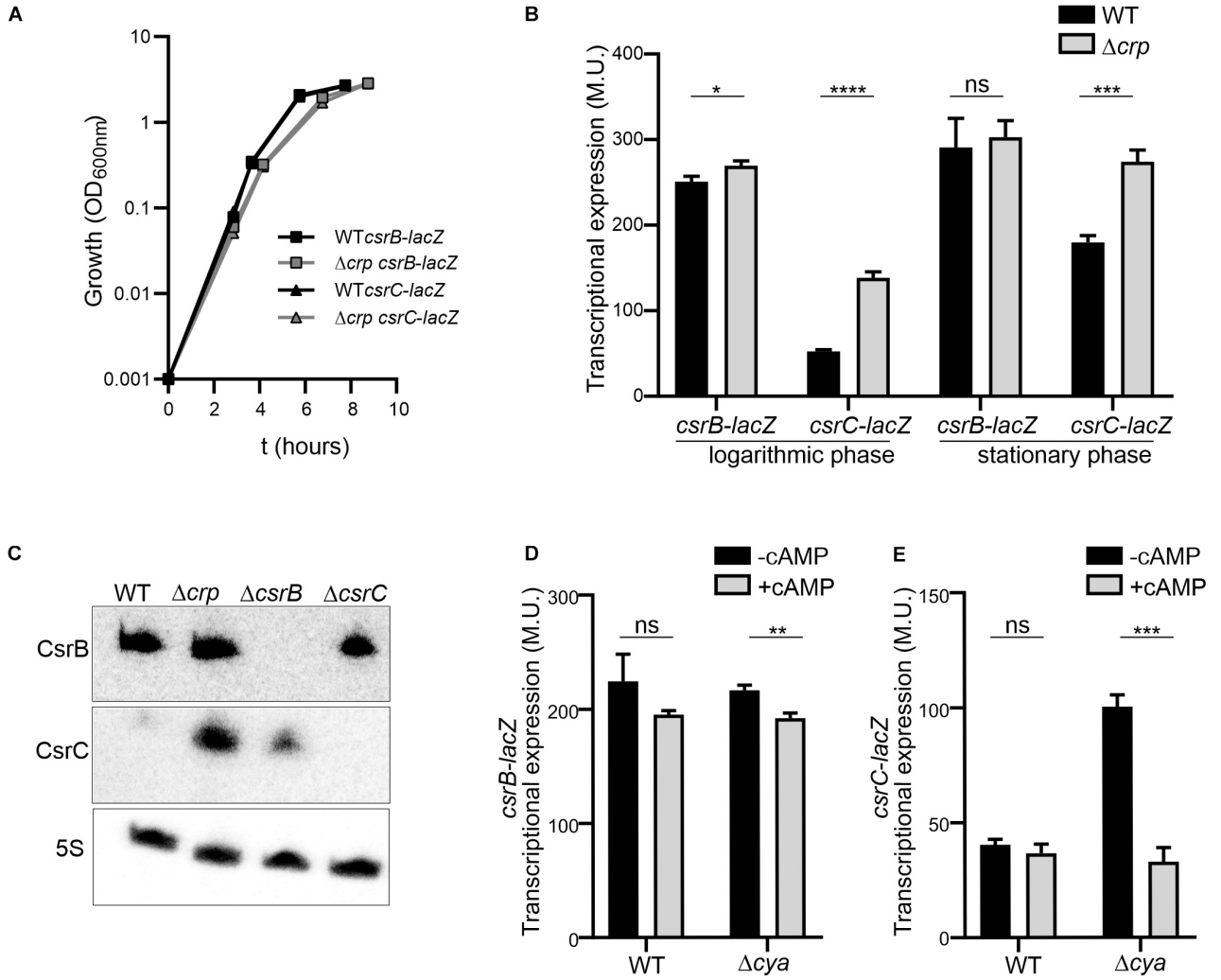


FIGURE 1 | The expression of *csrC* is repressed at the logarithmic growth phase in a CRP-cAMP mediated manner. **(A)** Growth curve in LB at 37°C of *Salmonella* SV5015 strain (WT) and its isogenic Δ*crp* mutant, carrying either a plasmid fusion of *csrB-lacZ* or *csrC-lacZ*. **(B)** Transcriptional expression of *csrB* and *csrC*. Samples from cultures in A were taken at either an $OD_{600\ nm}$ of 0.4 or an $OD_{600\ nm}$ of 2.0 for determination of the β-galactosidase activity in Miller units (M.U.). **(C)** Northern blot analyses of the *csrB* and *csrC* transcripts. Total RNA samples from cultures of WT (SV5015) and its derivatives Δ*crp*, Δ*csrB,* and Δ*csrC* grown in LB up to an $OD_{600\ nm}$ of 0.4 were analyzed. Detection of the 5S transcript was used as a control. Transcriptional expression of *csrB* **(D)** and *csrC* **(E)** in the *cya* mutant strain and chemical complementation. Cultures of WT, Δ*crp* and Δ*cya* derivatives were grown in LB either in the absence or in the presence of cAMP (5 mM) at 37°C up to logarithmic phase ($OD_{600\ nm}$ of 0.4) and samples were taken for β-galactosidase measurement. In B, C and D; β-galactosidase activity was determined for three independent cultures, average and standard deviation are presented. $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$; $^{****}p < 0.0001$; ns, not significant.

To further confirm the involvement of CRP-cAMP in CsrC regulation, chemical complementation of Δ*cya* was carried out by ectopic addition of cAMP. The addition of cAMP repressed the expression of CsrC in the Δ*cya* derivative strain while no effect was observed in CsrB expression (**Figure 1E**). Our data indicate that CRP-cAMP is involved in the growth-dependent regulation of CsrC by repressing its expression.

CRP-cAMP-Mediated Repression of CsrC via a SirA-Independent Pathway

To further characterize the CRP-cAMP mediated regulation of CsrC, a chromosomal *csrC-lacZ* fusion was generated. In agreement with earlier results, chromosomal *csrC-lacZ* has a growth dependent expression pattern, where it is lowly expressed at the logarithmic growth phase and induced upon entry into early stationary phase (**Figure 2A**). Interestingly, we further observed an eightfold induction of the chromosomal *csrC-lacZ* fusion in the Δ*crp* background when compared to the WT, which was only true during logarithmic growth (**Figure 2A**).

The BarA-SirA two-component system was described to positively regulate the expression of CsrC (Teplitski et al., 2006). The possible involvement of BarA-SirA in the regulation of CsrC by CRP-cAMP was assessed. SirA deletion leads to a decrease of the overall transcriptional expression of the chromosomal *csrC-lacZ* fusion, 28.2 ± 0.8 Miller units in Δ*sirA*

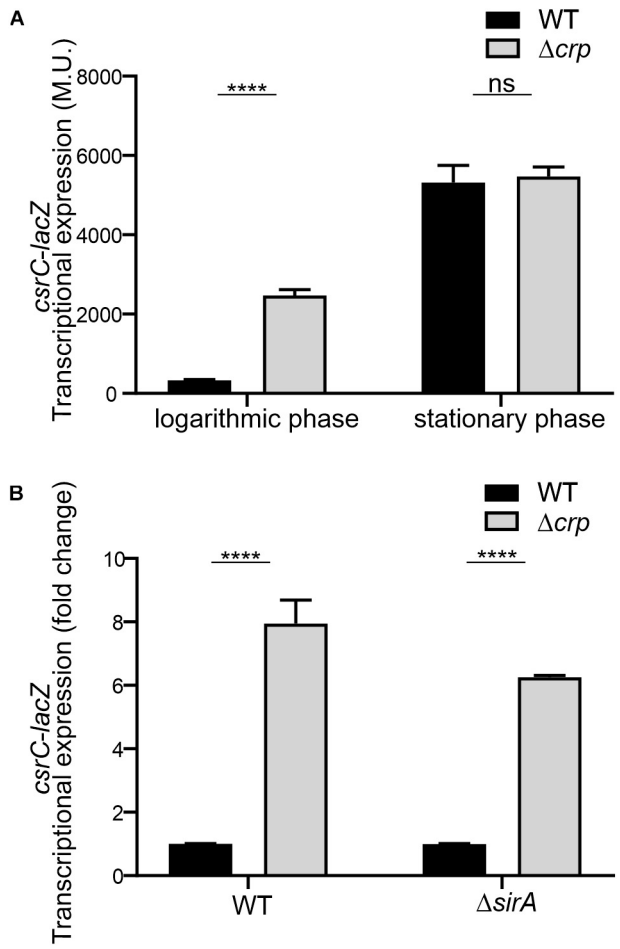


FIGURE 2 | CRP-cAMP mediated repression of *csrC* is independent of SirA. **(A)** Transcriptional expression of chromosomal *csrC-lacZ* fusion was monitored in WT and Δ*crp* mutant genetic backgrounds. Cultures were grown in LB at 37°C up to an $OD_{600\ nm}$ of either 0.4 or 2.0. **(B)** Transcriptional expression of the chromosomal *csrC-lacZ* fusion was monitored in WT, Δ*crp*, Δ*sirA,* and Δ*crp*Δ*sirA* genetic backgrounds. The transcriptional expression is shown in relative values. In each case, the reference (crp^+, WT) was set as 1.0. Miller units of the crp^+ backgrounds were 308 ± 19 and 28 ± 0.8 for WT and Δ*sirA,* respectively. β-galactosidase activity was determined for three independent cultures, average and standard deviation is presented. ****$p < 0.0001$; ns, not significant.

compared to 308.8 ± 19.3 Miller units in WT. Interestingly, despite the overall lower expression levels of CsrC in the Δ*sirA* background, the deletion of Δ*crp* led to a sixfold increase in the transcriptional expression of *csrC-lacZ* compared to WT (**Figure 2B**), indicating that CRP-cAMP modulates *csrC* expression via a SirA-independent mechanism.

In *E. coli*, it was proposed that CRP-cAMP represses CsrC expression through binding to the upstream region of *csrC* and competing with UvrY, the SirA homolog (Pannuri et al., 2016). In *Salmonella*, SirA induces CsrC expression by binding ∼160–168 bp upstream of the *csrC* promoter (**Figure 3A**; Martínez et al., 2014). Alignment of the upstream promoter regions of CsrC from *E. coli* and *Salmonella* showed that SirA binding site is conserved, whereas the CRP-cAMP binding site displays a lesser extent of conservation (**Supplementary Figure S2**). To characterize if CRP-cAMP binding to the *csrC* promoter is required for the described regulation, an additional *csrC-lacZ* fusion was generated where the putative binding sites of these transcriptional factors are not present. As shown in **Figure 3A**, the pQF50 cloned fragment $csrC_{347}$-*lacZ* (−347, +60) maintains SirA/CRP binding sites while $csrC_{91}$-*lacZ* (−91, +60) does not. Remarkably, in both $csrC_{347}$-*lacZ* and $csrC_{91}$-*lacZ* an induction of expression is observed in the Δ*crp* strain when compared to WT (**Figure 3B**). These results suggest that CRP-cAMP represses the expression of CsrC during logarithmic growth phase via a mechanism that is independent of the competition between SirA and CRP for the *csrC* promoter.

The sRNA Spot 42 Positively Regulates CsrC

In *Salmonella*, CRP-cAMP represses the expression of the regulatory sRNA Spot 42 in a growth-dependent manner (El Mouali et al., 2018). The similarity of the expression profiles of Spot 42 and CsrC, let us to hypothesize that CRP-cAMP might regulate CsrC via modulation of Spot 42 levels. To this end, we determined CsrC expression during logarithmic growth upon ectopic expression of Spot 42. Remarkably, CsrC transcript levels were strongly induced upon Spot 42 overexpression (pBRSpot 42) when compared to the vector control strain (pBRVC) (**Figure 4A**). In contrast, no accumulation of CsrB was detected upon overexpression of Spot 42 (**Figure 4A**). Further supporting our hypothesis, Spot 42 seemed to differentially regulate CsrB and CsrC expression (**Figure 4A**), similar to the described CRP-cAMP-mediated regulation (**Figure 1C**). To verify this observation, we overexpressed Spot 42 in the chromosomal *csrC-lacZ* fusion background, resulting in a threefold induction of CsrC when compared to the vector control strain (**Figure 4B**). The Spot 42-mediated regulation of CsrC may account for the described repression of CsrC by CRP-cAMP. The contribution of Spot 42 on the induction of CsrC in Δ*crp* was further assessed. The CsrC derepression in a Δ*crp* mutant strain drops in absence of Spot 42 (Δ*spf*) as compared to a Spot 42 proficient strain indicating that Spot 42 contributes to the CsrC derepression in a Δ*crp* background (**Figure 4C**).

Interestingly, it has been described that SirA can bind to the promoter region of *spf* (Zere et al., 2015), suggesting that SirA might be regulating Spot 42. However, transcriptomic data of Δ*sirA* compared to WT in *Salmonella* indicates that Spot 42 it is not regulated by SirA (Colgan et al., 2016). The effect of SirA on Spot 42 expression was assessed, using a chromosomal *spf-lacZ* in the presence and absence of SirA in both WT and Δ*crp* genetic backgrounds at logarithmic growth phase. As previously shown, Spot 42 expression is induced in the Δ*crp* mutant compared to WT (El Mouali et al., 2018; **Figure 5A**). Remarkably, no regulation was observed in the absence of SirA and induction of Spot 42 expression in the Δ*crp* mutant strain is still observed in absence of SirA (**Figure 5A**). These results indicate that SirA it is not regulating Spot 42 expression under the studied conditions. Furthermore, SirA was not required for the positive

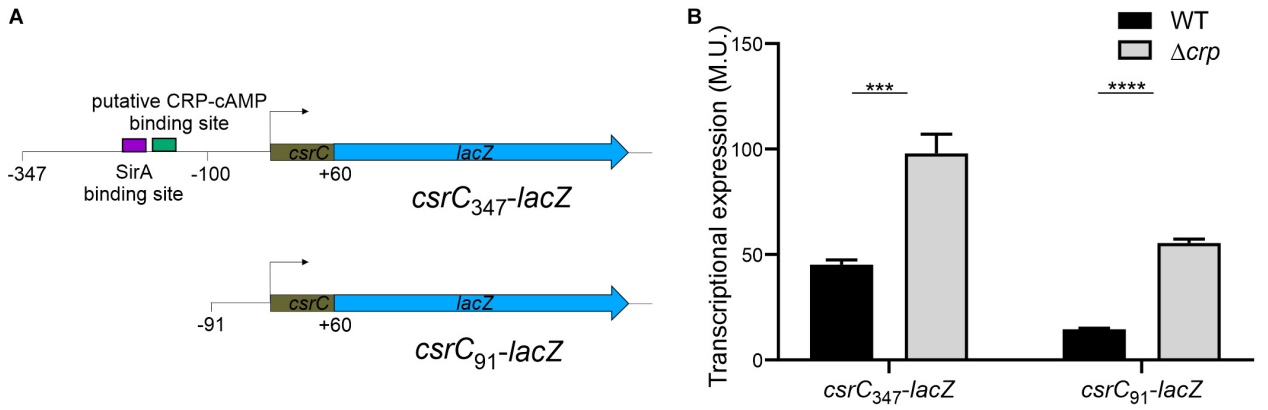


FIGURE 3 | CRP-cAMP represses CsrC in absence of SirA binding site. **(A)** Diagram of the *lacZ* transcriptional fusions present in the pQF*csrC* and pQF*csrC-91* plasmids. The *csrC* upstream sequences cloned in pQF50, –347/ + 60 and –91/+60, are depicted. The relative position of the SirA binding site and the putative CRP-cAMP binding site are indicated with a purple and green rectangle. **(B)** Transcriptional expression of the *lacZ* fusions described in A in WT and Δ*crp* genetic backgrounds. In all cases, cultures were grown in LB at 37°C up to an $OD_{600\,nm}$ of 0.4. β-galactosidase activity was determined for three independent cultures, average and standard deviation are presented. ***$p < 0.001$; ****$p < 0.0001$.

effect of Spot 42 on CsrC expression as concluded from *csrC-lacZ* expression studies under Spot 42 overexpression in both WT and Δ*sirA*. In both genetic backgrounds, the overexpression of Spot 42 induces the expression of *csrC-lacZ* to a similar fold (**Figure 5B**), consistent with no involvement of SirA in the CRP-cAMP mediated regulation of CsrC.

Altogether, our data indicate that CRP-cAMP differentially regulates CsrB and CsrC at the logarithmic growth phase where it represses specifically CsrC but not CsrB. The CRP-cAMP regulated sRNA Spot 42 arise as new regulator of the Csr regulon in *Salmonella*.

DISCUSSION

CRP-cAMP is a global regulator initially described to regulate metabolic genes in response to nutrient stimuli. Adding to its more prominent role as transcriptional regulator of mRNAs, CRP-cAMP was also described to regulate the expression of sRNAs such as CyaR, Spot 42, and FnrS (Polayes et al., 1988; Papenfort et al., 2008; De Lay and Gottesman, 2009; Durand and Storz, 2010). CRP-cAMP also modulates the expression of the sRNAs CsrB and CsrC which regulate CsrA protein activity via titration. Previous studies reveal apparent discrepancies in the role of CRP-cAMP on CsrB and CsrC expression. In *E. coli*, CsrB and CsrC are both repressed by CRP-cAMP, being CsrC expression directly repressed by CRP-cAMP via binding competition with the transcriptional activator UvrY (Pannuri et al., 2016). CRP-cAMP activity is modulated, among others, by the phosphorylation state of the $EIIA^{Glc}$, which is involved in glucose transport. When phosphorylated, it stimulates cAMP production, promoting the activity of CRP-cAMP. Surprisingly, dephosphorylated $EIIA^{Glc}$ interacts with CsrD and promotes the degradation of CsrB and CsrC by RNAseE (Leng et al., 2016). In other words, conditions that promote the transcriptional derepression of CsrC and CsrB via CRP-cAMP also promote the degradation of these transcripts (Leng et al., 2016; Pannuri et al., 2016). In contrast, in *Salmonella* CRP-cAMP has been described to play a positive role on CsrB and CsrC expression when it is grown on LB agar media (Teplitski et al., 2006). Although we corroborate these data, we also demonstrate that in LB liquid media, CRP-cAMP differentially regulates CsrB and CsrC. At the logarithmic phase, CsrC but not CsrB is repressed. The alterations in the cell physiology when growing planktonically and within a colony may account for the different regulation described for CsrB and CsrC in *Salmonella.* In addition of being differentially regulated by CRP-cAMP, CsrB, and CsrC depicted distinct expression profiles during the growth curve. CsrB expression seems to be constitutive, whereas, CsrC expression is silenced during logarithmic growth and induced in early stationary phase. The fact that production of CsrC but not CsrB is growth phase dependent indicates that each RNA responds differently to specific environmental inputs. CRP-cAMP is involved in the growth-dependent regulation of CsrC, suggesting that specific physiological signals that alter CRP-cAMP levels would cause alterations in the levels of CsrC and the concomitant alterations of free CsrA levels that would modulate gene expression to promote adaptation to the new conditions. The differential regulation of CsrB and CsrC, not only by CRP-cAMP, but potentially by additional regulators would provide sensitivity to the Csr system. Far from an ON/OFF state, the Csr regulon would display a scale of grays that would allow the bacteria to fine-tune gene expression in response to environmental stimuli. Interestingly, differential regulation of CsrB and CsrC by CRP-cAMP has been reported in *Yersinia pseudotuberculosis*, where CRP-cAMP activates the expression of CsrC and represses the expression of CsrB, highlighting complex species-specific regulation of the Csr system (Heroven et al., 2012).

SirA is an activator of CsrB and CsrC (Teplitski et al., 2006; Martínez et al., 2011, 2014. While we observe that SirA is required for full activation of CsrC, our data let us conclude that SirA is not involved in the deregulation of

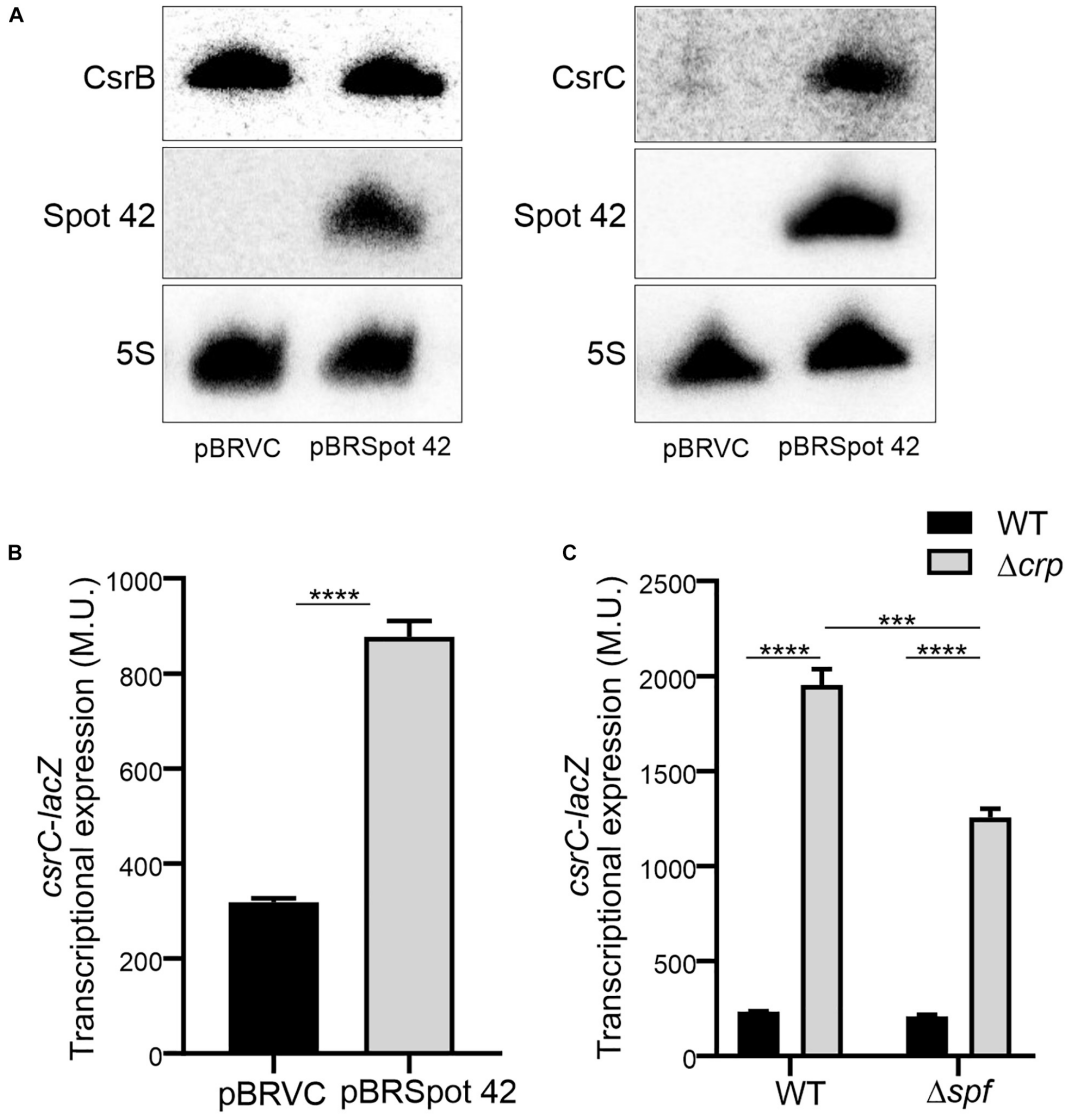


FIGURE 4 | Spot 42 positively regulates the expression of CsrC. **(A)** Northern blot detection of CsrB, CsrC, and Spot 42 was carried out in strains carrying either the pBRplacVC (control) or pBRplac Spot 42 (overexpressing the sRNA Spot 42). 5S RNA was monitored as loading control. **(B)** Transcriptional expression of the chromosomal *csrC-lacZ* fusion upon ectopic expression of the sRNA Spot 42 compared to the strain carrying the control vector (pBRplacVC). **(C)** Transcriptional expression of the chromosomal *csrC-lacZ* fusion was monitored in WT, Δ*spf*, Δ*crp,* and Δ*spf* Δ*crp* genetic backgrounds. In all cases cultures were grown in LB at 37°C up to an $OD_{600\,nm}$ of 0.4. β-galactosidase activity was determined for three independent cultures, average and standard deviation is presented. $^{****}p < 0.0001$, $^{***}p < 0.001$.

CsrC when CRP-cAMP is depleted. The CsrC deregulation was detected both in absence of SirA and when the SirA binding site was removed from the *csrC* promoter. Our data indicate that the sRNA Spot 42, which is transcriptionally repressed by CRP-cAMP at the logarithmic growth phase, positively regulates CsrC expression but it does not affect CsrB (**Figure 6**). This suggests a model where CRP-cAMP differentially regulates CsrB and CsrC via derepression of the trans-encoded sRNA Spot 42. Of note, Spot 42 does not seem to be solely responsible for the derepression of CsrC in absence of *crp*, as a partial derepression of CsrC is still observed in absence of *crp* and *spf*. Suggesting that, CRP-cAMP additionally represses CsrC expression through a Spot 42 independent mechanism (**Figure 6**).

The role of Spot 42 in the regulation of CsrC seem to be restricted to logarithmic growth phase and it is not involved in the stationary phase dependent induction. Expression of *csrC-lacZ* is induced at stationary phase compared to logarithmic growth phase in both the WT and the Δ*spf* backgrounds (**Supplementary Figure S3**). Consistently, transcriptomic data indicates that the expression level of Spot 42 is downregulated 100-fold upon entry into stationary phase when compared to logarithmic growth (Kröger et al., 2013). In agreement, Spot 42 might be responsible to fine-tune the expression of CsrC

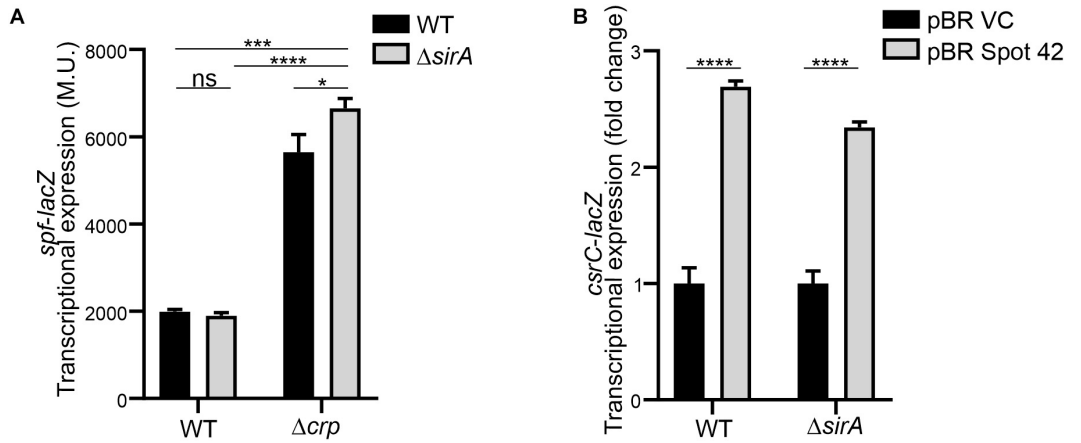


FIGURE 5 | Spot 42 regulates CsrC in a SirA-independent manner. **(A)** Transcriptional expression of the chromosomal *spf-lacZ* fusion (Spot 42 expression) was monitored in WT, Δ*crp*, Δ*sirA,* and Δ*crp*Δ*sirA* genetic backgrounds. **(B)** Transcriptional expression of the chromosomal *csrC-lacZ* fusion upon ectopic expression of the sRNA Spot 42 compared to the strain carrying the control vector (pBRplacVC) in a WT and a Δ*sirA* backgrounds. The transcriptional expression is shown in relative values. In each case, the reference (pBRVC) was set as 1.0. Miller units of the pBRVC backgrounds were 159 ± 21 and 20 ± 2.2 for WT and Δ*sirA,* respectively. In all cases, cultures were grown in LB at 37°C up to an $OD_{600\ nm}$ of 0.4. β-galactosidase activity was determined for three independent cultures, average and standard deviation are presented. *$p < 0.05$; ***$p < 0.001$; ****$p < 0.0001$; ns, not significant.

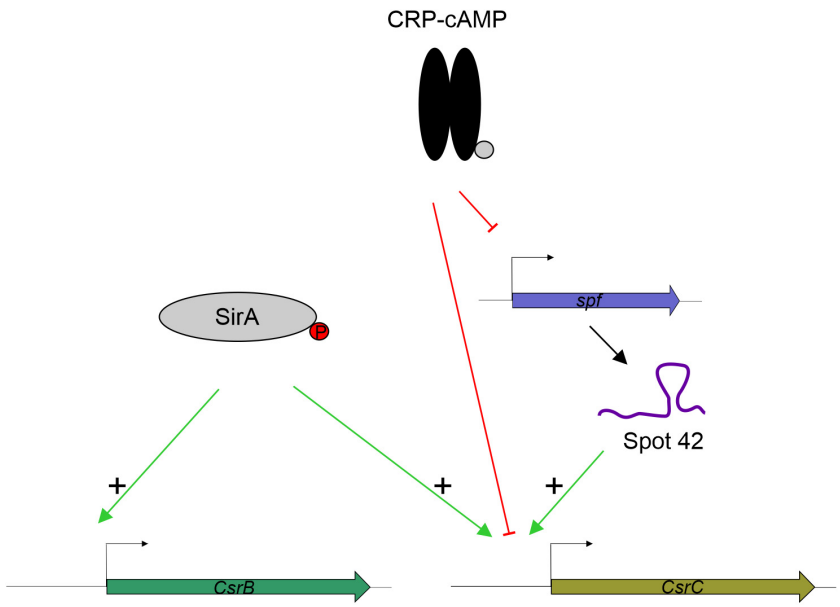


FIGURE 6 | Proposed model for CRP-cAMP mediated repression of CsrC at the logarithmic growth phase. In green, positive regulation is indicated. In red, negative regulation is indicated. SirA, when phosphorylated, it positively regulates CsrB and CsrC. CRP-cAMP represses CsrC and Spot 42. In addition, Spot 42 positively regulates CsrC levels.

in conditions where the CRP-cAMP activity is downregulated similarly as it occurs for the Spot 42-regulated *hilD* mRNA (El Mouali et al., 2018).

In this report, we provide new insights in the link between the CRP-cAMP and Csr regulons in *Salmonella*. We demonstrate that CRP-cAMP differentially regulates CsrB and CsrC. These two sRNAs are considered functionally redundant but our results suggest that they respond distinctly under specific environmental and physiological conditions in *Salmonella*. The presence of CsrB alone would affect the free pool of CsrA to a lesser extent than when both CsrB and CsrC sRNAs are present. This fine-tuning of the free CsrA levels would affect the Csr regulon where high affinity mRNA targets will require reduced amounts of free CsrA whereas low affinity targets will require full derepression of CsrA. We show that CRP-cAMP and its target sRNA Spot 42 contribute to the overall levels of CsrB and CsrC. Its contribution allows *Salmonella* to tightly control the levels of free CsrA in response to environmental stimuli. These features contribute to

the versatility of pathogenic bacteria such as the model organism *Salmonella enterica*.

MATERIALS AND METHODS

Bacterial Strains, Plasmids, and Growth Conditions

The bacterial strains, derivatives of *Salmonella enterica* serovar Typhimurium SL1344, were cultivated in Lysogeny broth (LB; tryptone 10 g/l, yeast extract 5 g/l and sodium chloride 10 g/l). When required, media was supplemented with the antibiotics indicated, ampicillin (Amp) 100 μg/ml, chloramphenicol (Cm) 15 μg/ml or kanamycin (Km) 50 μg/ml. Bacterial cultures were inoculated at an $OD_{600\,nm}$ of 0.001, cultures reaching $OD_{600\,nm}$ 0.4 were considered logarithmic growth and reaching $OD_{600\,nm}$ 2.0 was considered early stationary phase. The bacterial strains and plasmids used in this study are listed in **Supplementary Tables S1, S2**, respectively.

Genetic Manipulations

Deletion strains were generated by standard gene replacement as previously described (Datsenko and Wanner, 2000). For chromosomal *csrC-lacZ* transcriptional fusion. The *csrC*:Cm strain was cured from the antibiotic resistance as previously described (Cherepanov and Wackernagel, 1995). The resulting *csrC*:*frt* strain carrying the FRT scar was further used to obtain chromosomal *csrC-lacZ* by integration of pKG136 as previously described (Ellermeier et al., 2002). Chromosomal *spf-lacZ* was obtained as for CsrC (El Mouali et al., 2018). Oligonucleotides used for strains construction are listed in **Supplementary Table S3**.

Plasmidic transcriptional fusions were generated for *csrB* and *csrC*. The regulatory regions of interest were PCR amplified, *Bam*HI/*Hin*dIII digested and ligated within pQF50 (Farinha and Kropinski, 1989). Spot 42 was cloned in pBRplac (Guillier and Gottesman, 2006; Beisel and Storz, 2011). Spot 42 was PCR amplified, *Aat*II/*Eco*RI digested and ligated in pBRplac. Spot 42 was expressed constitutively in *Salmonella* (El Mouali et al., 2018). Oligonucleotides used for cloning are listed in **Supplementary Table S3**.

β-Galactosidase Activity Assay

Strains of interest were grown to logarithmic growth phase ($OD_{600\,nm}$ 0.4) or early stationary phase ($OD_{600\,nm}$ 2.0). β-galactosidase activity was measured as described previously (Miller, 1992). Shortly, 100 μl of culture was added to 900 μl of buffer Z (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, and 50 mM 2-mercaptoethanol, pH 7) and cells were lysed by the addition of 10 μl of toluene. Reactions were incubated at 28°C and β-galactosidase activity was measured upon addition of 200 μl of ONPG (4 mg/ml). Reactions were stopped by the addition of 500 μl of Na_2CO_3 (1 M). The $OD_{420\,nm}$ and $OD_{550\,nm}$ was measured and Miller Units were calculated as previously described (Miller, 1992). β-galactosidase activity determination was performed in technical duplicates for each of three biological replicates.

Total RNA Isolation and Northern Blot

Strains of interest were grown to logarithmic growth phase ($OD_{600\,nm}$ 0.4). The biomass of 4 units of $OD_{600\,nm}$ was collected, and total RNA extracted by classic hot phenol method. Shortly, the cells were resuspended in 600 μl of TE (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) containing 0.5 mg/ml lysozyme. Then, 60 μl of 10% SDS (w/v) was added; mixed by inversion and incubated at 64°C for 1–2 min. After the incubation, 66 μl of sodium acetate pH 5.2 (1 M) was added and mixed by inversion. For RNA extraction, 750 μl of Roti-Aqua phenol was added to the samples, mixed by inversion and incubated at 64°C for 6 min. Upon centrifugation (15 min, 13,000 rpm, 4°C), the top aqueous layer was transferred to a fresh 2 ml microcentrifuge tube, and 750 μl of chloroform was added. The samples were mixed and upon centrifugation (12 min, 13,000 rpm, 15°C), the upper aqueous layer was transferred into a new tube and precipitated with 30:1 mix of ethanol and sodium acetate (1 M, pH 6.5). The samples were incubated for 2–3 h or overnight at −20°C. The samples were then centrifuged and precipitated RNA was resuspended in water and concentration measured by NanoDrop. Samples of 10 μg of total RNA were subjected to electrophoretic separation in Tris-Borate-EDTA (TBE) 8% acrylamide gels containing 8.3 M urea. RNAs were transferred to Hybond N + (GE Healthcare) filters by semi-dry TBE based transference. Transcripts of interest were detected by hybridization with 5′ radiolabeled oligos as probes. Images were obtained with the FLA-5100 imaging system (Fujifilm). Oligonucleotides used as probes are listed in **Supplementary Table S3**.

Statistical Analysis

Graph Pad 8.0 software was used for data analysis. Unpaired two-tailed Student's *t*-test were carried out for two groups comparison and $p < 0.05$ were considered significant.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

YE, GE-M, and DG-P contributed to the investigation. YE and CB contributed to the conceptualization, investigation, formal analysis and writing the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Spanish Ministry of Economy and Competitiveness (grant AGL2013-45339R), the Spanish Ministry of Science, Innovation and Universities (MCIU), State Bureau of Investigation (AIE), and European Regional Development Fund (FEDER) (grant PGC2018-096958-B-I00), and the Catalonian Government (grant 2017SGR499). YE was recipient of an APIF fellowship from the University of Barcelona.

ACKNOWLEDGMENTS

We thank Caroline Tawk (Yale University) and Jens Hör (University of Würzburg) for critical reading of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.570536/full#supplementary-material

REFERENCES

Beisel, C. L., and Storz, G. (2011). The base-pairing RNA Spot 42 participates in a multioutput feedforward loop to help enact catabolite repression in *Escherichia coli*. *Mol. Cell* 41, 286–297. doi: 10.1016/j.molcel.2010.12.027

Cherepanov, P. P., and Wackernagel, W. (1995). Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene* 158, 9–14. doi: 10.1016/0378-1119(95)00193-a

Colgan, A. M., Kröger, C., Diard, M., Hardt, W. D., Puente, J. L., Sivasankaran, S. K., et al. (2016). The impact of 18 ancestral and horizontally-acquired regulatory proteins upon the transcriptome and sRNA landscape of *Salmonella enterica* serovar Typhimurium. *PLoS Genet.* 12:e1006258. doi: 10.1371/journal.pgen.1006258

Datsenko, K. A., and Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A.* 97, 6640–6645. doi: 10.1073/pnas.120163297

De Lay, N., and Gottesman, S. (2009). The Crp-activated small noncoding regulatory RNA CyaR (RyeE) links nutritional status to group behavior. *J. Bacteriol.* 191, 461–476. doi: 10.1128/JB.01157-08

Durand, S., and Storz, G. (2010). Reprogramming of anaerobic metabolism by the FnrS small RNA. *Mol. Microbiol.* 75, 1215–1231. doi: 10.1111/j.1365-2958.2010.07044.x

El Mouali, Y., Gaviria-Cantin, T., Sánchez-Romero, M. A., Gibert, M., Westermann, A. J., Vogel, J., 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

Ellermeier, C. D., Janakiraman, A., and Slauch, J. M. (2002). Construction of targeted single copy lac fusions using lambda Red and FLP-mediated site-specific recombination in bacteria. *Gene* 290, 153–161. doi: 10.1016/s0378-1119(02)00551-6

Farinha, M. A., and Kropinski, A. M. (1989). Construction of broad-host-range vectors for general cloning and promoter selection in *Pseudomonas* and *Escherichia coli*. *Gene* 77, 205–210. doi: 10.1016/0378-1119(89)90068-1

Fortune, D. R., Suyemoto, M., and Altier, C. (2006). Identification of CsrC and characterization of its role in epithelial cell invasion in *Salmonella enterica* serovar Typhimurium. *Infect. Immun.* 74, 331–339. doi: 10.1128/IAI.74.1.331-339.2006

Görke, B., and Stülke, J. (2008). Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. *Nat. Rev. Microbiol.* 6, 613–624. doi: 10.1038/nrmicro1932

Guillier, M., and Gottesman, S. (2006). Remodelling of the *Escherichia coli* outer membrane by two small regulatory RNAs. *Mol. Microbiol.* 59, 231–247. doi: 10.1111/j.1365-2958.2005.04929.x

Heroven, A. K., Sest, M., Pisano, F., Scheb-Wetzel, M., Steinmann, R., Böhme, K., et al. (2012). Crp Induces Switching of the CsrB and CsrC RNAs in *Yersinia pseudotuberculosis* and links nutritional status to virulence. *Front. Cell. Infect. Microbiol.* 2:158. doi: 10.3389/fcimb.2012.00158

Hör, J., Matera, G., Vogel, J., Gottesman, S., and Storz, G. (2020). Trans-acting small RNAs and their effects on gene expression in *Escherichia coli* and *Salmonella enterica*. *EcoSal Plus* 9. doi: 10.1128/ecosalplus.ESP-0030-2019

Johnston, C., Pegues, D. A., Hueck, C. J., Lee, C. A., and Miller, S. I. (1996). Transcriptional activation of *Salmonella* Typhimurium invasion genes by a member of the phosphorylated response-regulator superfamily. *Mol. Microbiol.* 22, 715–727. doi: 10.1046/j.1365-2958.1996.d01-1719.x

Kröger, C., Colgan, A., Srikumar, S., Händler, K., Sivasankaran, S. K., Hammarlöf, D. L., et al. (2013). An infection-relevant transcriptomic compendium for *Salmonella enterica* serovar Typhimurium. *Cell Host Microbe* 14, 683–695. doi: 10.1016/j.chom.2013.11.010

Leng, Y., Vakulskas, C. A., Zere, T. R., Pickering, B. S., Watnick, P. I., Babitzke, P., et al. (2016). Regulation of CsrB/C sRNA decay by EIIA(Glc) of the phosphoenolpyruvate: carbohydrate phosphotransferase system. *Mol. Microbiol.* 99, 627–639. doi: 10.1111/mmi.13259

Liu, M. Y., Gui, G., Wei, B., Preston, J. F., Oakford, L., Yüksel, Ü, et al. (1997). The RNA molecule CsrB binds to the global regulatory protein CsrA and antagonizes its activity in *Escherichia coli*. *J. Biol. Chem.* 272, 17502–17510. doi: 10.1074/jbc.272.28.17502

Martínez, L. C., Martinez-Flores, I., Salgado, H., Fernandez-Mora, M., Medina-Rivera, A., Puente, J. L., et al. (2014). *In Silico* identification and experimental characterization of regulatory elements controlling the expression of the *Salmonella csrB* and *csrC* Genes. *J. Bacteriol.* 196, 325–336. doi: 10.1128/JB.00806-13

Martínez, L. C., Yakhnin, H., Camacho, M. I., Georgellis, D., Babitzke, P., Puente, J. L., 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, 1637–1656. doi: 10.1111/j.1365-2958.2011.07674.x

Miller, J. H. (1992). *A Short Course in Bacterial Genetics: A Laboratory Manual*. Cold Spring Harbor, NY: CSHL Press.

Pannuri, A., Vakulskas, C. A., Zere, T., McGibbon, L. C., Edwards, A. N., Georgellis, D., et al. (2016). Circuitry linking the catabolite repression and Csr global regulatory systems of *Escherichia coli*. *J. Bacteriol.* 198, 3000–3015. doi: 10.1128/JB.00454-416

Papenfort, K., Pfeiffer, V., Lucchini, S., Sonawane, A., Hinton, J. C. D., and Vogel, J. (2008). Systematic deletion of *Salmonella* small RNA genes identifies CyaR, a conserved CRP-dependent riboregulator of OmpX synthesis. *Mol. Microbiol.* 68, 890–906. doi: 10.1111/j.1365-2958.2008.06189.x

Polayes, D. A., Rice, P. W., Garner, M. M., and Dahlberg, J. E. (1988). Cyclic AMP-cyclic AMP receptor protein as a repressor of transcription of the spf gene of *Escherichia coli*. *J. Bacteriol.* 170, 3110–3114. doi: 10.1128/jb.170.7.3110-3114.1988

Romeo, T., and Babitzke, P. (2018). Global regulation by CsrA and its RNA antagonists. *Microbiol. Spectr.* 6:e0009-2017. doi: 10.1128/microbiolspec.rwr-0009-2017

Teplitski, M., Goodier, R. I., and Ahmer, B. M. M. (2003). Pathways leading from BarA/SirA to motility and virulence gene expression in *Salmonella*. *J. Bacteriol.* 185, 7257–7265. doi: 10.1128/jb.185.24.7257-7265.2003

Teplitski, M., Goodier, R. I., and Ahmer, B. M. M. (2006). Catabolite repression of the SirA regulatory cascade in *Salmonella enterica*. *Int. J. Med. Microbiol.* 296, 449–466. doi: 10.1016/j.ijmm.2006.06.001

Weilbacher, T., Suzuki, K., Dubey, A. K., Wang, X., Gudapaty, S., Morozov, I., et al. (2003). A novel sRNA component of the carbon storage regulatory system of *Escherichia coli*. *Mol. Microbiol.* 48, 657–670. doi: 10.1046/j.1365-2958.2003.03459.x

Zere, T. R., Vakulskas, C. A., Leng, Y., Pannuri, A., Potts, A. H., Dias, R., et al. (2015). Genomic targets and features of BarA-UvrY (-SirA) signal transduction systems. *PLoS One* 10:e0145035. doi: 10.1371/journal.pone.0145035

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

REVIEW
published: 21 October 2020
doi: 10.3389/fmicb.2020.596914

Naturally Occurring tRNAs With Non-canonical Structures

Natalie Krahn[1†], *Jonathan T. Fischer*[1†] *and Dieter Söll*[1,2*]

[1] Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT, United States, [2] Department of Chemistry, Yale University, New Haven, CT, United States

Transfer RNA (tRNA) is the central molecule in genetically encoded protein synthesis. Most tRNA species were found to be very similar in structure: the well-known cloverleaf secondary structure and L-shaped tertiary structure. Furthermore, the length of the acceptor arm, T-arm, and anticodon arm were found to be closely conserved. Later research discovered naturally occurring, active tRNAs that did not fit the established 'canonical' tRNA structure. This review discusses the non-canonical structures of some well-characterized natural tRNA species and describes how these structures relate to their role in translation. Additionally, we highlight some newly discovered tRNAs in which the structure–function relationship is not yet fully understood.

Keywords: tRNA, non-canonical, genetic code expansion, identity elements, translation, selenocysteine, pyrrolysine, mitochondria

OPEN ACCESS

Edited by:
Omar Orellana, University of Chile, Chile

Reviewed by:
Manuel Santos, University of Aveiro, Portugal
Lennart Randau, University of Marburg, Germany

***Correspondence:**
Dieter Söll
dieter.soll@yale.edu

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *20 August 2020*
Accepted: *29 September 2020*
Published: *21 October 2020*

Citation:
Krahn N, Fischer JT and Söll D (2020) Naturally Occurring tRNAs With Non-canonical Structures. Front. Microbiol. 11:596914. doi: 10.3389/fmicb.2020.596914

INTRODUCTION

Transfer RNAs (tRNAs) range in length between 70 and 100 nucleotides. tRNAs are acylated with the cognate amino acid by their cognate aminoacyl-tRNA synthetase (aaRS), and the resulting aminoacyl-tRNAs are substrates for ribosomal protein synthesis. tRNAs were determined early on to have a highly conserved cloverleaf secondary structure (**Figure 1A**) (Holley et al., 1965) and an L-shaped tertiary structure (**Figure 1B**) (Cramer et al., 1969). The cloverleaf secondary structure is formed from Watson–Crick base pairs (bp) which create helical stems typically ending in unpaired bases to form loops. These arms (stem and loop) include the acceptor arm, D-arm, anticodon arm, TΨC arm (T-arm), and a variable arm. Of these features, the acceptor arm, anticodon arm and T-arm are highly conserved in size, while the D-arm and variable arm can differ.

The stems of the acceptor arm and T-arm are found to have 7 and 5 bp, respectively, to give the canonical 7/5 configuration. Occasionally, a base-pair mismatch is found in the acceptor stem, which does not disrupt the conformation of the helix. Furthermore, the acceptor arm has four 3′-terminal residues which are not base paired: the discriminator base (Crothers et al., 1972) and the CCA tail. The 3′-terminal adenosine is in the form of a slightly activated ester to bind to the amino acid. The 3′-strand of the acceptor stem is directly connected to the T-stem, while the 5′-strand is connected to the D-stem by two unpaired bases. The length of the D-stem varies amongst tRNAs; most of them have between three and 5 bp leading up to the D-loop. The anticodon stem is also highly conserved in length, being found to consistently have 5 bp before ending with the anticodon loop. The anticodon loop has 7 nucleotides, with the three residues in the center of the anticodon loop (anticodon) participating in mainly Watson–Crick (but also sometimes non-Watson–Crick or wobble base) interactions with the codon of the mRNA. Finally, the variable loop is the least conserved amongst all tRNAs (Sigler, 1975). tRNAs are classified into two groups based on the size of their variable loops. Most tRNAs fall into class I and have four or five nucleotides in the variable loop, while class II tRNAs including $tRNA^{Ser}$, $tRNA^{Leu}$, and $tRNA^{Tyr}$ have long variable loops consisting of 10 or more nucleotides (Sprinzl et al., 1998).

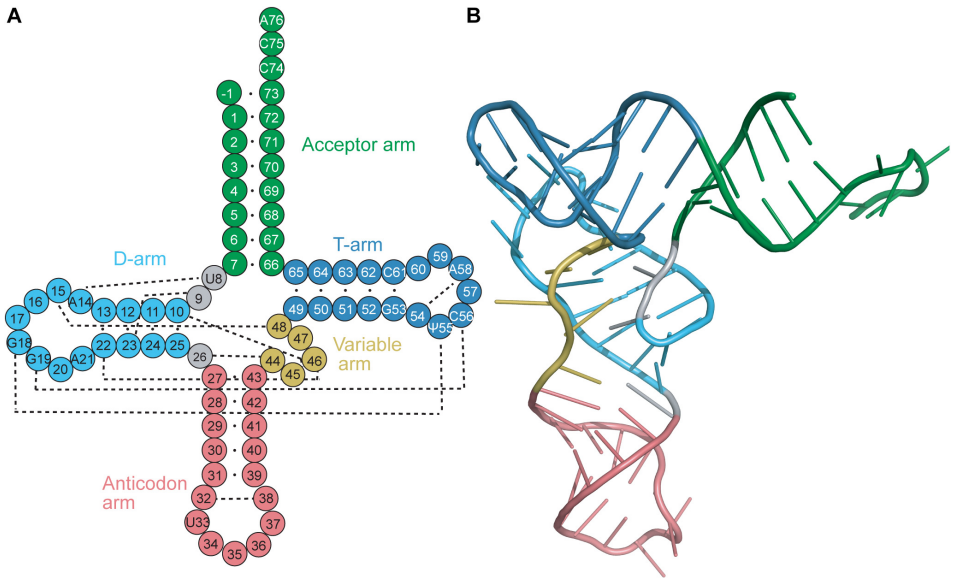


FIGURE 1 | (A) Secondary and **(B)** tertiary representation of a tRNA molecule colored based on specific regions: acceptor arm (green), T-arm (dark blue), D-arm (light blue), variable arm (yellow), anticodon arm (pink). The cloverleaf (secondary structure) model includes the standard tRNA numbering. Dashed lines correspond to the tertiary interactions which form the L-shape observed with tRNAs.

The conserved L-shape of tRNA is facilitated by base stacking and tertiary interactions between conserved or semi-conserved nucleotides. One arm of the L is formed from stacking of the acceptor stem with the T-stem, while the other arm is formed from stacking of the D-stem and anticodon stem. The conserved and semi-conserved nucleotides which form the tertiary interactions are found in the D- and T-loops; specifically G18 with Ψ55 and G19 with C56 (**Figure 1A**) (Hou, 1993). Additional tertiary interactions are found throughout the tRNA, including interactions between the D-stem and variable loop, the connecting base U8 and A14 in the D-loop, and stabilizing interactions within the T-loop and anticodon loop. Furthermore, stabilization often occurs at the level of the codon:anticodon interaction via tRNA modifications typically found at position 34, the wobble base in the anticodon, or position 37 which is just prior to the anticodon (Lyons et al., 2018). The conserved tRNA structure and sequences are crucial for functionality of the tRNA, including interaction with modifying enzymes (i.e., CCA-adding enzyme) and positioning in the ribosome (Lorenz et al., 2017).

Although tRNA structures are highly conserved, they do contain distinguishing elements which allow recognition by their cognate aaRS. These distinguishing elements, referred to as identity elements, are the only residues required for recognition by that aaRS. Common identity elements include the discriminator base and the anticodon; however, they are not limited to those regions. Identity elements of tRNAs have been extensively reviewed (Giegé et al., 1998) and therefore will not be discussed in detail here.

Instead, we focus on tRNAs with structures that deviate from canonical tRNAs. The nature of the proper secondary structure model of tRNAs was widely discussed (Hubert et al., 1998). Diverse experimental strategies (chemical and enzymatic RNA probing, phylogenetic analyses, and finally structural studies) showed the presence of natural tRNAs which lack the canonical 7/5 structure. As genomic studies expanded, many more tRNA genes with unique features were discovered. Some of these were poorly annotated due to the presence of unusual recognition elements, an anticodon sequence that disagreed with the other identity elements of the tRNA, or an irregular secondary structure.

tRNA^{Sec}

Discovered as the 21st amino acid in 1976, incorporation of selenocysteine (Sec) into proteins occurs naturally in all domains of life (Cone et al., 1976). Unlike the translational mechanism of inserting the first 20 identified amino acids into proteins, incorporation of Sec into selenoproteins is more nuanced and involves additional steps. First, a specialized tRNA^{Sec} initially becomes aminoacylated with serine (Ser) by seryl-tRNA synthetase (SerRS) to form Ser-tRNA^{Sec}. The serine hydroxyl group is then substituted with selenium to form Sec. In bacteria, this occurs in a single step with the enzyme selenocysteine synthase (SelA), while in archaea and eukaryotes, it is a two-step process involving first phosphorylation of the Ser with phosphoseryl-tRNA^{Sec} kinase (PSTK) followed by its replacement with selenium by *O*-phosphoseryl-tRNA^{Sec}:Sec synthase (SepSecS) (**Figure 2A**). The fully aminoacylated Sec-tRNA^{Sec} product is then transported to the ribosome and incorporated into the nascent peptide at a UGA codon via a specialized elongation factor, SelB (also referred to as eEFSec in eukaryotes). SelB distinguishes UGA codons for Sec incorporation over UGA stop codons through recognition of a

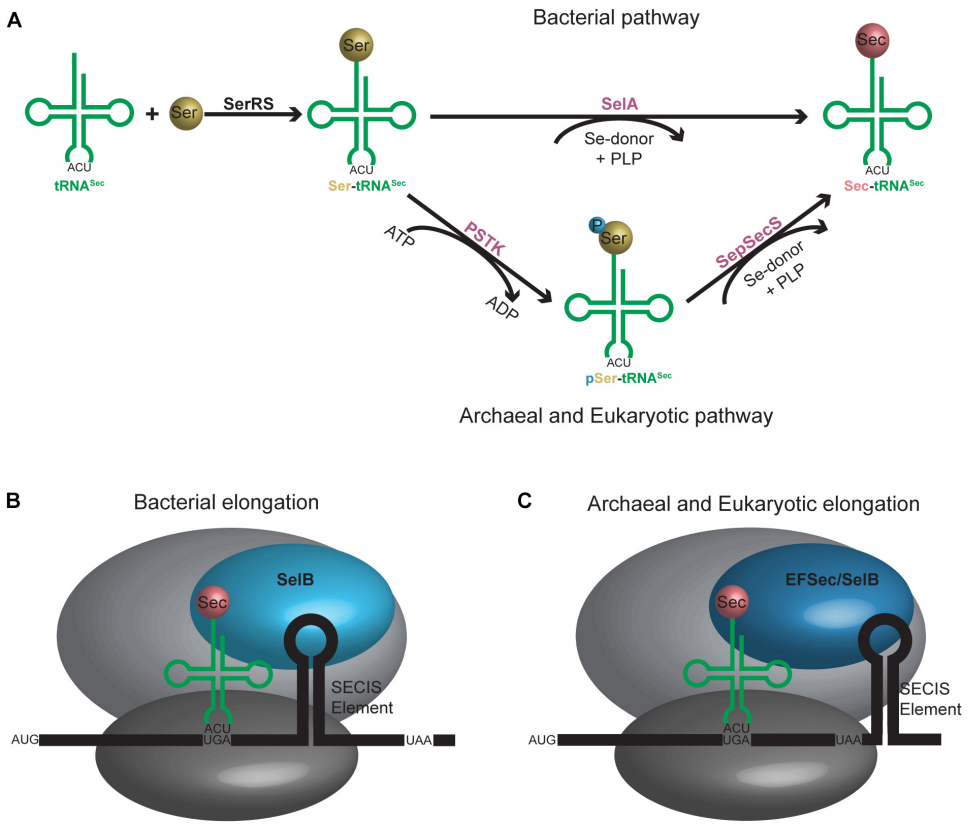


FIGURE 2 | (A) Aminoacylation pathway for tRNA^{Sec} is different in bacteria compared to archaea and eukaryotes. An additional step is required in the latter, resulting in an intermediate tRNA^{Sec} containing a phosphoserine moiety. **(B)** Bacterial and **(C)** archaeal and eukaryotic elongation pathways for tRNA^{Sec}. A selenocysteine insertion sequence (SECIS) element is required in the mRNA sequence which forms a hairpin in the 3′ translated region for bacteria or 3′ untranslated region for archaea and eukaryotes. A unique elongation factor [SelB (sometimes referred to as EFSec)] is required in all systems.

hairpin in the mRNA [**Sec** **I**nsertion **S**equence (SECIS) element] (**Figures 2B,C**) (reviewed in Serrao et al., 2018).

This rather complicated translational process is distinct from the translation pathway of the other 20 canonical amino acids. Therefore, it does not come as a surprise that tRNA^{Sec} does not conform to the structure of a canonical tRNA. From secondary structure predictions, it was evident that bacterial and archaeal tRNA^{Sec} both have a 13 bp acceptor domain (acceptor stem and T-stem) with an 8/5 and 9/4 structure, respectively (Schön et al., 1989; Sturchler et al., 1993; Hubert et al., 1998). However, this was not so clear for eukaryotes. The information obtained through modeling was torn between tRNA^{Sec} adopting a 9/4 structure like the archaeal tRNA^{Sec} or the canonical 7/5 structure observed in all other eukaryotic tRNAs known at the time (Ioudovitch and Steinberg, 1998; Steinberg et al., 1998). Without three-dimensional data to gather information from, researchers turned to phylogenetics to settle the debate. With eukaryotic and archaeal translational machineries being very similar to one another, it suggested that their tRNAs would have similar structures.

This provided the necessary evidence for eukaryotic tRNA^{Sec} to be accepted as having a 9/4 structure even before the structural data was able to confirm it over a decade later (Hubert et al., 1998; Itoh et al., 2009; Palioura et al., 2009). The 9/4 structure of eukaryotic tRNA^{Sec} would imply that the 8th residue would be base paired in the acceptor stem and unable to participate in binding to SerRS, a common interaction in canonical pairs of aaRSs and tRNAs. This is in contrast to what is found in *E. coli* tRNA^{Sec} with G8 forming a novel tertiary interaction with A21 and U14 (Sturchler et al., 1993). However, even with the proposed 9/4 arrangement, the overall tRNA secondary structure still has non-paired residues in positions 8 and 9 which in theory could participate in the aforementioned interactions. Crystal data show that this is not the case and there is instead an open cavity in the tertiary core with positions 8 and 9 not participating in any tertiary interactions. Instead, a different unique base triple is found in eukaryotic tRNA^{Sec} with U20 forming an interaction with the commonly found G19:C56 pair (Itoh et al., 2009). In addition to the canonical tertiary interactions between the D- and T-loops, a novel interaction was found between C16 in the D-loop and C59 in the T-loop for *E. coli* which is similar to the U16:U59 interaction in eukaryotic tRNA^{Sec}. Moreover, the tertiary interactions found between the variable arm and D-arm in canonical tRNAs are absent in tRNA^{Sec} (**Figure 3**). Although this was predicted to create a different orientation of the variable arm with respect to the overall L-shape of the tRNA through

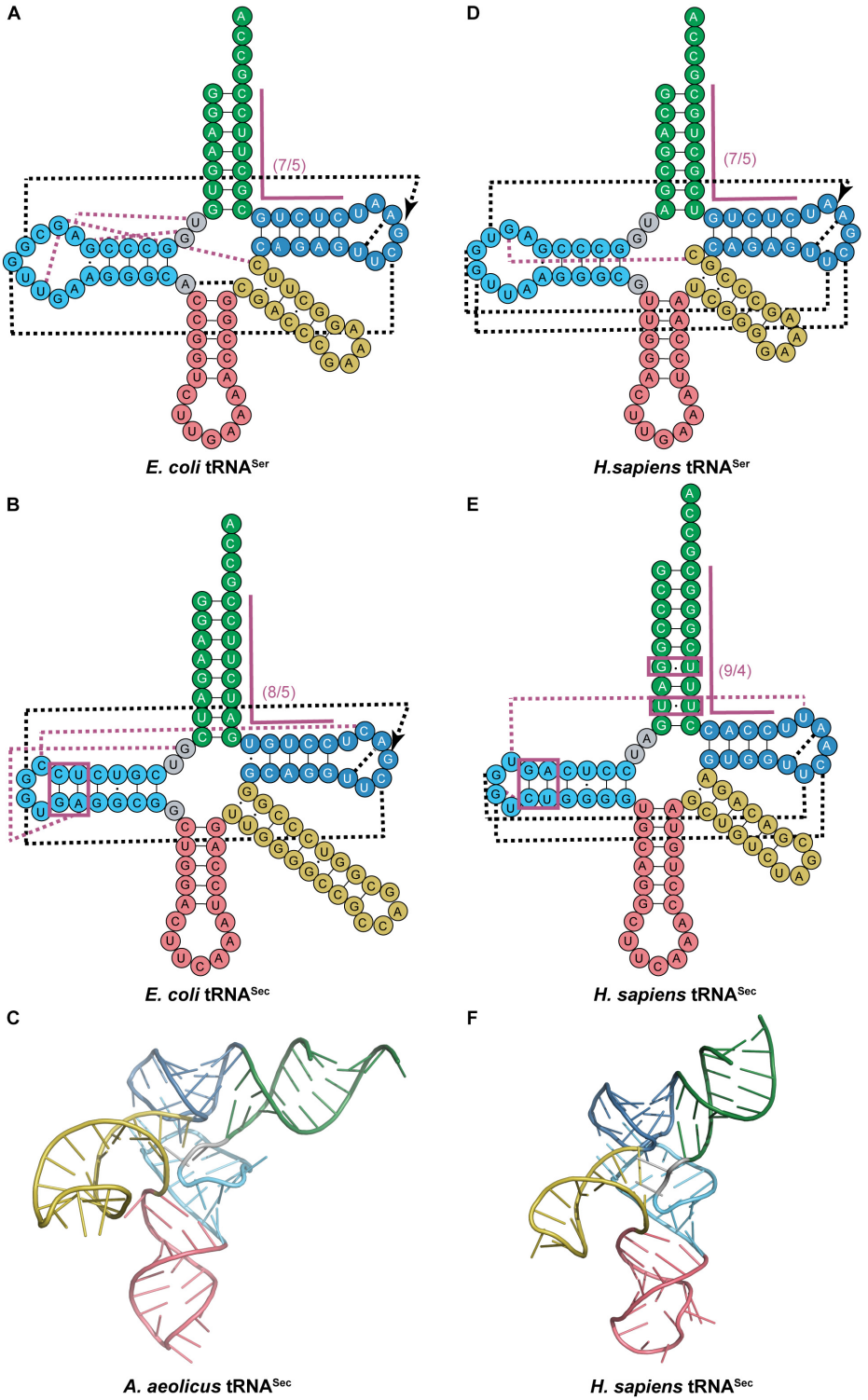


FIGURE 3 | Comparing cloverleaf structures of tRNA^{Ser} with tRNA^{Sec} from bacterial and eukaryotic species with specific focus on the tertiary structure of tRNA^{Sec}. Cloverleaf structures of **(A)** *E. coli* tRNA^{Ser} and **(B)** *E. coli* tRNA^{Sec} highlight differences in the acceptor domain and D-arm as well as different tertiary interactions. **(C)** Bacterial *A. aeolicus* tRNA^{Sec} (PDB ID: 3W3S; Itoh et al., 2013) forms the canonical L-shaped tertiary structure. A similar comparision is shown of **(D)** *H. sapiens* tRNA^{Ser} and **(E)** *H. sapiens* tRNA^{Sec}. **(F)** The tertiary structure of *H. sapiens* tRNA^{Sec} (PDB ID: 3A3A; Itoh et al., 2009) also shows a similar L-shape. The tRNA structure elements are colored accordingly: acceptor arm (green), T-arm (dark blue), D-arm (light blue), variable arm (yellow), and anticodon arm (pink). Tertiary interactions are represented by dashed lines with black lines being conserved interactions between the two tRNAs while magenta lines are unique to that tRNA. Magenta boxes highlight important regions of tRNA^{Sec} for interaction with aminoacylation and elongation machinery.

biochemical studies (Baron et al., 1993), structures indicate that indeed the variable arm of $tRNA^{Ser}$ and $tRNA^{Sec}$ are in an almost identical orientation with respect to the T-arm (Itoh et al., 2009).

The unique structure of $tRNA^{Sec}$, mainly the 13 bp acceptor domain structure and long variable arm, are essential for its function in translation (Mizutani and Goto, 2000). First, $tRNA^{Sec}$ must be recognized by SerRS, which suggests it must have the same identity elements as $tRNA^{Ser}$. Since SerRS only recognizes class II tRNAs, it follows that $tRNA^{Sec}$ is also a class II tRNA (Schatz et al., 1991; Heckl et al., 1998). More specifically, it is the long variable arm and G73 discriminator base which SerRS recognizes for serylation of both $tRNA^{Ser}$ and $tRNA^{Sec}$ (Breitschopf and Gross, 1994). The mechanism of how the long variable arm serves as an identity element remains to be elucidated. It is hypothesized that the orientation of the arm is more important than the actual sequence based off of the low conservation over different $tRNA^{Ser}$ sequences (Wu and Gross, 1993), but further studies suggest that there might be some sequence identity in the variable arm of $tRNA^{Sec}$ (Ohama et al., 1994). This coincides with the evidence that SerRS preferentially binds 12 bp acceptor domain tRNAs. To get efficient serylation of $tRNA^{Sec}$, which has a 13 bp acceptor domain, specific residues in the variable arm (and in the D-arm for eukaryotes) work together to promote synthetase binding and thus serylation (Ohama et al., 1994; Amberg et al., 1996; Fu et al., 2018).

Following serylation by SerRS, bacterial Ser-$tRNA^{Sec}$ must then be recognized by SelA for conversion to Sec-$tRNA^{Sec}$. One of the striking features that distinguishes $tRNA^{Sec}$ from $tRNA^{Ser}$ other than the acceptor domain, is the secondary structure of the D-arm. The D-stem in $tRNA^{Sec}$ has 6 bp with a 4 base D-loop while $tRNA^{Ser}$ has a 4 bp D-stem and 8–11 base D-loop (**Figures 3A,B**). Kinetic data initially suggested that the unique D-arm of bacterial $tRNA^{Sec}$ was the important feature for SelA discrimination against $tRNA^{Ser.}$ This was later confirmed by crystal structures to show that the N-terminal domain (NTD) of SelA is responsible for interacting with the D-arm of $tRNA^{Sec}$ (**Figure 3C**) (Itoh et al., 2013). Further studies found that the interaction of SelA is not sequence-based but rather structural. The presence of the 5^{th} and 6^{th} bp in the D-arm of $tRNA^{Sec}$ (regardless of sequence) had a positive impact on its activity (Ishii et al., 2013).

In eukaryotes, it is PSTK that must recognize Ser-$tRNA^{Sec}$ for phosphorylation while at the same time excluding Ser-$tRNA^{Ser}$. Similar to bacteria, the length and secondary structure of the D-arm differs between eukaryotic $tRNA^{Sec}$ and $tRNA^{Ser}$. The D-stem in $tRNA^{Sec}$ has 6 bp with a 4-base D-loop, while $tRNA^{Ser}$ has a 4 bp D-stem and 8-base D-loop (**Figures 3D,E**). Studies show that by simply adding 2 bp into the D-stem of $tRNA^{Ser}$, it becomes a substrate for PSTK. Alternatively, by increasing the 4-base D-loop of $tRNA^{Sec}$ to 8 bases, phosphorylation is decreased (Wu and Gross, 1994). These results suggest that the D-arm is a major identity element for PSTK recognition. It was also discovered that a minor contributor for successful phosphorylation is the T-stem length ($tRNA^{Sec}$ has a 4 bp T-stem and $tRNA^{Ser}$ has 5 bp). This was observed by the slight decrease in phosphorylation found by increasing the length of the T-stem in $tRNA^{Sec}$ (Wu and Gross, 1994). These findings were confirmed by the complex crystal structure. Crystal contacts were found between PSTK and the D-arm and T-stem of $tRNA^{Sec}$ (**Figure 3F**) (Chiba et al., 2010). In contrast to eukaryotic $tRNA^{Sec}$, the archaeal D-stem is 7 bp and not considered a major identity element for PSTK recognition (reduction to 5 bp caused only a minor decrease in phosphorylation). Instead, the 13 bp acceptor domain binds to the NTD of PSTK as the major contributor (Sherrer et al., 2011) while the minor contributor was formed through the D-stem binding the C-terminal domain (CTD) of PSTK (Sherrer et al., 2008).

After phosphorylation by PSTK, archaeal and eukaryotic $tRNA^{Sec}$ must also be recognized by SepSecS to form Sec-$tRNA^{Sec}$ for incorporation into the polypeptide chain. As with the other enzymes involved in Sec incorporation, SepSecS distinguishes $tRNA^{Sec}$ from the canonical tRNAs by the 13 bp acceptor domain (**Figure 3E**) (Amberg et al., 1996). The role of the 13 bp acceptor domain was determined to play a role in stabilization of $tRNA^{Sec}$ for interaction with SepSecS. The co-crystal structure revealed that the main interaction of $tRNA^{Sec}$ with SepSecS is through its acceptor domain, where it approaches the tRNA from the variable arm side and not through recognition of the D-arm (Palioura et al., 2009).

The final step in translation of $tRNA^{Sec}$ is elongation. The unique structure of $tRNA^{Sec}$ is also required for recognition by SelB and rejection from the traditional elongation factors (EF-Tu in bacteria and EF-1α in eukaryotes) used for canonical tRNAs. In bacteria, EF-Tu was found to bind with 100-fold weaker affinity to Ser-$tRNA^{Sec}$ than Ser-$tRNA^{Ser}$. This was found to be a result of the longer acceptor stem in $tRNA^{Sec}$ (8 bp compared to 7 bp). This long acceptor stem is also the major structural determinant for SelB binding and the feature which distinguishes $tRNA^{Sec}$ from the canonical tRNAs (Forster et al., 1990; Baron and Böck, 1991; Sturchler-Pierrat et al., 1995). With eukaryotes and archaea, the structural determinant for elongation is more specific. Phylogenetic considerations showed conservation of U6:U67 and a non-Watson–Crick base pair at 5a:67b in the acceptor stem of vertebrates, *Drosophila* and *Caenorhabditis elegans*. In vertebrates and *C. elegans* the sequence is a conserved wobble base pair G5a:U67b while in *Drosophila* it is replaced with A5a:G67b (highlighted in **Figure 3E**). Through randomization of both regions of the acceptor stem, it was found that the 5a:67b non-Watson–Crick interaction was imperative for function and that the U6:U67 pair was dispensable. The 5a:67b pair is believed to provoke structural modification of the phosphodiester backbone of the RNA helix for interaction with SelB (Mizutani et al., 1998b).

From the above evidence, the unique structure of $tRNA^{Sec}$ is warranted by the specific interactions it encounters compared to canonical tRNAs. Interestingly, although multiple enzymes interact with $tRNA^{Sec}$, none of them bind to the anticodon arm. Therefore, it follows that although $tRNA^{Sec}$ was initially found to have a UCA anticodon and that majority of species conform to this, there are quite a few $tRNA^{Sec}$ species with sense anticodons (Mukai et al., 2016). Through ongoing research in the field of Sec incorporation, a genomic search for $tRNA^{Sec}$ in other species revealed two other tRNA variants: *selC** $tRNA^{Cys}$ and allo-tRNA (Mukai et al., 2017).

tRNA^{Sec}-LIKE STRUCTURES

In a metagenomic search for additional tRNA^{Sec} species, some tRNAs were found that are structurally similar to, but do not function as tRNA^{Sec}. Further investigation into these unique structures led to the classification of two tRNA species: *selC** tRNA^{Cys} and allo-tRNA. Both tRNA groups contain the same distinctive tRNA^{Sec} structural features; a longer variable arm, acceptor stem, and anticodon stem compared to canonical tRNAs (Mukai et al., 2017).

selC* tRNA^{Cys}

*selC** tRNA^{Cys} are found in anaerobic bacteria from the phyla *Firmicutes*, *Thermodesulfobacteria*, *Nitrospirae*, and *Proteobacteria*. *selC** tRNAs were named after the *selC* gene which encodes tRNA^{Sec} in *E. coli* (Mukai et al., 2017). *selC** tRNA^{Cys} isoacceptors have similar structure to tRNA^{Sec} but contain identity elements of tRNA^{Cys} (notably the U73 discriminator base and cysteine (GCA) anticodon) (Pallanck et al., 1992; Komatsoulis and Abelson, 1993). Although the GCA anticodon is a strong identity element for CysRS recognition, some of the *selC** tRNA^{Cys} were found to have an opal (UCA) anticodon instead. Further evidence showed that some CysRS variants can cysteinylate tRNA$^{Cys}{}_{UCA}$, therefore including this group of tRNAs in this category (Turanov et al., 2009). The most striking feature of *selC** tRNA^{Cys} is their modified 8/4 structure. In the 4 bp T-stem an unpaired adenosine produces a bulge at position 51a. This is similar to what is found in the structure of minor bacterial (8/4) tRNA^{Sec} species (**Figure 4A**) (Mukai et al., 2017).

Further genomic analysis revealed that in two δ-proteobacterial subgroups, *Syntrophobacterales* and *Desulfobacterales*, a second copy of *selB* (*selB**) was found downstream of the *selC** genes. From this, it was hypothesized that *selC** tRNA^{Cys} is recognized by SelB* in a similar way as the 8/4 tRNA^{Sec} is recognized by SelB. Moreover, in *Desulfobacterales*, *selC** tRNA^{Cys} was found to contain an A1:U72 pair and an opal (UCA) anticodon. As previously mentioned, some CysRS variants (encoded by *cysS*) would be able to recognize the opal anticodon, however, they would be unable to recognize the A1:U72 pair. Therefore in these species a second copy of CysRS was found to be encoded downstream of *selB** (*cysS**) (**Figure 4B**) (Mukai et al., 2017). CysRS* lacks an anticodon binding domain, which allows for recognition of *selC** tRNA^{Cys} with an opal anticodon. Recognition of A1:U72 is possible due to mutations in the CP1 domain (**Figure 4C**) (Liu et al., 2012). These discoveries suggest that CysRS* specifically evolved to recognize and aminoacylate *selC** tRNA^{Cys}. *In vivo* analysis confirmed that CysRS* can aminoacylate *selC** tRNA^{Cys} through recognition of the 8 bp acceptor stem and the unique A51a bulge,

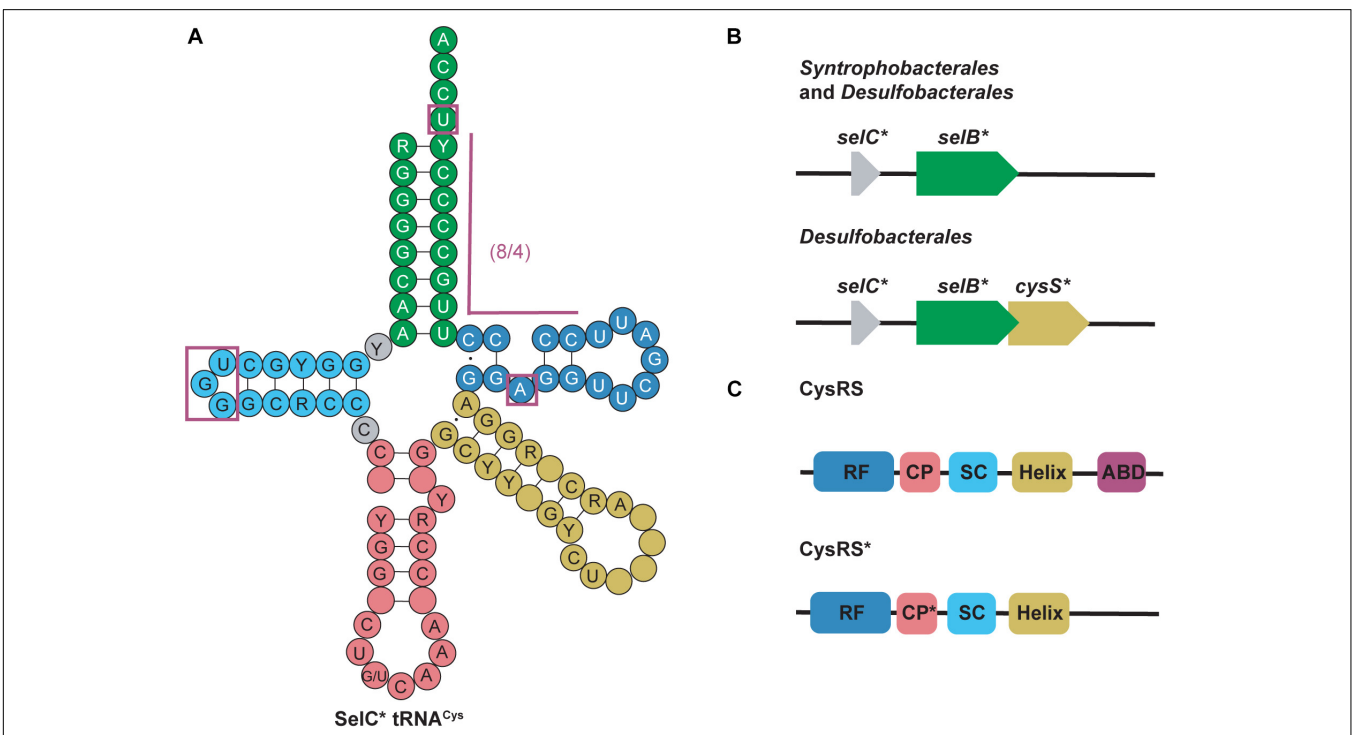


FIGURE 4 | (A) Cloverleaf structure of SelC* tRNA^{Cys} highlights its unique structure compared with canonical tRNAs. Magenta boxes emphasize these specific regions. R and Y denote A/G and U/C, respectively and empty circles represent no conservation in sequence. **(B)** Genomic structure of selC* reveals an additional elongation factor (*selB**) in *Syntrophobacterales* and *Desulfobacterales* while an additional aaRS (*cysS**) is present in only *Desulfobacterales*. **(C)** *cysS** (which codes for CysRS*) is an aaRS that has a mutated connective polypeptide (CP*) domain and the anticodon binding domain (ABD) is absent. The Rossman fold (RF) and stem-contact (SC) fold are consistent between CysRS and CysRS*

characteristic of *selC** tRNA^{Cys}, was found to be dispensable (Mukai et al., 2017).

Allo-tRNA

Allo-tRNA genes belong to bacteria from *Clostridia*, *Proteobacteria,* and *Acidobacteria*. They encode a unique tRNA whose striking feature is their 12 bp acceptor domain, which is found in either an 8/4 or 9/3 conformation (Mukai et al., 2017). Based on the presence of identity elements for SerRS recognition, allo-tRNAs were suggested to be Ser isoacceptors. With the knowledge that SerRS can recognize not only 7/5 tRNA^{Ser} but also 8/5, 9/4, and 8/4 tRNA^{Sec}, this was a reasonable hypothesis (Mizutani et al., 1998a). Moreover, most allo-tRNAs were found to have non-serine anticodons which SerRS does not recognize (Breitschopf and Gross, 1994). In fact, the anticodons are highly diverse and span 35 out of 64 codons. The most predominant anticodons for the 8/4 allo-tRNA species are UAU, GCG, and GUC which correspond to isoleucine (Ile), arginine (Arg), and aspartic acid (Asp), respectively. Conversely, anticodons UUC, GUC, CAC, and AAA corresponding to phenylalanine (Phe), valine (Val), histidine (His), and lysine (Lys) were only found once in the metagenomic data analyzed. Fewer 9/3 species were found and they contained anticodons which corresponded to Arg, leucine (Leu) and the ochre stop codon (UAA) (**Figure 5**) (Mukai et al., 2017).

In vivo studies began to examine the utility of these tRNAs in the bacterial translation system. Initially, two allo-tRNAs (an 8/4 and 9/3 structure) from *Silvibacterium bohemicum* were expressed in *E. coli* with position 2 of super-folder green fluorescent protein (sfGFP) mutated to Leu codons (CUC and UUA). Interestingly, Ser was efficiently incorporated into sfGFP as confirmed by fluorescence and mass spectrometry data. Other 8/4 and 9/3 allo-tRNAs were found to contain the major identity element for recognition by alanyl-tRNA synthetase (AlaRS), specifically the G3:U70 wobble base pair (Hou and Schimmel, 1988; Mcclain and Foss, 1988). Testing their capabilities *in vivo*, it was found that Ala and Ser were the main residues incorporated at an amber codon; however, insertion of other amino acids including Asn, Gln, Lys, Cys, Ile, and Glu were also detected. These studies showed that allo-tRNAs derived from other bacterial species could be efficiently used as a substrate in the *E. coli* translation system, and the nature of the incorporated

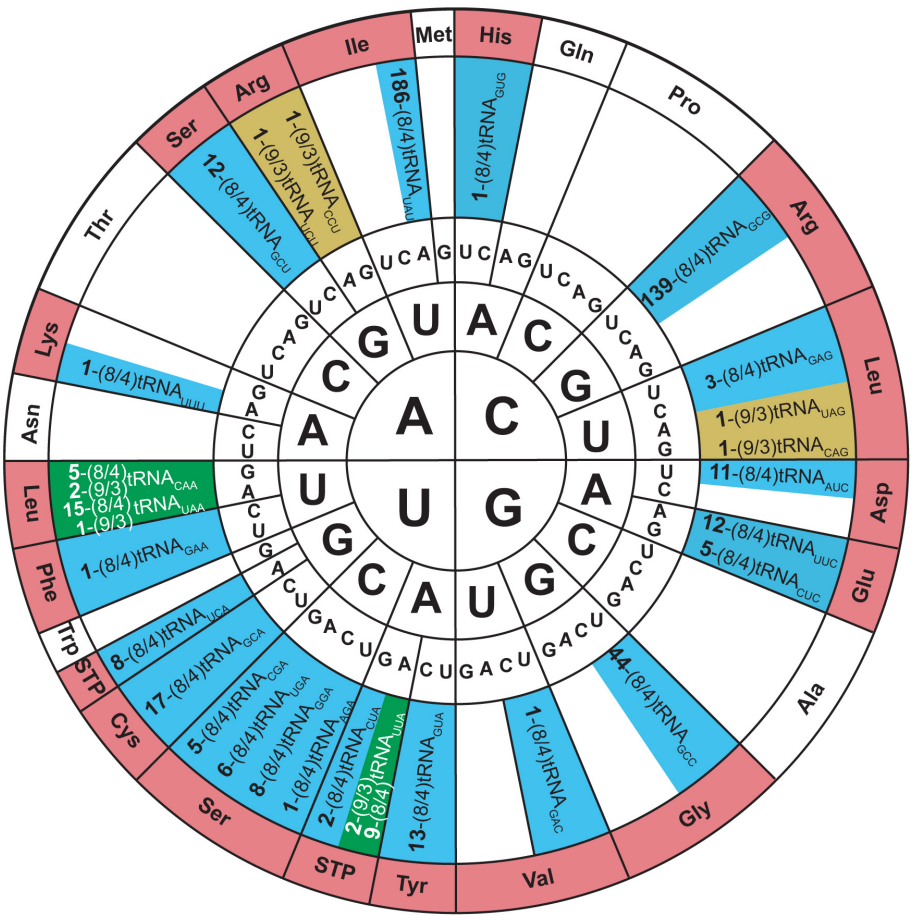


FIGURE 5 | Genetic code wheel highlights the codons and subsequent amino acids which can be incorporated by allo-tRNAs of 9/3 structure (yellow), 8/4 structure (blue) or both (green) (adapted from Mukai et al., 2017). The number of allo-tRNAs found with the indicated anticodon (subscript) is shown as a bold number in front of the allo-tRNA name. Some allo-tRNAs with a specific anticodon are thought to be able to read through multiple codons.

amino acid is based on the allo-tRNA identity elements rather than on the anticodon present (Mukai et al., 2017).

An interesting discovery was found in the *Edaphobacter* strain C40. An allo-tRNA$_{UAU}$ pseudogene with several base-pair disruptions was found overlapping with the open reading frame of a transposon-related protein. This allo-tRNA species was found to be the most abundant group among the allo-tRNA genes observed in the soil and sediment metagenomic sequences. Their cloverleaf structures were unlike allo-tRNA^{Ser}, containing stem-destabilizing mutations as in the *Edaphobacter* strain C40 and possible five-stem-junction structures (**Figure 6**). Unlike the previously described allo-tRNAs, these were unable to be used for translation in *E. coli* and were hypothesized to be associated with transposable elements or toxin-antitoxin systems. Moreover, polycistrons of allo-tRNA-like sequences and other irregular tRNA sequences were discovered from forest and peat soil metatranscriptomic data. Many of these tRNAs were predicted to have an 8/4 structure, but with additional features were found, including an extra loop in between the acceptor stem and D-stem as well as a G-1 base. These allo-tRNAs could not be aminoacylated by *E. coli* aaRSs *in vitro*. However, this does not answer the question whether they are used for translation in their original hosts with an aaRS capable of recognizing these unique differences in the tRNA structure (Mukai et al., 2017).

tRNA^{Pyl}

Pyrrolysine (Pyl), the 22nd proteinogenic amino acid, was discovered in the active site of methylamine methyltransferase in the archaeal methanogen *Methanosarcina barkeri* (Hao et al., 2002). Pyl is genetically encoded via an in-frame amber (UAG) codon, which is normally used as a stop codon to terminate protein synthesis. This is possible due to an amber suppressor tRNA found in certain archaeal and bacterial species, pyrrolysine tRNA (tRNA^{Pyl}) (Srinivasan et al., 2002). tRNA^{Pyl} is aminoacylated by its cognate pyrrolysyl-tRNA synthetase (PylRS), a class II aaRS (Polycarpo et al., 2004). Unlike Sec, which requires a multi-step enzymatic process to be incorporated into a protein during translation (see section "tRNA^{Sec}"), incorporation of Pyl utilizes the same translational machinery as canonical tRNAs (Théobald-Dietrich et al., 2004; Zhang et al., 2005; Longstaff et al., 2007). The PylRS-tRNA^{Pyl} pair has been studied extensively; it is frequently utilized as a tool for genetic code expansion due to its ability to charge a wide variety of non-canonical amino acids (ncAAs) as well as its orthogonality in both bacterial and eukaryotic hosts (Wan et al., 2014; Tharp et al., 2018).

PylRS is typically composed of two domains: a CTD catalytic domain (PylSc) and an NTD (PylSn) (**Figure 7A**) (Herring et al., 2007a). The organization of these domains varies between organisms. In species from the archaeal genus *Methanosarcina*, PylRS is encoded as a single protein featuring both an NTD and CTD connected with a linker (Herring et al., 2007a). On the other hand, Pyl-utilizing bacteria such as *Desulfitobacterium hafniense* encode two individual proteins, PylSc and PylSn, for each domain (Nozawa et al., 2009). Finally, seventh-order methanogens such as *Methanomethylophilus alvus* encode a protein homologous to PylSc, but no homolog of PylSn exists in these archaea (Borrel et al., 2014). In general, PylSc is responsible for catalyzing the aminoacylation of tRNA^{Pyl}, while PylSn forms additional contacts with the tRNA (**Figure 7B**) (Herring et al., 2007a; Nozawa et al., 2009; Suzuki et al., 2017). Regardless of the domain structure of the enzyme, tRNA^{Pyl} structure and its interaction with PylRS varies from canonical tRNAs, and at the same time vary from one another.

Unlike previously mentioned tRNAs, most tRNA^{Pyl} species characterized to date have the canonical 7/5 tRNA structure, which allows translation with the same machinery as canonical tRNAs. Crystal structures as well as structure mapping and melting curve assays show that tRNA^{Pyl} adopts a tertiary conformation similar to the canonical L-shape (Théobald-Dietrich et al., 2004; Nozawa et al., 2009). The distinguishing features of tRNA^{Pyl} are the three-nucleotide variable arm, an elongated anticodon stem (from 5 to 6 bp), and a CUA anticodon. More specifically, the universal tRNA^{Pyl} identity elements are the discriminator base G73, and the first bp in the acceptor stem G1:C72 (Ambrogelly et al., 2007; Herring et al., 2007b). However, since these are also identity elements of many other tRNAs (Giegé et al., 1998; Giegé and Frugier, 2000–2013), additional identity elements are necessary for PylRS to distinguish tRNA^{Pyl} from the canonical tRNAs. These additional identity elements can differ for each PylRS-tRNA^{Pyl} pair and therefore will be explored in more detail below.

The *M. barkeri* tRNA^{Pyl} (*Mb* tRNA^{Pyl}) contains the above-mentioned features of tRNA^{Pyl} with a 6 bp anticodon stem (**Figure 8A**). However, it contains some additional features that differ from canonical tRNAs and distinguishes it from other tRNA^{Pyl} species. Canonical tRNAs contain two nucleotides between the acceptor stem and D-stem, while *Mb* tRNA^{Pyl} only has one. However, the connecting nucleotide is a U, consistent with the highly conserved U8 in canonical tRNAs. Furthermore, the D-loop is small, with only five nucleotides, and lacking the widely conserved G18, G19 sequence motif. Since the D- and T-loop are known to interact with each other, it follows that the T-loop is missing the corresponding U54, Ψ55, and C56 sequence. The absence of G19 and C56 (which forms a tertiary interaction in canonical tRNAs) indicates that an unusual interaction occurs between the D- and T-loops in *Mb* tRNA^{Pyl}. Details on the identity elements of *Mb* tRNA^{Pyl} were elucidated by screening its amber suppression efficiency (Ambrogelly et al., 2007). This study revealed that the nucleotides adjacent to the anticodon U33 and A37, and the T-stem bp G51:C63 are identity elements. Mutation of these identity elements significantly decreased the binding of *Mb* tRNA^{Pyl} to *Mb* PylRS in addition to their suppression efficiency. Furthermore, transplanting these identity elements into bovine mitochondrial tRNA^{Ser} yielded an active chimeric tRNA that could be aminoacylated by *Mb* PylRS both *in vitro* and *in vivo*.

The crystal structure of *M. mazei* tRNA^{Pyl} (*Mm* tRNA^{Pyl}) in complex with *M. mazei* PylSn (NTD) revealed the importance of the small, three-nucleotide variable arm. A tight interaction is formed between *Mm* PylSn and the variable arm of tRNA^{Pyl} (Suzuki et al., 2017). As previously mentioned, the small

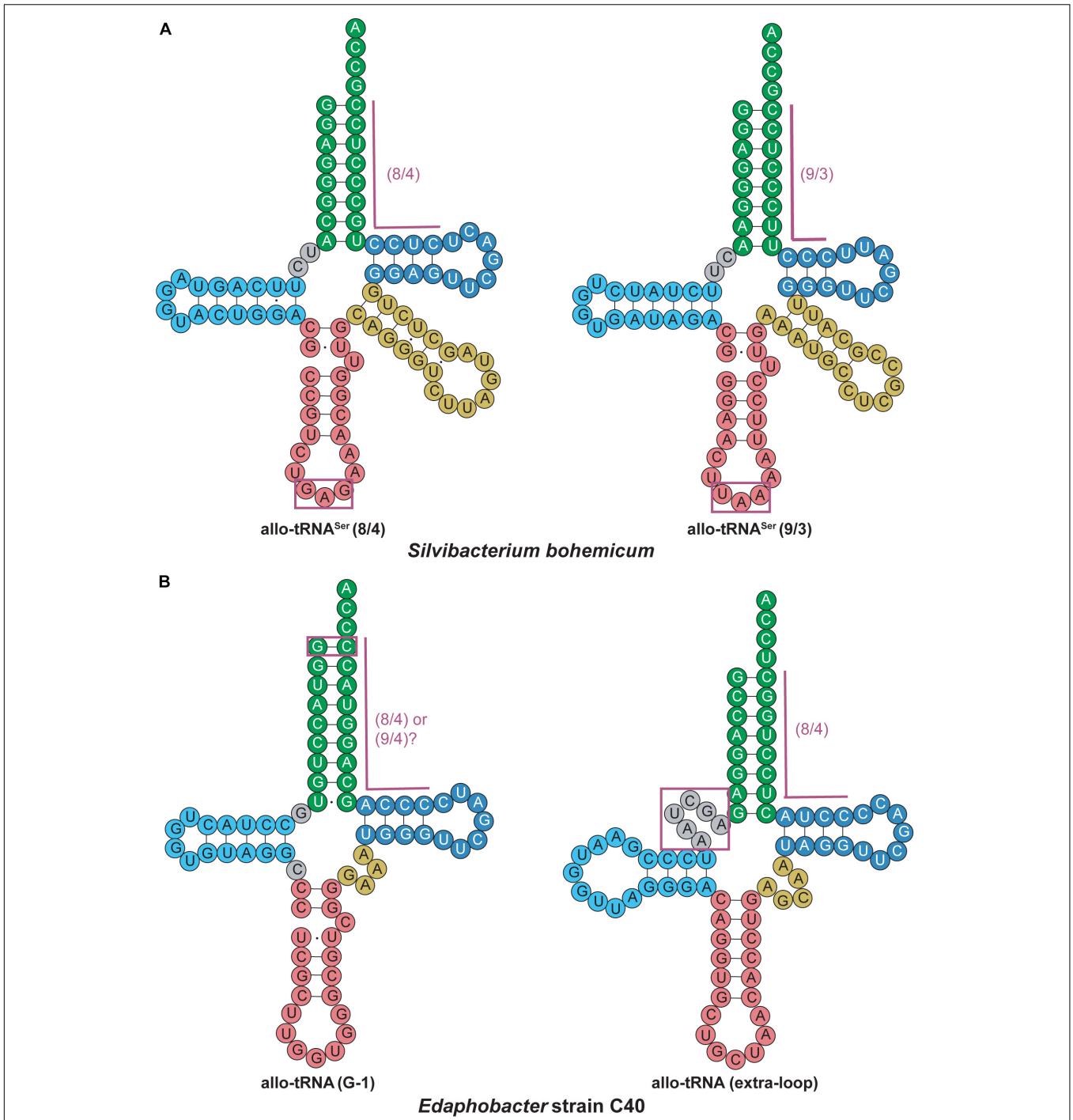


FIGURE 6 | Allo-tRNA secondary structures for **(A)** *Silvibacterium bohemicum* and **(B)** *Edaphobacter* strain C40 have unique features from canonical tRNAs. **(A)** Both 8/4 and 9/3 structures are observed for allo-tRNA^{Ser} with anticodons (boxed in magenta) that do not correlate with decoding of the Ser amino acid. **(B)** Two completely different tRNA secondary structures are proposed with either a G at the –1 position to form an additional base pair (boxed in magenta) providing either a 8/4 or 9/4 structure and a tRNA structure with an additional fifth loop between the acceptor arm and D-arm (gray boxed in magenta).

variable arm is a unique feature of tRNA^{Pyl}, as the variable arms of canonical tRNAs typically have 4–5 nucleotides for class I tRNAs, or greater than 10 nucleotides in the case of class II tRNAs (Sprinzl et al., 1998). Therefore, PylSn discriminates against canonical tRNAs based on the size of their variable arm. Addition of a fourth nucleotide to the variable arm of *Mb* tRNA^{Pyl} significantly decreases its suppression efficiency, providing further evidence that the interaction

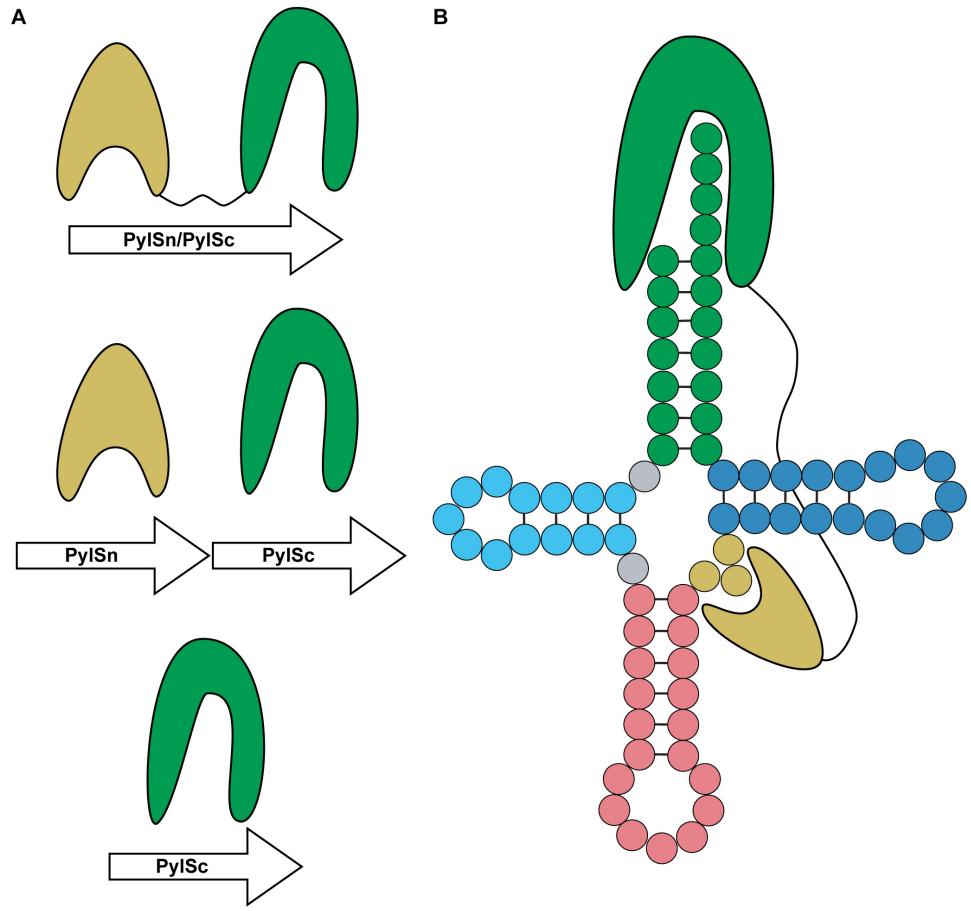


FIGURE 7 | Domain organization and binding mode of PylRS. **(A)** PylRS is composed of two domains, an N-terminal domain (PylSn) and a catalytic domain (PylSc). PylRS is either composed of a fusion of these two domains, two standalone proteins, or as a lone PylSc. **(B)** PylSc interacts with the acceptor stem and catalyzes the aminoacylation of $tRNA^{Pyl}$. PylSn forms a tight interaction with the variable arm.

between PylRS and the variable arm is critical for aminoacylation (Ambrogelly et al., 2007).

Although there are several differences in nucleotide sequences, the secondary structure of *Mb* $tRNA^{Pyl}$ is quite similar to the homologous tRNA from *D. hafniense*. Like *Mb* $tRNA^{Pyl}$, *Dh* $tRNA^{Pyl}$ has an elongated anticodon stem, shortened D-loop, small variable arm, and lacks the conserved nucleotide sequences G18, G19, and TΨC (**Figure 8B**). However, *Dh* $tRNA^{Pyl}$ is unique from $tRNA^{Pyl}$ from *methanosarciniae* in that the single nucleotide separating the acceptor and D-stem is G8 as opposed to U8. In canonical tRNAs, this position is widely conserved as U8, which stabilizes tertiary structure through base pairing with A14. Thus, the absence of U8 in *Dh* $tRNA^{Pyl}$ abolishes the highly conserved U8:A14 bp (Herring et al., 2007b; Nozawa et al., 2009).

The crystal structure of the *D. hafniense* PylSc in complex with *Dh* $tRNA^{Pyl}$ shows that the change of U8 to G8 allows an unusual interaction to occur between the D- and T-loop, wherein G13 interacts with C55 to stabilize the tertiary conformation of the tRNA (Nozawa et al., 2009). This also enables G8 to serve as an identity element for the interaction with PylSc, specifically through interaction with residues Arg140, Arg144, and Glu145 (Herring et al., 2007b; Nozawa et al., 2009). Despite these differences, *Dh* $tRNA^{Pyl}$ folds into an L-shape similar to canonical tRNAs (**Figure 8B**), with a compact core that is accommodated by the PylSc active site.

Structural and biochemical data on the interaction between *Dh* $tRNA^{Pyl}$ and *Dh* PylSc have revealed several tRNA identity elements (**Figure 8B**). In addition to the universal $tRNA^{Pyl}$ identity elements, a direct interaction occurs between *Dh* PylSc and the D-stem base pairs G10:C25 and A11:U24, as well as the previously mentioned G8 (Herring et al., 2007b; Nozawa et al., 2009). Although *in vitro* aminoacylation assays indicate that the nucleotides flanking the anticodon U33 and A37 are identity elements for *Mb* PylRS (Ambrogelly et al., 2007), *Dh* PylSc and PylSn do not directly interact with these residues (Nozawa et al., 2009; Jiang and Krzycki, 2012). Furthermore, while the anticodon is normally a tRNA identity element, *Dh* PylSc is found not to interact with the anticodon, which is also the case for all other characterized PylRS-$tRNA^{Pyl}$ pairs (Ambrogelly et al., 2007; Herring et al., 2007b; Nozawa et al., 2009). This desirable trait allowed for general codon reassignment, and thus opened the door for synthetic biologists to incorporate multiple

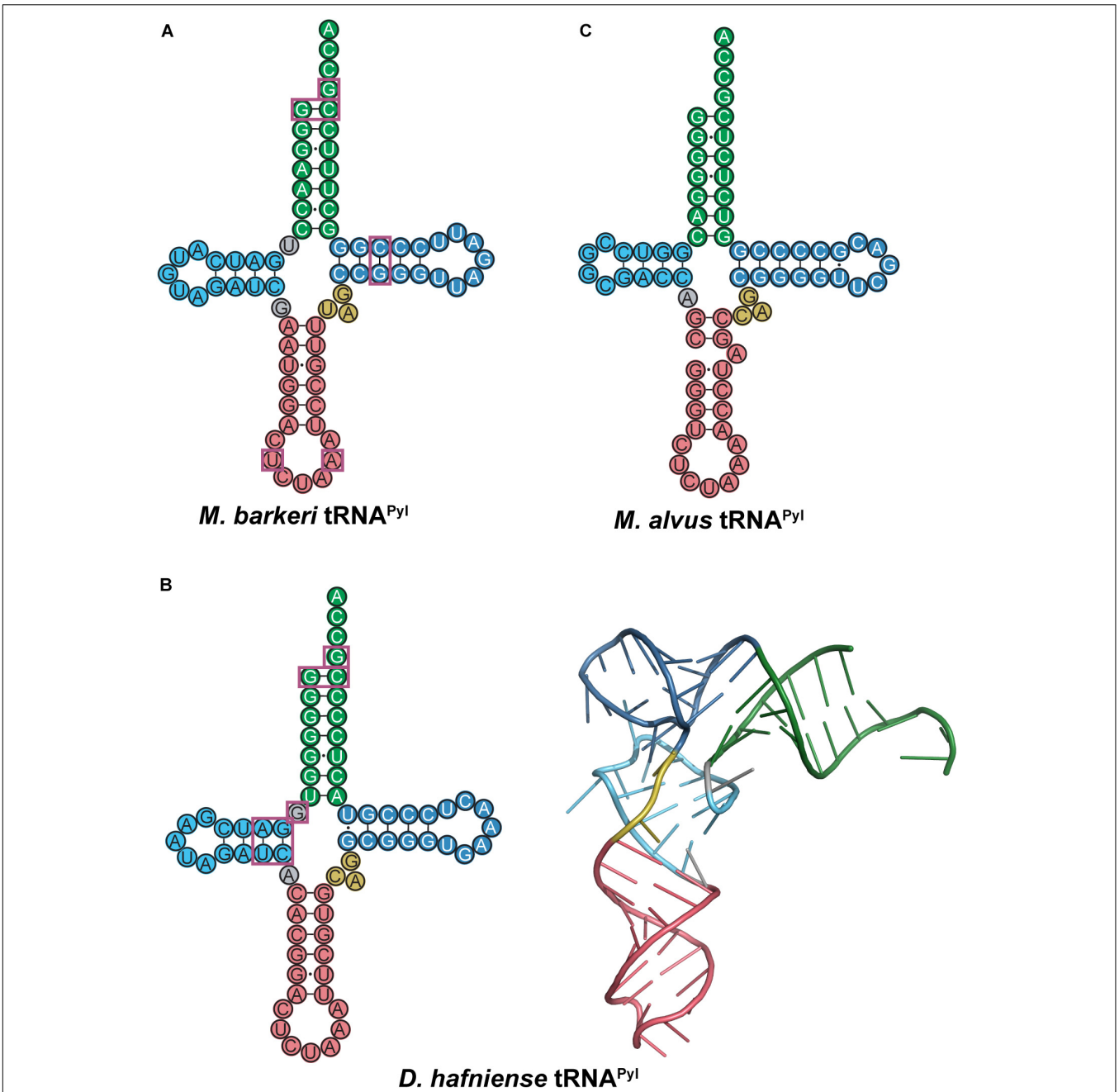


FIGURE 8 | Cloverleaf structures of tRNA^{Pyl} from **(A)** *M. barkeri*, **(B)** *D. hafniense*, and **(C)** *M. alvus*. Identity elements for each tRNA^{Pyl} are highlighted by magenta boxes. The crystal structure of *D. hafniense* tRNA^{Pyl} (PDB ID: 2ZNI; Nozawa et al., 2009) is also shown in **(B)**.

ncAAs into a single protein using different PylRS-tRNA^{Pyl} pairs (Wan et al., 2010; Chin, 2017). Ultimately, *Dh* PylSc binds to tRNA^{Pyl} through contacts with the acceptor and D-stem, and has no direct contact with the anticodon stem, variable loop, or T-stem (Nozawa et al., 2009).

The PylRS-tRNA^{Pyl} pair in the seventh-order methanogen *M. alvus* has recently been explored as an additional tool for genetic code expansion with advantages over its previously studied counterparts (Meineke et al., 2018; Willis and Chin, 2018; Yamaguchi et al., 2018; Beránek et al., 2019; Dunkelmann et al., 2020; Seki et al., 2020). *Ma* tRNA^{Pyl} has many unusual features that distinguish it from canonical tRNAs as well as previously characterized tRNA^{Pyl} (**Figure 8C**). The anticodon stem of *Ma* tRNA^{Pyl} features 6 bp in the anticodon stem like other tRNA^{Pyl}, but the stem is broken by an unpaired adenosine on the 3′ side of the stem. Other seventh order methanogens such as *Methanomassiliicoccus intestinalis* and *Methanomassiliicoccus lumenyensis* tRNA^{Pyl} feature larger breaks that form small loops

within the anticodon stem (Borrel et al., 2014). Also, *Ma* tRNA^{Pyl} does not have a nucleotide separating the acceptor and D-stem of the tRNA. This differs considerably from canonical tRNAs as well as from tRNA^{Pyl} species previously mentioned. An additional difference of *M. alvus* tRNA^{Pyl} is the four nucleotide D-loop (instead of five observed in the other tRNA^{Pyl} discussed).

On the surface, the break in the base pairing of the anticodon stem as well as the lack of a connecting base between the acceptor and D-stem profile as potential identity elements for *Ma* tRNA^{Pyl}. Interestingly, deletion of the unpaired nucleotide in the anticodon stem did not significantly alter the translation efficiency of *Ma* PylRS-tRNA^{Pyl} in a cell-free translation system (Yamaguchi et al., 2018). Insertion of a C or U between the acceptor and D-stem (position 8) moderately decreased translation, but inserting an A or G had no effect (Yamaguchi et al., 2018). This indicates that the absence of a base in this position may not be an identity element for *Ma* tRNA^{Pyl}. Therefore, in this system, the functional role, if any exists, of these unique features of *Ma* tRNA^{Pyl} is unclear.

Unlike *M. barkeri* and *D. hafniense*, *M. alvus* does not encode a protein homologous to PylSn, either as a standalone protein or a fusion to PylSc. PylSn binds tightly to the variable loop of tRNA^{Pyl} (Suzuki et al., 2017) and is essential for *in vivo* aminoacylation of *Dh* tRNA^{Pyl} (Herring et al., 2007a). However, *Ma* PylRS is highly active toward its cognate tRNA^{Pyl} even though it does not feature PylSn. Despite significant structural differences between *Ma* and *Mm* tRNA^{Pyl}, *Ma* tRNA^{Pyl} can serve as a substrate for both PylRS enzymes (Yamaguchi et al., 2018). However, lengthening the variable arm of *Ma* tRNA^{Pyl} prevents aminoacylation by *Mm* PylRS, due to steric constraints between PylSn and the enlarged variable arm as discussed earlier (Suzuki et al., 2017). Since *Ma* PylRS does not have a PylSn to interact with the variable arm, it still readily aminoacylates the tRNA despite the larger variable arm (Willis and Chin, 2018).

MITOCHONDRIAL tRNAs

Mitochondria are responsible for energy production in eukaryotic cells. As a semi-autonomous organelle descendent from bacteria, mitochondria have their own genome. Mitochondrial genomes not only encode proteins essential for energy production, but also encode parts of the translation machinery, including mitochondrial tRNAs (mt-tRNAs) (Gray et al., 1999). The number of mt-tRNA genes encoded in the mitochondrion varies between organisms. In most cases, mitochondria import additional, nuclear-encoded tRNA and proteins that are required for translation (Alfonzo and Söll, 2009; Dudek et al., 2013; Salinas-Giegé et al., 2015). In addition to mt-tRNAs, mitochondrial translation occurs via a specialized translation machinery, including mitoribosomes and mitochondrial initiation and elongation factors (Salinas-Giegé et al., 2015; D' Souza and Minczuk, 2018). Although canonical tRNAs require conserved structural elements for proper folding, many mt-tRNAs possess highly unusual secondary structures that deviate greatly from canonical tRNAs.

Most tRNAs found in organisms are type 0 tRNAs, which have a conserved cloverleaf structure and fold into a tertiary L-shape due to interactions between the D- and T-loops. On the other hand, mt-tRNAs can be classified into three types based off of their secondary structure (Watanabe, 2010; Suzuki et al., 2011). Type I mt-tRNAs have an atypical anticodon stem. This includes mt-tRNA$^{Ser}{}_{UCN}$, which has 6 bp in the anticodon stem instead of the typical 5 bp (**Figure 9A**). Mammalian mt-tRNA$^{Ser}{}_{UCN}$ has many similarities with *Mb* tRNA^{Pyl} that are not seen in most characterized tRNAs. Both tRNA structures have only a single nucleotide separating the acceptor and D-stem, have smaller than normal D-loops, elongated anticodon stems, and variable arms consisting of only three nucleotides. However, unlike *Mb* tRNA^{Pyl}, mt-tRNA$^{Ser}{}_{UCN}$ features the G18, G19, and TΨC sequences in its D- and T-loops (**Figures 7A**, **9A**). Furthermore, type I mt-tRNAs have the L-shaped tertiary structure which resembles that of canonical tRNA (Watanabe et al., 1994a; Hayashi et al., 1998; Mustoe et al., 2015). The most common mt-tRNAs, type II mt-tRNAs lack interaction between the D- and T-loops (**Figure 9B**). In these mt-tRNAs, the canonical G18, G19, and TΨC sequence motifs in the D- and T-loop, respectively, are not conserved. Instead, interactions occur between the D-loop and the variable stem to stabilize the mt-tRNA tertiary structure (Wakita et al., 1994; Messmer et al., 2009; Watanabe, 2010). Finally, type III mt-tRNAs lack a D-stem; they are the only mammalian mt-tRNAs without the canonical cloverleaf structure. An example of a type III mt-tRNA is mt-tRNA$^{Ser}{}_{AGY}$ (**Figure 9C**). Despite lacking a D-stem, this mt-tRNA is functional *in vitro* and adopts a conformation that is suitable for the ribosome (Hanada et al., 2001; Frazer-Abel and Hagerman, 2008).

The interaction between mt-tRNAs and mt-aaRSs is not well-understood, as there is limited structural information available on the binding of mt-tRNAs to their cognate aaRSs. However, identity elements have been established for mammalian mt-tRNA^{Tyr} (Bonnefond et al., 2005, 2007), mt-tRNA^{Leu} (Sohm et al., 2003, 2004), mt-tRNA^{Ser} (AGY and UCN) (Shimada et al., 2001), mt-tRNA^{Ala} (Lovato et al., 2001), and mt-tRNA^{Asp} (Fender et al., 2006; Neuenfeldt et al., 2013) species. Interestingly, mammalian mt-aaRSs appear to have evolved relaxed specificity for their cognate tRNAs. Specifically, bovine mt-aaRSs have been shown to acylate the corresponding *E. coli* tRNAs, while the *E. coli* aaRSs cannot acylate the equivalent mt-tRNA (Kumazawa et al., 1991; Watanabe et al., 1994a). Mt-SerRS is even more promiscuous, as it serylates several *E. coli* tRNAs as well as mt-tRNA^{Gln} (Shimada et al., 2001). Mt-tRNA^{Gln} is an orphan tRNA; in addition to being a substrate for mt-SerRS, mt-tRNA^{Gln} is also mischarged by mt-GluRS to Glu-tRNA^{Gln}, which is subsequently transamidated to Gln-tRNA^{Gln} (Nagao et al., 2009). In canonical aaRS-tRNA pairs, the first bp is a common identity element. However, in the more promiscuous human mt-TyrRS it was found not to recognize the first bp of mt-tRNA^{Tyr} (Bonnefond et al., 2005, 2007). Taken together, these findings indicate that in mammals, mt-aaRSs do not strongly discriminate against non-cognate tRNAs. This apparent lack of specificity may be attributed to the high substrate diversity of mt-tRNAs, or possibly a lack of evolutionary pressure due to the smaller pool of mt-tRNAs present in the cell.

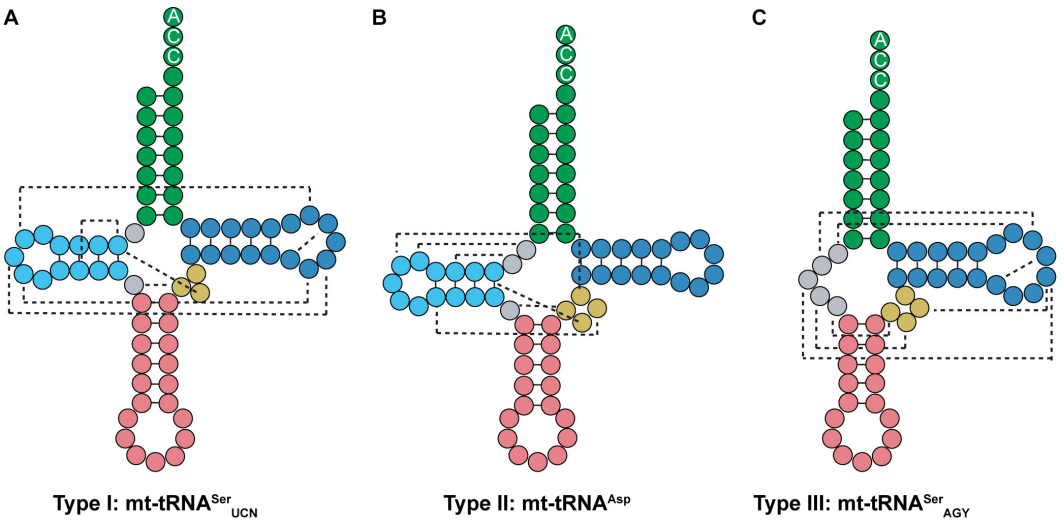


FIGURE 9 | Mammalian mt-tRNA can be classified into three types. **(A)** Type I mt-tRNA, represented by mt-tRNA$^{Ser}_{UCN}$, shares similarities with canonical tRNAs, featuring the same conserved D- and T-loop interactions. **(B)** Type II mt-tRNA, represented by mt-tRNA^{Asp}, is the most commonly found mt-tRNA. **(C)** Type III mt-tRNA, represented by mt-tRNA$^{Ser}_{AGY}$, do not have a D-arm. Instead, the connecting region between the acceptor and anticodon stem interacts with the variable and T-loop to promote folding.

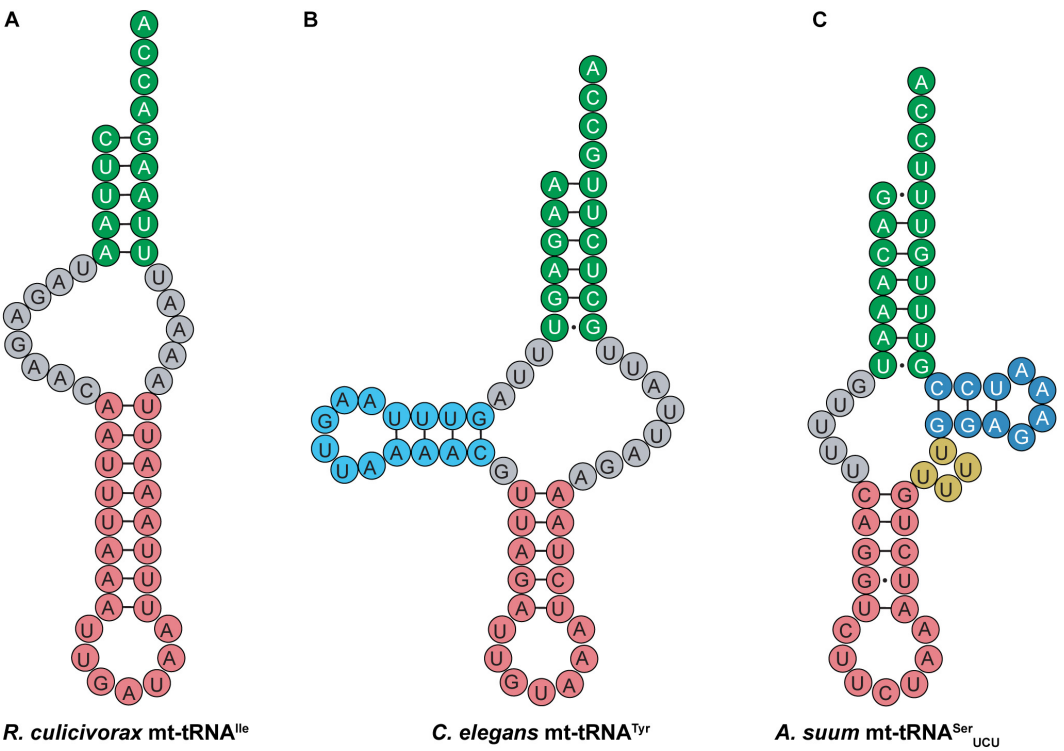


FIGURE 10 | Nematode mt-tRNAs have diverse and highly unusual secondary structures. Examples of these abnormal mt-tRNA structures are shown here. **(A)** *R. culicivorax* mt-tRNA^{Ile} has no D- or T-arm. **(B)** *C. elegans* mt-tRNA^{Tyr} has a D-arm, but no T-arm. **(C)** *A. suum* mt-tRNA$^{Ser}_{UCU}$ has a short T-arm and a variable loop, but no D-arm.

Like their mammalian counterparts, nematode mt-tRNAs have unusual structural features that are distinct from canonical tRNAs. Nematodes encode short mt-tRNAs with diverse cloverleaf structures. In addition to nematodes, highly truncated mt-tRNAs have been found in the genomes of mites and arachnids, where short tRNAs missing both the D- and T-arms

have been identified (**Figure 10A**) (Klimov and Oconnor, 2009; Jühling et al., 2012; Palopoli et al., 2014; Wende et al., 2014; Salinas-Giegé et al., 2015; Juhling et al., 2018). Despite greatly deviating from the canonical tRNA cloverleaf structure, evidence suggests that mt-tRNAs lacking one or both sidearms can still interact with tRNA processing enzymes such as CCA-adding enzyme, and are aminoacylated by their cognate synthetases (Wolfson et al., 1999; Tomari et al., 2002; Wende et al., 2014). Most nematode mt-tRNAs lack the entire T-arm (**Figure 10B**), except mt-tRNA^{Ser}, which have a short T-arm consisting of only 10–13 nucleotides (Wolstenholme et al., 1987; Okimoto and Wolstenholme, 1990; Watanabe et al., 1994b). For instance, *Ascaris suum* mt-tRNA$^{Ser}{}_{UCU}$ has a short, 10-nucleotide T-arm, and completely lacks a D-arm (**Figure 10C**) (Ohtsuki et al., 2002). The short T-arm in *A. suum* tRNA$^{Ser}{}_{UCU}$ as well as the connector region which replaces the D-arm, confer flexibility to the mt-tRNA. This flexibility allows the mt-tRNA to adopt a less rigid tertiary structure than the canonical L-shape, enabling the distance between the 3′ end of the tRNA and the anticodon to properly adjust to fit into the ribosome (Ohtsuki et al., 2002). Similar findings of flexible tertiary structure have been reported for mammalian mt-tRNA$^{Ser}{}_{AGY}$ (**Figure 9C**), which also lacks a D-arm (Steinberg and Cedergren, 1994; Frazer-Abel and Hagerman, 2008). In addition to these observations, recent structural data also indicate that *R. culicivorax* mt-tRNA^{Ile} (**Figure 10A**) folds into a stable, boomerang-shaped tertiary structure (Juhling et al., 2018). Further, the 3D structure reveals that the distance between the anticodon and 3′ end of *R. culicivorax* mt-tRNA^{Ile} is comparable to that of canonical, cytosolic tRNA^{Phe} from *Saccharomyces cerevisiae* (Juhling et al., 2018). Thus, evidence suggests that the D- and T-arms are not required for tRNA to fold into a tertiary conformation suitable for enzymatic activity, and the flexibility of these truncated mt-tRNAs helps to achieve functionality.

In addition to the flexible tertiary structure discussed above, post-transcriptional modifications appear to play an important role in stabilizing mt-tRNAs. Many unmodified mt-tRNAs will not fold properly, but proper modification allows folding and interaction with tRNA processing enzymes to occur (Lorenz et al., 2017). For instance, 1-methyl adenosine at position 9 (m^1A9) is found in many mt-tRNA species, including those lacking one or both sidearms, and this modification is important for proper cloverleaf folding to occur (Helm et al., 1999; Sakurai et al., 2005; Lorenz et al., 2017; Juhling et al., 2018). Nematode mt-tRNA lacking the D-arm possess m^1A9 as well as several pseudouridine in the acceptor and anticodon stem (Sakurai et al., 2005; Lorenz et al., 2017; Juhling et al., 2018). In these mt-tRNA, m^1A9 facilitates aminoacylation and interaction with EF-Tu (Sakurai et al., 2005), while pseudouridine likely supports tRNA stability (Lorenz et al., 2017). Ultimately, post-translational modification appears to play an important role in facilitating mt-tRNA activity and stability, including the truncated mt-tRNAs that lack a canonical cloverleaf structure.

Deviations from the standard genetic code have been reported in the mitochondria of green plant algae from the phylum Chlorophyta (Noutahi et al., 2019). Similar to the pyrrolysine incorporation system in archaea and bacteria, the stop codons UAG and UGA are reassigned to sense codons in some Chlorophyta. In several of these species, UAG is reassigned to Ala or Leu, and UGA is reassigned to Trp (Fučíková et al., 2014). These mt-tRNAs with CUA or UCA anticodons, feature identity elements of tRNA^{Ala}, tRNA^{Leu} or tRNA^{Trp} species, thus allowing for stop codon suppression and elongation with the corresponding amino acid.

Recent evidence suggests that in addition to stop codon reassignment, sense codons may also be reassigned in green algae. AGG, which is normally an Arg codon, appears to be reassigned in Sphaeropleales (Noutahi et al., 2019). In these green algae, mt-tRNAs with a CCU anticodon do not share any structural or sequence similarities with canonical tRNA^{Arg}. Analysis of the mt-tRNA$_{CCU}$ secondary structures reveals that many of these mt-tRNAs instead share identity elements with Chlorophyta mt-tRNA$^{Ala}{}_{UGC}$, including the invariant G3:U70 pair and the discriminator base A73.

Sense codon reassignment has also been observed in *S. cerevisiae* mitochondria as well as that of *Ashbya gossypii*, a relative of yeast (Alfonzo and Söll, 2009; Su et al., 2011; Ling et al., 2014, 2015). In *S. cerevisiae* mitochondria, CUN codons, which normally decode Leu, are reassigned to Thr. This reassignment is facilitated by an unusual mt-tRNA$^{Thr}{}_{UAG}$ that features an enlarged 8-nt anticodon loop and a UAG anticodon. *S. cerevisiae* lacks mt-tRNA$^{Leu}{}_{UAG}$, thus allowing complete reassignment of the CUN codon from Leu to Thr. Interestingly, phylogenetic and mutational analyses of yeast mt-tRNAs indicate that mt-tRNA$^{Thr}{}_{UAG}$ evolved from mt-tRNA$^{His}{}_{GUG}$ as opposed to mt-tRNA$^{Leu}{}_{UAG}$ or mt-tRNA$^{Thr}{}_{UGU}$ (Su et al., 2011). In *A. gossypii,* the codons CUU and CUA are reassigned to decode Ala. Like the Chlorophyta mt-tRNA$_{CCU}$ described above, *A. gossypii* mt-tRNA$^{Ala}{}_{UAG}$ features the strictly conserved Ala identity element G3:U70. This bp is critical for recognition of mt-tRNA$^{Ala}{}_{UAG}$ by AlaRS, with a G3A mutation abolishing aminoacylation (Ling et al., 2014). The observation that codon reassignment occurs in mitochondria across kingdoms underscores the dynamic nature of the mitochondrial genome.

OUTLOOK

This evidence shows that not all tRNAs have the canonical 7/5 structure that was originally portrayed. The unique structures found in these non-canonical tRNAs appear to be a result of their necessary function and the enzymes that they interact with. In some cases (tRNA^{Sec}), these details are well-understood while in others (allo-tRNAs) it is a bit more speculative. Despite deviating from the canonical structure, majority of the tRNAs presented in this review have been found to be functional in translation. While tRNA^{Sec} has specialized translational requirements for Sec to be incorporated into proteins, tRNA^{Pyl} utilizes the same translational machinery as canonical tRNAs (Théobald-Dietrich et al., 2004; Zhang et al., 2005; Longstaff et al., 2007). In mitochondria, highly unusual mt-tRNAs with diverse structures are used along with specialized mitochondrial translation machinery to translate proteins encoded by the mitochondrial genome (Gray et al., 1999). Taken together, these

observations clearly indicate that the canonical tRNA structure is not a prerequisite for translation, and it is evident that although canonical tRNAs are in the 7/5 structure, their translation systems can accommodate diverse structures including 8/4, 8/5, 9/3, and 9/4 structures.

The translational machinery has evolved to accept a wide variety of tRNA structures for efficient translation of proteins in the desired host. A significant amount of effort has been put forth to expand the genetic code, pushing the boundaries of what functionality can be incorporated into proteins. To that end, engineered aaRS-tRNA pairs have been utilized to incorporate numerous diverse ncAAs into proteins both *in vitro* and *in vivo*. In many cases, the host's translation machinery readily accepts foreign and modified tRNAs featuring diverse structures and charged with a ncAA (Chin, 2017). This plasticity may indicate a lack of evolutionary pressure to discriminate against unknown or unusual tRNAs that are rarely, if ever, encountered by the host cell. Thus, it is plausible that the unusual structures of specialized or non-canonical tRNAs such as the ones described in this review are made possible by a lack of evolutionary pressure to maintain the canonical structure. An alternative possibility is that many of these non-canonical tRNAs originated from an ancient, more diverse genetic code, and because of their specialized and infrequent usage, they were never pressured to evolve into a canonical tRNA structure. In either case, if deviations from the canonical tRNA structure are well-tolerated by the aminoacyl-synthetase and translation machinery, mutations or structural changes to the tRNA can potentially occur without consequence and lead to polymorphisms over time. This can be seen in mitochondria. Highly variable mt-tRNAs are well-known to be susceptible to mutations, and while mt-tRNA mutations to critical nucleotides can cause diseases, neutral or slightly deleterious polymorphisms frequently occur and are inconsequential (Lynch, 1996; Wittenhagen and Kelley, 2003; Yarham et al., 2010). Ultimately, despite their many differences from the canonical tRNA structure, non-canonical tRNAs are readily utilized in translation and enable the cell to produce proteins that are, in many cases, essential for survival (Longstaff et al., 2007).

AUTHOR CONTRIBUTIONS

NK and JF wrote the manuscript. DS edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

Work in the authors' laboratory was supported by the National Institute of General Medical Sciences (R35GM122560) and the DOE Office of Basic Energy Sciences (DE-FG02-98ER20311).

ACKNOWLEDGMENTS

We thank Oscar Vargas-Rodriguez for discussions.

REFERENCES

Alfonzo, J. D., and Söll, D. (2009). Mitochondrial tRNA import–the challenge to understand has just begun. *Biol. Chem.* 390, 717–722.

Amberg, R., Mizutani, T., Wu, X. Q., and Gross, H. J. (1996). Selenocysteine synthesis in mammalia: an identity switch from tRNASer to tRNASec. *J. Mol. Biol.* 263, 8–19. doi: 10.1006/jmbi.1996.0552

Ambrogelly, A., Gundllapalli, S., Herring, S., Polycarpo, C., Frauer, C., and Söll, D. (2007). Pyrrolysine is not hardwired for cotranslational insertion at UAG codons. *Proc. Natl. Acad. Sci. U.S.A.* 104, 3141–3146. doi: 10.1073/pnas.0611634104

Baron, C., and Böck, A. (1991). The length of the aminoacyl-acceptor stem of the selenocysteine-specific tRNASec of *Escherichia coli* is the determinant for binding to elongation factors SELB or Tu. *J. Biol. Chem.* 266, 20375–20379.

Baron, C., Westhof, E., Bock, A., and Giegé, R. (1993). Solution structure of selenocysteine-inserting tRNASec from *Escherichia coli*. Comparison with canonical tRNASer. *J. Mol. Biol.* 231, 274–292. doi: 10.1006/jmbi.1993.1282

Beránek, V., Willis, J. C. W., and Chin, J. W. (2019). An evolved *Methanomethylophilus alvus* pyrrolysyl-tRNA synthetase/tRNA pair is highly active and orthogonal in mammalian cells. *Biochemistry* 58, 387–390. doi: 10.1021/acs.biochem.8b00808

Bonnefond, L., Frugier, M., Giegé, R., and Rudinger-Thirion, J. (2005). Human mitochondrial TyrRS disobeys the tyrosine identity rules. *RNA* 11, 558–562. doi: 10.1261/rna.7246805

Bonnefond, L., Frugier, M., Touzé, E., Lorber, B., Florentz, C., Giegé, R., et al. (2007). Crystal structure of human mitochondrial tyrosyl-tRNA synthetase reveals common and idiosyncratic features. *Structure* 15, 1505–1516. doi: 10.1016/j.str.2007.09.018

Borrel, G., Gaci, N., Peyret, P., O' Toole, P. W., Gribaldo, S., and Brugère, J.-F. (2014). Unique characteristics of the pyrrolysine system in the 7th order of methanogens: implications for the evolution of a genetic code expansion cassette. *Archaea* 2014:374146.

Breitschopf, K., and Gross, H. J. (1994). The exchange of the discriminator base A73 for G is alone sufficient to convert human tRNALeu into a serine-acceptor *in vitro*. *EMBO J.* 13, 3166–3169. doi: 10.1002/j.1460-2075.1994.tb06615.x

Chiba, S., Itoh, Y., Sekine, S., and Yokoyama, S. (2010). Structural basis for the major role of O-phosphoseryl-tRNA kinase in the UGA-specific encoding of selenocysteine. *Mol. Cell* 39, 410–420. doi: 10.1016/j.molcel.2010.07.018

Chin, J. W. (2017). Expanding and reprogramming the genetic code. *Nature* 550, 53–60. doi: 10.1038/nature24031

Cone, J. E., Del Rio, R. M., Davis, J. N., and Stadtman, T. C. (1976). Chemical characterization of the selenoprotein component of clostridial glycine reductase: identification of selenocysteine as the organoselenium moiety. *Proc. Natl. Acad. Sci. U.S.A.* 73, 2659–2663. doi: 10.1073/pnas.73.8.2659

Cramer, F., Erdmann, V. A., Von Der Haar, F., and Schlimme, E. (1969). Structure and reactivity of tRNA. *J. Cell. Physiol.* 74(Suppl. 1), 163.

Crothers, D. M., Seno, T., and Söll, G. (1972). Is there a discriminator site in transfer RNA? *Proc. Natl. Acad. Sci. U.S.A.* 69, 3063–3067. doi: 10.1073/pnas.69.10.3063

D' Souza, A. R., and Minczuk, M. (2018). Mitochondrial transcription and translation: overview. *Essays Biochem.* 62, 309–320. doi: 10.1042/ebc20170102

Dudek, J., Rehling, P., and Van Der Laan, M. (2013). Mitochondrial protein import: common principles and physiological networks. *Biochim. Biophys. Acta* 1833, 274–285. doi: 10.1016/j.bbamcr.2012.05.028

Dunkelmann, D. L., Willis, J. C. W., Beattie, A. T., and Chin, J. W. (2020). Engineered triply orthogonal pyrrolysyl-tRNA synthetase/tRNA pairs enable the genetic encoding of three distinct non-canonical amino acids. *Nat. Chem.* 12, 535–544. doi: 10.1038/s41557-020-0472-x

Fender, A., Sauter, C., Messmer, M., Pütz, J., Giegé, R., Florentz, C., et al. (2006). Loss of a primordial identity element for a mammalian mitochondrial

aminoacylation system. *J. Biol. Chem.* 281, 15980–15986. doi: 10.1074/jbc.m511633200

Forster, C., Ott, G., Forchhammer, K., and Sprinzl, M. (1990). Interaction of a selenocysteine-incorporating tRNA with elongation factor Tu from *E. coli*. *Nucleic Acids Res.* 18, 487–491. doi: 10.1093/nar/18.3.487

Frazer-Abel, A. A., and Hagerman, P. J. (2008). Core flexibility of a truncated metazoan mitochondrial tRNA. *Nucleic Acids Res.* 36, 5472–5481. doi: 10.1093/nar/gkn529

Fu, X., Crnkoviæ, A., Sevostyanova, A., and Söll, D. (2018). Designing seryl-tRNA synthetase for improved serylation of selenocysteine tRNAs. *FEBS Lett.* 592, 3759–3768. doi: 10.1002/1873-3468.13271

Fučíková, K., Lewis, P. O., González-Halphen, D., and Lewis, L. A. (2014). Gene arrangement convergence, diverse intron content, and genetic code modifications in mitochondrial genomes of sphaeropleales (chlorophyta). *Genome Biol. Evol.* 6, 2170–2180. doi: 10.1093/gbe/evu172

Giegé, R., and Frugier, M. (2000–2013). "Transfer RNA structure and identity," in *Madame Curie Bioscience Database [Internet]*, (Austin, TX: Landes Bioscience). Available online at: https://www.ncbi.nlm.nih.gov/books/NBK6236/

Giegé, R., Sissler, M., and Florentz, C. (1998). Universal rules and idiosyncratic features in tRNA identity. *Nucleic Acids Res.* 26, 5017–5035. doi: 10.1093/nar/26.22.5017

Gray, M. W., Burger, G., and Lang, B. F. (1999). Mitochondrial evolution. *Science* 283, 1476–1481.

Hanada, T., Suzuki, T., Yokogawa, T., Takemoto-Hori, C., Sprinzl, M., and Watanabe, K. (2001). Translation ability of mitochondrial tRNASer with unusual secondary structures in an in vitro translation system of bovine mitochondria. *Genes Cells* 6, 1019–1030.

Hao, B., Gong, W., Ferguson, T. K., James, C. M., Krzycki, J. A., and Chan, M. K. (2002). A new UAG-encoded residue in the structure of a methanogen methyltransferase. *Science* 296, 1462–1466. doi: 10.1126/science.1069556

Hayashi, I., Kawai, G., and Watanabe, K. (1998). Higher-order structure and thermal instability of bovine mitochondrial tRNAUGASer investigated by proton NMR spectroscopy. *J. Mol. Biol.* 284, 57–69. doi: 10.1006/jmbi.1998.2151

Heckl, M., Busch, K., and Gross, H. J. (1998). Minimal tRNASer and tRNASec substrates for human seryl-tRNA synthetase: contribution of tRNA domains to serylation and tertiary structure. *FEBS Lett.* 427, 315–319. doi: 10.1016/s0014-5793(98)00435-9

Helm, M., Giegé, R., and Florentz, C. (1999). A Watson-Crick base-pair-disrupting methyl group (m1A9) is sufficient for cloverleaf folding of human mitochondrial tRNALys. *Biochemistry* 38, 13338–13346. doi: 10.1021/bi991061g

Herring, S., Ambrogelly, A., Gundllapalli, S., O' Donoghue, P., Polycarpo, C. R., and Söll, D. (2007a). The amino-terminal domain of pyrrolysyl-tRNA synthetase is dispensable *in vitro* but required for *in vivo* activity. *FEBS Lett.* 581, 3197–3203. doi: 10.1016/j.febslet.2007.06.004

Herring, S., Ambrogelly, A., Polycarpo, C. R., and Söll, D. (2007b). Recognition of pyrrolysine tRNA by the *Desulfitobacterium hafniense* pyrrolysyl-tRNA synthetase. *Nucleic Acids Res.* 35, 1270–1278. doi: 10.1093/nar/gkl1151

Holley, R. W., Apgar, J., Everett, G. A., Madison, J. T., Marquisee, M., Merrill, S. H., et al. (1965). Structure of a ribonucleic acid. *Science* 147, 1462–1465.

Hou, Y. M. (1993). The tertiary structure of tRNA and the development of the genetic code. *Trends Biochem. Sci.* 18, 362–364. doi: 10.1016/0968-0004(93)90088-5

Hou, Y. M., and Schimmel, P. (1988). A simple structural feature is a major determinant of the identity of a transfer RNA. *Nature* 333, 140–145. doi: 10.1038/333140a0

Hubert, N., Sturchler, C., Westhof, E., Carbon, P., and Krol, A. (1998). The 9/4 secondary structure of eukaryotic selenocysteine tRNA: more pieces of evidence. *RNA* 4, 1029–1033. doi: 10.1017/s1355838298980888

Ioudovitch, A., and Steinberg, S. V. (1998). Modeling the tertiary interactions in the eukaryotic selenocysteine tRNA. *RNA* 4, 365–373.

Ishii, T. M., Kotlova, N., Tapsoba, F., and Steinberg, S. V. (2013). The long D-stem of the selenocysteine tRNA provides resilience at the expense of maximal function. *J Biol. Chem.* 288, 13337–13344. doi: 10.1074/jbc.m112.434704

Itoh, Y., Bröcker, M. J., Sekine, S., Hammond, G., Suetsugu, S., Söll, D., et al. (2013). Decameric SelA*tRNASec ring structure reveals mechanism of bacterial selenocysteine formation. *Science* 340, 75–78. doi: 10.1126/science.1229521

Itoh, Y., Chiba, S., Sekine, S., and Yokoyama, S. (2009). Crystal structure of human selenocysteine tRNA. *Nucleic Acids Res.* 37, 6259–6268. doi: 10.1093/nar/gkp648

Jiang, R., and Krzycki, J. A. (2012). PylSn and the homologous N-terminal domain of pyrrolysyl-tRNA synthetase bind the tRNA that is essential for the genetic encoding of pyrrolysine. *J Biol. Chem.* 287, 32738–32746. doi: 10.1074/jbc.m112.396754

Jühling, F., Pütz, J., Florentz, C., and Stadler, P. F. (2012). Armless mitochondrial tRNAs in Enoplea (Nematoda). *RNA Biol.* 9, 1161–1166. doi: 10.4161/rna.21630

Juhling, T., Duchardt-Ferner, E., Bonin, S., Wohnert, J., Putz, J., Florentz, C., et al. (2018). Small but large enough: structural properties of armless mitochondrial tRNAs from the nematode *Romanomermis culicivorax*. *Nucleic Acids Res.* 46, 9170–9180. doi: 10.1093/nar/gky593

Klimov, P. B., and Oconnor, B. M. (2009). Improved tRNA prediction in the American house dust mite reveals widespread occurrence of extremely short minimal tRNAs in acariform mites. *BMC Genomics* 10:598. doi: 10.1186/1471-2164-10-598

Komatsoulis, G. A., and Abelson, J. (1993). Recognition of tRNACys by *Escherichia coli* cysteinyl-tRNA synthetase. *Biochemistry* 32, 7435–7444. doi: 10.1021/bi00080a014

Kumazawa, Y., Himeno, H., Miura, K., and Watanabe, K. (1991). Unilateral aminoacylation specificity between bovine mitochondria and eubacteria. *J. Biochem.* 109, 421–427. doi: 10.1093/oxfordjournals.jbchem.a123397

Ling, J., Daoud, R., Lajoie, M. J., Church, G. M., Söll, D., and Lang, B. F. (2014). Natural reassignment of CUU and CUA sense codons to alanine in Ashbya mitochondria. *Nucleic Acids Res.* 42, 499–508. doi: 10.1093/nar/gkt842

Ling, J., O' Donoghue, P., and Söll, D. (2015). Genetic code flexibility in microorganisms: novel mechanisms and impact on physiology. *Nat. Rev. Microbiol.* 13, 707–721. doi: 10.1038/nrmicro3568

Liu, C., Sanders, J. M., Pascal, J. M., and Hou, Y. M. (2012). Adaptation to tRNA acceptor stem structure by flexible adjustment in the catalytic domain of class I tRNA synthetases. *RNA* 18, 213–221. doi: 10.1261/rna.029983.111

Longstaff, D. G., Blight, S. K., Zhang, L., Green-Church, K. B., and Krzycki, J. A. (2007). *In vivo* contextual requirements for UAG translation as pyrrolysine. *Mol. Microbiol.* 63, 229–241. doi: 10.1111/j.1365-2958.2006.05500.x

Lorenz, C., Lunse, C. E., and Morl, M. (2017). tRNA Modifications: impact on structure and thermal adaptation. *Biomolecules* 7:35. doi: 10.3390/biom7020035

Lovato, M. A., Chihade, J. W., and Schimmel, P. (2001). Translocation within the acceptor helix of a major tRNA identity determinant. *EMBO J.* 20, 4846–4853. doi: 10.1093/emboj/20.17.4846

Lynch, M. (1996). Mutation accumulation in transfer RNAs: molecular evidence for Muller's ratchet in mitochondrial genomes. *Mol. Biol. Evol.* 13, 209–220. doi: 10.1093/oxfordjournals.molbev.a025557

Lyons, S. M., Fay, M. M., and Ivanov, P. (2018). The role of RNA modifications in the regulation of tRNA cleavage. *FEBS Lett.* 592, 2828–2844. doi: 10.1002/1873-3468.13205

Mcclain, W. H., and Foss, K. (1988). Changing the identity of a tRNA by introducing a G-U wobble pair near the 3' acceptor end. *Science* 240, 793–796. doi: 10.1126/science.2452483

Meineke, B., Heimgärtner, J., Lafranchi, L., and Elsässer, S. J. (2018). *Methanomethylophilus alvus Mx1201* provides basis for mutual orthogonal pyrrolysyl tRNA/aminoacyl-tRNA synthetase pairs in mammalian cells. *ACS Chem. Biol.* 13, 3087–3096. doi: 10.1021/acschembio.8b00571

Messmer, M., Pütz, J., Suzuki, T., Suzuki, T., Sauter, C., Sissler, M., et al. (2009). Tertiary network in mammalian mitochondrial tRNAAsp revealed by solution probing and phylogeny. *Nucleic Acids Res.* 37, 6881–6895. doi: 10.1093/nar/gkp697

Mizutani, T., and Goto, C. (2000). Eukaryotic selenocysteine tRNA has the 9/4 secondary structure. *FEBS Lett.* 466, 359–362. doi: 10.1016/s0014-5793(00)01104-2

Mizutani, T., Kanaya, K., Ikeda, S., Fujiwara, T., Yamada, K., and Totsuka, T. (1998a). The dual identities of mammalian tRNASec for SerRS and selenocysteine synthase. *Mol. Biol. Rep.* 25, 211–216.

Mizutani, T., Tanabe, K., and Yamada, K. (1998b). A G.U base pair in the eukaryotic selenocysteine tRNA is important for interaction with SePF, the putative selenocysteine-specific elongation factor. *FEBS Lett.* 429, 189–193. doi: 10.1016/s0014-5793(98)00589-4

Mukai, T., Englert, M., Tripp, H. J., Miller, C., Ivanova, N. N., Rubin, E. M., et al. (2016). Facile recoding of selenocysteine in nature. *Angew. Chem. Int. Ed. Engl.* 55, 5337–5341. doi: 10.1002/anie.201511657

Mukai, T., Vargas-Rodriguez, O., Englert, M., Tripp, H. J., Ivanova, N. N., Rubin, E. M., et al. (2017). Transfer RNAs with novel cloverleaf structures. *Nucleic Acids Res.* 45, 2776–2785.

Mustoe, A. M., Liu, X., Lin, P. J., Al-Hashimi, H. M., Fierke, C. A., and Brooks, C. L. (2015). Noncanonical secondary structure stabilizes mitochondrial tRNAUCNSer by reducing the entropic cost of tertiary folding. *J. Am. Chem. Soc.* 137, 3592–3599. doi: 10.1021/ja5130308

Nagao, A., Suzuki, T., Katoh, T., Sakaguchi, Y., and Suzuki, T. (2009). Biogenesis of glutaminyl-mt tRNAGln in human mitochondria. *Proc. Natl. Acad. Sci. U.S.A.* 106, 16209–16214. doi: 10.1073/pnas.0907602106

Neuenfeldt, A., Lorber, B., Ennifar, E., Gaudry, A., Sauter, C., Sissler, M., et al. (2013). Thermodynamic properties distinguish human mitochondrial aspartyl-tRNA synthetase from bacterial homolog with same 3D architecture. *Nucleic Acids Res.* 41, 2698–2708. doi: 10.1093/nar/gks1322

Noutahi, E., Calderon, V., Blanchette, M., El-Mabrouk, N., and Lang, B. F. (2019). Rapid genetic code evolution in green algal mitochondrial genomes. *Mol. Biol. Evol.* 36, 766–783. doi: 10.1093/molbev/msz016

Nozawa, K., O' Donoghue, P., Gundllapalli, S., Araiso, Y., Ishitani, R., Umehara, T., et al. (2009). Pyrrolysyl-tRNA synthetase:tRNAPyl structure reveals the molecular basis of orthogonality. *Nature* 457, 1163–1167. doi: 10.1038/nature07611

Ohama, T., Yang, D. C., and Hatfield, D. L. (1994). Selenocysteine tRNA and serine tRNA are aminoacylated by the same synthetase, but may manifest different identities with respect to the long extra arm. *Arch. Biochem. Biophys.* 315, 293–301. doi: 10.1006/abbi.1994.1503

Ohtsuki, T., Kawai, G., and Watanabe, K. (2002). The minimal tRNA: unique structure of Ascaris suum mitochondrial tRNA(Ser)(UCU) having a short T arm and lacking the entire D arm. *FEBS Lett.* 514, 37–43. doi: 10.1016/s0014-5793(02)02328-1

Okimoto, R., and Wolstenholme, D. R. (1990). A set of tRNAs that lack either the TΨC arm or the dihydrouridine arm: towards a minimal tRNA adaptor. *EMBO J.* 9, 3405–3411. doi: 10.1002/j.1460-2075.1990.tb07542.x

Palioura, S., Sherrer, R. L., Steitz, T. A., Söll, D., and Simonovic, M. (2009). The human SepSecS-tRNASec complex reveals the mechanism of selenocysteine formation. *Science* 325, 321–325. doi: 10.1126/science.1173755

Pallanck, L., Li, S., and Schulman, L. H. (1992). The anticodon and discriminator base are major determinants of cysteine tRNA identity *in vivo*. *J. Biol. Chem.* 267, 7221–7223.

Palopoli, M. F., Minot, S., Pei, D., Satterly, A., and Endrizzi, J. (2014). Complete mitochondrial genomes of the human follicle mites *Demodex brevis* and *D. folliculorum*: novel gene arrangement, truncated tRNA genes, and ancient divergence between species. *BMC Genomics* 15:1124. doi: 10.1186/1471-2164-15-1124

Polycarpo, C., Ambrogelly, A., Bérubé, A., Winbush, S. M., Mccloskey, J. A., Crain, P. F., et al. (2004). An aminoacyl-tRNA synthetase that specifically activates pyrrolysine. *Proc. Natl. Acad. Sci. U.S.A.* 101, 12450–12454. doi: 10.1073/pnas.0405362101

Sakurai, M., Ohtsuki, T., and Watanabe, K. (2005). Modification at position 9 with 1-methyladenosine is crucial for structure and function of nematode mitochondrial tRNAs lacking the entire T-arm. *Nucleic Acids Res.* 33, 1653–1661. doi: 10.1093/nar/gki309

Salinas-Giegé, T., Giegé, R., and Giegé, P. (2015). tRNA biology in mitochondria. *Int. J. Mol. Sci.* 16, 4518–4559. doi: 10.3390/ijms16034518

Schatz, D., Leberman, R., and Eckstein, F. (1991). Interaction of *Escherichia coli* tRNASer with its cognate aminoacyl-tRNA synthetase as determined by footprinting with phosphorothioate-containing tRNA transcripts. *Proc. Natl. Acad. Sci. U.S.A.* 88, 6132–6136. doi: 10.1073/pnas.88.14.6132

Schön, A., Böck, A., Ott, G., Sprinzl, M., and Söll, D. (1989). The selenocysteine-inserting opal suppressor serine tRNA from *E. coli* is highly unusual in structure and modification. *Nucleic Acids Res.* 17, 7159–7165. doi: 10.1093/nar/17.18.7159

Seki, E., Yanagisawa, T., Kuratani, M., Sakamoto, K., and Yokoyama, S. (2020). Fully productive cell-free genetic code expansion by structure-based engineering of *Methanomethylophilus alvus* pyrrolysyl-tRNA synthetase. *ACS Synth. Biol.* 9, 718–732. doi: 10.1021/acssynbio.9b00288

Serrao, V. H. B., Silva, I. R., Da Silva, M. T. A., Scortecci, J. F., De Freitas Fernandes, A., and Thiemann, O. H. (2018). The unique tRNASec and its role in selenocysteine biosynthesis. *Amino Acids* 50, 1145–1167. doi: 10.1007/s00726-018-2595-6

Sherrer, R. L., Araiso, Y., Aldag, C., Ishitani, R., Ho, J. M., Söll, D., et al. (2011). C-terminal domain of archaeal *O*-phosphoseryl-tRNA kinase displays large-scale motion to bind the 7-bp D-stem of archaeal tRNASec. *Nucleic Acids Res.* 39, 1034–1041. doi: 10.1093/nar/gkq845

Sherrer, R. L., Ho, J. M., and Söll, D. (2008). Divergence of selenocysteine tRNA recognition by archaeal and eukaryotic *O*-phosphoseryl-tRNASec kinase. *Nucleic Acids Res.* 36, 1871–1880. doi: 10.1093/nar/gkn036

Shimada, N., Suzuki, T., and Watanabe, K. (2001). Dual mode recognition of two isoacceptor tRNAs by mammalian mitochondrial seryl-tRNA synthetase. *J Biol. Chem.* 276, 46770–46778. doi: 10.1074/jbc.m105150200

Sigler, P. B. (1975). An analysis of the structure of tRNA. *Annu. Rev. Biophys. Bioeng.* 4, 477–527.

Sohm, B., Frugier, M., Brulé, H., Olszak, K., Przykorska, A., and Florentz, C. (2003). Towards understanding human mitochondrial leucine aminoacylation identity. *J. Mol. Biol.* 328, 995–1010. doi: 10.1016/s0022-2836(03)00373-5

Sohm, B., Sissler, M., Park, H., King, M. P., and Florentz, C. (2004). Recognition of human mitochondrial tRNAUURLeu by its cognate leucyl-tRNA synthetase. *J. Mol. Biol.* 339, 17–29. doi: 10.1016/j.jmb.2004.03.066

Sprinzl, M., Horn, C., Brown, M., Ioudovitch, A., and Steinberg, S. (1998). Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res.* 26, 148–153. doi: 10.1093/nar/26.1.148

Srinivasan, G., James, C. M., and Krzycki, J. A. (2002). Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA. *Science* 296, 1459–1462. doi: 10.1126/science.1069588

Steinberg, S., and Cedergren, R. (1994). Structural compensation in atypical mitochondrial tRNAs. *Nat. Struct. Biol.* 1, 507–510. doi: 10.1038/nsb0894-507

Steinberg, S. V., Ioudovitch, A., and Cedergren, R. (1998). The secondary structure of eukaryotic selenocysteine tRNA: 7/5 versus 9/4. *RNA* 4, 241–245.

Sturchler, C., Westhof, E., Carbon, P., and Krol, A. (1993). Unique secondary and tertiary structural features of the eucaryotic selenocysteine tRNASec. *Nucleic Acids Res.* 21, 1073–1079. doi: 10.1093/nar/21.5.1073

Sturchler-Pierrat, C., Hubert, N., Totsuka, T., Mizutani, T., Carbon, P., and Krol, A. (1995). Selenocysteylation in eukaryotes necessitates the uniquely long aminoacyl acceptor stem of selenocysteine tRNASec. *J. Biol. Chem.* 270, 18570–18574. doi: 10.1074/jbc.270.31.18570

Su, D., Lieberman, A., Lang, B. F., Simonovic, M., Söll, D., and Ling, J. (2011). An unusual tRNAThr derived from tRNAHis reassigns in yeast mitochondria the CUN codons to threonine. *Nucleic Acids Res.* 39, 4866–4874. doi: 10.1093/nar/gkr073

Suzuki, T., Miller, C., Guo, L. T., Ho, J. M. L., Bryson, D. I., Wang, Y. S., et al. (2017). Crystal structures reveal an elusive functional domain of pyrrolysyl-tRNA synthetase. *Nat. Chem. Biol.* 13, 1261–1266. doi: 10.1038/nchembio.2497

Suzuki, T., Nagao, A., and Suzuki, T. (2011). Human mitochondrial tRNAs: biogenesis, function, structural aspects, and diseases. *Annu. Rev. Genet.* 45, 299–329. doi: 10.1146/annurev-genet-110410-132531

Tharp, J. M., Ehnbom, A., and Liu, W. R. (2018). tRNAPyl: structure, function, and applications. *RNA Biol.* 15, 441–452. doi: 10.1080/15476286.2017.1356561

Théobald-Dietrich, A., Frugier, M., Giegé, R., and Rudinger-Thirion, J. (2004). Atypical archaeal tRNA pyrrolysine transcript behaves towards EF-Tu as a typical elongator tRNA. *Nucleic Acids Res.* 32, 1091–1096. doi: 10.1093/nar/gkh266

Tomari, Y., Suzuki, T., and Ueda, T. (2002). tRNA recognition by CCA-adding enzyme. *Nucleic Acids Res.* 2002, 77–78. doi: 10.1093/nass/2.1.77

Turanov, A. A., Lobanov, A. V., Fomenko, D. E., Morrison, H. G., Sogin, M. L., Klobutcher, L. A., et al. (2009). Genetic code supports targeted insertion of two amino acids by one codon. *Science* 323, 259–261. doi: 10.1126/science.1164748

Wakita, K., Watanabe, Y., Yokogawa, T., Kumazawa, Y., Nakamura, S., Ueda, T., et al. (1994). Higher-order structure of bovine mitochondrial tRNAPhe lacking the 'conserved' GG and TΨCG sequences as inferred by enzymatic and chemical probing. *Nucleic Acids Res.* 22, 347–353. doi: 10.1093/nar/22.3.347

Wan, W., Huang, Y., Wang, Z., Russell, W. K., Pai, P. J., Russell, D. H., et al. (2010). A facile system for genetic incorporation of two different noncanonical

amino acids into one protein in *Escherichia coli*. *Angew. Chem. Int. Ed. Engl.* 49, 3211–3214. doi: 10.1002/anie.201000465

Wan, W., Tharp, J. M., and Liu, W. R. (2014). Pyrrolysyl-tRNA synthetase: an ordinary enzyme but an outstanding genetic code expansion tool. *Biochim. Biophys. Acta* 1844, 1059–1070. doi: 10.1016/j.bbapap.2014.03.002

Watanabe, K. (2010). Unique features of animal mitochondrial translation systems. The non-universal genetic code, unusual features of the translational apparatus and their relevance to human mitochondrial diseases. *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.* 86, 11–39. doi: 10.2183/pjab.86.11

Watanabe, Y., Kawai, G., Yokogawa, T., Hayashi, N., Kumazawa, Y., Ueda, T., et al. (1994a). Higher-order structure of bovine mitochondrial tRNAUGASer: chemical modification and computer modeling. *Nucleic Acids Res.* 22, 5378–5384. doi: 10.1093/nar/22.24.5378

Watanabe, Y., Tsurui, H., Ueda, T., Furushima, R., Takamiya, S., Kita, K., et al. (1994b). Primary and higher order structures of nematode (*Ascaris suum*) mitochondrial tRNAs lacking either the T or D stem. *J. Biol. Chem.* 269, 22902–22906.

Wende, S., Platzer, E. G., Jühling, F., Pütz, J., Florentz, C., Stadler, P. F., et al. (2014). Biological evidence for the world's smallest tRNAs. *Biochimie* 100, 151–158. doi: 10.1016/j.biochi.2013.07.034

Willis, J. C. W., and Chin, J. W. (2018). Mutually orthogonal pyrrolysyl-tRNA synthetase/tRNA pairs. *Nat. Chem.* 10, 831–837. doi: 10.1038/s41557-018-0052-5

Wittenhagen, L. M., and Kelley, S. O. (2003). Impact of disease-related mitochondrial mutations on tRNA structure and function. *Trends Biochem. Sci.* 28, 605–611. doi: 10.1016/j.tibs.2003.09.006

Wolfson, A. D., Khvorova, A. M., Sauter, C., Florentz, C., and Giege, R. (1999). Mimics of yeast tRNAAsp and their recognition by aspartyl-tRNA synthetase. *Biochemistry* 38, 11926–11932. doi: 10.1021/bi9908383

Wolstenholme, D. R., Macfarlane, J. L., Okimoto, R., Clary, D. O., and Wahleithner, J. A. (1987). Bizarre tRNAs inferred from DNA sequences of mitochondrial genomes of nematode worms. *Proc. Natl. Acad. Sci. U.S.A.* 84, 1324–1328. doi: 10.1073/pnas.84.5.1324

Wu, X. Q., and Gross, H. J. (1993). The long extra arms of human tRNA(Ser)Sec and tRNASer function as major identify elements for serylation in an orientation-dependent, but not sequence-specific manner. *Nucleic Acids Res.* 21, 5589–5594. doi: 10.1093/nar/21.24.5589

Wu, X. Q., and Gross, H. J. (1994). The length and the secondary structure of the D-stem of human selenocysteine tRNA are the major identity determinants for serine phosphorylation. *EMBO J.* 13, 241–248. doi: 10.1002/j.1460-2075.1994.tb06254.x

Yamaguchi, A., Iraha, F., Ohtake, K., and Sakamoto, K. (2018). Pyrrolysyl-tRNA synthetase with a unique architecture enhances the availability of lysine derivatives in synthetic genetic codes. *Molecules* 23:2460. doi: 10.3390/molecules23102460

Yarham, J. W., Elson, J. L., Blakely, E. L., Mcfarland, R., and Taylor, R. W. (2010). Mitochondrial tRNA mutations and disease. *Wiley Interdiscip. Rev. RNA* 1, 304–324.

Zhang, Y., Baranov, P. V., Atkins, J. F., and Gladyshev, V. N. (2005). Pyrrolysine and selenocysteine use dissimilar decoding strategies. *J. Biol. Chem.* 280, 20740–20751. doi: 10.1074/jbc.m501458200

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

REVIEW
published: 26 October 2020
doi: 10.3389/fmicb.2020.596092

Bacterial Spore mRNA – What's Up With That?

Peter Setlow[1]* *and Graham Christie*[2]*

[1]Department of Molecular Biology and Biophysics, UConn Health, Farmington, CT, United States, [2]Department of Chemical Engineering and Biotechnology, University of Cambridge, Cambridge, United Kingdom

Bacteria belonging to the orders Bacillales and Clostridiales form spores in response to nutrient starvation. From a simplified morphological perspective, the spore can be considered as comprising a central protoplast or core, that is, enveloped sequentially by an inner membrane (IM), a peptidoglycan cortex, an outer membrane, and a proteinaceous coat. All of these structures are characterized by unique morphological and/or structural features, which collectively confer metabolic dormancy and properties of environmental resistance to the quiescent spore. These properties are maintained until the spore is stimulated to germinate, outgrow and form a new vegetative cell. Spore germination comprises a series of partially overlapping biochemical and biophysical events – efflux of ions from the core, rehydration and IM reorganization, disassembly of cortex and coat – all of which appear to take place in the absence of *de novo* ATP and protein synthesis. If the latter points are correct, why then do spores of all species examined to date contain a diverse range of mRNA molecules deposited within the spore core? Are some of these molecules "functional," serving as translationally active units that are required for efficient spore germination and outgrowth, or are they just remnants from sporulation whose sole purpose is to provide a reservoir of ribonucleotides for the newly outgrowing cell? What is the fate of these molecules during spore senescence, and indeed, are conditions within the spore core likely to provide any opportunity for changes in the transcriptional profile of the spore during dormancy? This review encompasses a historical perspective of spore ribonucleotide biology, from the earliest biochemical led analyses – some of which in hindsight have proved to be remarkably prescient – through the transcriptomic era at the turn of this century, to the latest next generation sequencing derived insights. We provide an overview of the key literature to facilitate reasoned responses to the aforementioned questions, and many others, prior to concluding by identifying the major outstanding issues in this crucial area of spore biology.

Keywords: *Bacillus*, spores, mRNA, germination, sporulation

OPEN ACCESS

Edited by:
Orna Amster-Choder,
Hebrew University of Jerusalem, Israel

Reviewed by:
Jonathan Dworkin,
Columbia University, United States
Aimee Shen,
Tufts University School of Medicine, United States

***Correspondence:**
Peter Setlow
setlow@uchc.edu;
setlow@nso2.uchc.edu
Graham Christie
gc301@cam.ac.uk

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *18 August 2020*
Accepted: *28 September 2020*
Published: *26 October 2020*

Citation:
Setlow P and Christie G (2020) Bacterial Spore mRNA – What's Up With That? Front. Microbiol. 11:596092. doi: 10.3389/fmicb.2020.596092

PRELUDE

When RNA species in spores of Bacillales species were first characterized more than 50 years ago, spores were found to contain both rRNAs and tRNAs, although some of these nucleic acids exhibited a few apparent differences from their growing cell counterparts (Chambon et al., 1968; Deutscher et al., 1968; Setlow et al., 1974). The existence of functional spore mRNA, however, seemed less likely, since several early studies showed that spores germinating

in the presence of inhibitors of RNA polymerase made little if any detectable protein (Steinberg et al., 1965; Torriani and Levinthal, 1967; Setlow and Primus, 1975), although it is possible that RNA synthesis inhibitors may not enter the spore core until germination is complete. However, even ~50 years ago there was evidence that spores did contain some mRNA, and that some of this mRNA was translationally active, at least *in vitro* (Chambon et al., 1968; Deutscher et al., 1968; Jeng and Doi, 1974), even if this spore mRNA appeared to be non-functional during spore germination and subsequent outgrowth. However, the thinking about spore mRNA was changed dramatically beginning ~14 years ago in the transcriptomics revolution, as over the following years a number of laboratories found 100s to 1,000s of specific mRNAs in spores of a large number of Bacillales as well as Clostridiales species (Bergman et al., 2006; Bettegowda et al., 2006; Keijser et al., 2007; Jones et al., 2008; van Melis et al., 2011; Nicholson et al., 2012; Segev et al., 2012; Bassi et al., 2013; Dembek et al., 2013; Bate et al., 2014; Swarge et al., 2018). This review will trace the history of spore mRNA biology and summarize the more recent work, some of which is ongoing. The focus will be on what we do know, and more importantly what we do not know about spore mRNA, including how these mRNAs come to be present in dormant spores. Perhaps most significantly, we will also consider the function of the plethora of mRNAs in dormant bacterial spores, and, just as importantly, what this function is not.

INTRODUCTION

Spores are formed by some Firmicutes species, the most well-studied being those of the Bacillales, although studies on spores of Clostridiales species are increasing (Tan and Ramamurthi, 2014; Setlow and Johnson, 2019; Shen et al., 2019). Such spores are considered metabolically dormant and extremely resistant to all manner of potentially harmful treatments, including wet and dry heat, desiccation, high radiation levels, and a host of toxic chemicals, including antibiotics, and can survive for many years (Setlow, 2006; Ulrich et al., 2018). Notably, dormant spores' central core (**Figure 1**), comparable to the protoplast of a growing cell, has an extremely low water content, as low as 25% of wet wt in spores of some thermophiles and 35% of wet wt in the well-studied *Bacillus subtilis* spores; this is in contrast to the value of 80% of wet wt as water in growing cells or fully germinated spores (Gerhardt and Marquis, 1989). Spores have been studied in part because of their fascinating life cycle of growth, sporulation, and spore germination (Tan and Ramamurthi, 2014; Setlow et al., 2017; Shen et al., 2019). In addition, spores of a number of species are vectors for human diseases or intoxications, as well as food spoilage and food-borne disease. Consequently, there is also much applied interest in spore resistance and germination, the latter largely because a germinated spore has lost the high resistance of the dormant spore, and is easy to kill (Setlow and Johnson, 2019; Buhr et al., 2020). Together, both the basic science and applied interests in spores' formation, properties, and germination has made this one of the best studied developmental system in biology, with definitive work going back many years.

The Early Years

Given the technology available for use 50-plus years ago, much of early work on spores was descriptive (Setlow et al., 2017; Setlow and Johnson, 2019). This work involved: (i) isolation and mapping of sporulation mutants in *B. subtilis* because some strains of this species are naturally transformable (this is also the major reason that *B. subtilis* became the "model" spore former), (ii) characterizing biochemical and morphological events in sporulation and spore germination and their timing, including changes in metabolism, (iii) characterization and quantitation of the components in the dormant spore, including nucleic acids and small molecules, and using spores of multiple species, and (iv) examining early events in spore germination, including the release of small molecules, degradation of a specific peptidoglycan layer unique to spores, termed the cortex (**Figure 1**), followed by a return to metabolism and macromolecular synthesis in outgrowth, which converts a germinated spore into a growing cell (although enzyme activity in the spore core could well resume before completion of spore germination – see section Issues for the Future). Notably, one of the unique features of spores of all Bacillales and Clostridiales is the presence in the spore core of ~25% of core dry wt as CaDPA, a 1:1 chelate of Ca^{2+} and pyridine-2,6-dicarboxylic acid (dipicolinic acid, DPA; Gerhardt and Marquis, 1989). DPA is made in the mother cell compartment of the sporulating cell, and imported into developing spores probably as CaDPA *via* several spore-specific channels (Ramírez-Guadiana et al., 2017; Setlow and Johnson, 2019); this incorporation almost certainly requires energy and lowers the spore core water content appreciably. CaDPA also has roles in spore resistance, is rapidly released during spore germination, and has a signaling role in germination of spores of most, but probably not all Firmicutes (Setlow et al., 2017; Shen et al., 2019).

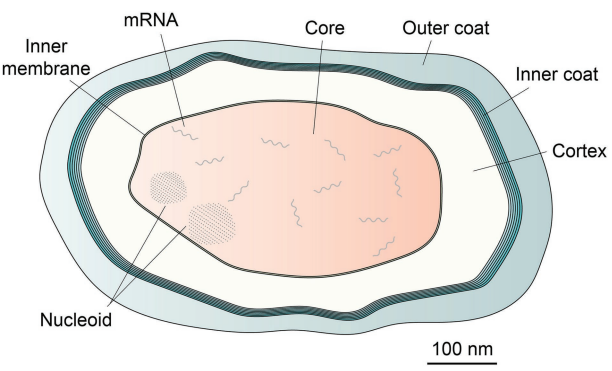


FIGURE 1 | Schematic of a dormant *Bacillus subtilis* spore. Spores of all species share similar morphological features, namely a coat, which in this species can be sub-divided into distinct inner and outer layers, a peptidoglycan cortex, and a membrane bound spore core. The nucleoid comprises DNA encrusted by protective SASP proteins, visible as clusters of ordered dots. mRNA content is present only within the spore core.

Among the important observations made in this early period of spore research was the characterization of spore nucleic acids, most thoroughly in *Bacillus megaterium*, an attractive subject for biochemical work because it sporulates profusely, its spores are easily purified and these spores germinate very rapidly (Levinson and Hyatt, 1966). Early analyses of nucleic acids found that spores contained the expected rRNAs, in some cases with minor modifications, and also tRNAs (Chambon et al., 1968; Deutscher et al., 1968). However, the 3' terminal A residue almost always present in the tRNA in vegetative cells is absent from approximately one third of total spore tRNAs and some adjacent C residues are also absent (Setlow, 1974). Several researchers also found a small amount of RNA in spores (2–4% of the total), which was larger than tRNA but was not small rRNA (Chambon et al., 1968; Deutscher et al., 1968; Jeng and Doi, 1974). One of the latter reports showed that this RNA hybridized throughout the spore genome, and another reported that this RNA directed a small amount of protein synthesis *in vitro*. This latter RNA was thus suggested to be mRNA or "mRNA-like." It was also clear from the early work that tRNA in spores is minimally aminoacylated, if at all, although aminoacyl-tRNA synthetases are present (Deutscher et al., 1968; Setlow, 1974).

In addition, studies of small molecules in spores ~50 years ago found that while spores do have the coenzymes generally found in all growing cells such as FAD, NAD, NADP, and the four common ribo and deoxyribonucleotides, these are all in what might be termed a "low energy state," with minimal if any NADH, NADPH, or ATP (Setlow and Kornberg, 1970a,b; Setlow and Setlow, 1977). Indeed, the absence of ATP (≤1% of the total adenine nucleotide pool) from spores shown ~50 years ago by chromatography of extracts from ^{32}P-labeled spores and enzymatic assays using firefly luciferase, was confirmed more recently using ^{31}P-NMR analysis of spore extracts (Ghosh et al., 2015). The latter assays were also carried out on spores incubated at physiological temperatures with care taken to prevent spore germination using appropriate mutant strains, and again with no ATP detected. A major conclusion from this older and newer work is that dormant spores in water have an energy charge [(ATP) + 0.5(ADP)/(ATP) + (ADP) + (AMP)] of ≤0.1, while in growing cells of both prokaryotes and eukaryotes' energy charge is maintained at ~0.9 (Chapman et al., 1971). The ratio of ATP relative to the total adenine nucleotide pool in the dormant spore is thus ≤0.01, while this is normally at 0.8 in growing cells. This low ratio would in all likelihood make normal biosynthetic reactions using ATP thermodynamically unfavorable in the dormant spore core, leaving aside the role of the low core water content that most likely greatly restrains protein movement and enzymatic activity there also (Gerhardt and Marquis, 1989; Cowan et al., 2004).

Another set of important observations made >40 years ago was that dormant spores of *B. megaterium* did not contain one or more key enzymes for the synthesis of ribonucleotides or amino acids (Setlow and Kornberg, 1970b; Setlow and Primus, 1975). As a consequence, despite synthesis of RNA and proteins beginning relatively soon after spores were triggered to germinate, there was no detectable *de novo* synthesis of ribonucleotides or amino acids at this time, and biosynthesis of these small molecules only began long after germination was complete and then only after the required biosynthetic enzymes were synthesized during spore outgrowth. Thus, any synthesis of RNA or proteins occurring soon after the triggering of spore germination must come from spores' reserves of ribonucleotides or amino acids, respectively. Indeed, spores contain a large amount of protein, classified as small, acid-soluble spore proteins (SASP), that are degraded soon after germination is initiated, and this provides amino acids for new protein synthesis at this time (Setlow, 2007). The energy for the earliest RNA and protein synthesis after spore germination is initiated, as well as tRNA repair, can come from the environment, if available, and also by utilizing spores' endogenous energy reserves, including 3-phosphoglyceric acid (3PGA) as well as degradation of some amino acids generated by proteolysis (Nelson and Kornberg, 1970b; Setlow and Kornberg, 1970a; Setlow and Primus, 1975). There were also suggestions in this early period of spore research that ribonucleotides needed for RNA synthesis soon after triggering spore germination would also come from breakdown of spore RNAs (Setlow and Kornberg, 1970b), but the specific RNA degraded was not definitively identified.

Notably, the absence of nucleotide biosynthetic enzymes from spores and that these proteins are synthesized at various times in spore outgrowth have also recently been shown for *B. subtilis* spores in several proteomic studies (Swarge et al., 2018, 2020b). The latter findings were also made for amino acid biosynthetic enzymes as well. Other important observations (Nelson and Kornberg, 1970b; Setlow and Kornberg, 1970a,b; Scott and Ellar, 1978) made early in biochemical analyses of spore germination were that: (i) ATP accumulation began only after germination was completed or very near completion, (ii) catabolism of dormant spores' endogenous energy reserves, most notably 3PGA, could support almost all ATP needed for at least the first few minutes after germination was triggered, and (iii) even if ATP accumulation was blocked, germination still took place. This latter observation was also confirmed more recently (Cowan et al., 2004). The thinking about mRNA in spores was also greatly influenced by experiments showing that spores of several species incubated with germinants in the presence of inhibitors of RNA synthesis still germinated, but protein synthesis was blocked >99% (Steinberg et al., 1965; Torriani and Levinthal, 1967; Setlow and Kornberg, 1970b). While this certainly does not prove that spores have no mRNA, it does suggest that the amount of functional mRNA in spores was either minimal, or this mRNA was very rapidly degraded early in spore germination.

Overall, based on the evidence available in the early years of biochemical studies on dormant spores and spore germination, the general thinking was that: (i) there was minimal if any ATP in spores and therefore minimal if any metabolic activity, (ii) this latter statement included even metabolism of endogenous energy reserves such as 3PGA – note that enzymes for converting

3PGA to acetate, NADH, and ATP are present in spores and generate ATP early after spore germination is initiated, but 3PGA is stable in dormant spores (Nelson and Kornberg, 1970a), and this latter point was revisited recently and confirmed by ^{31}P-NMR of extracts from spores stored at physiological temperatures for weeks (Ghosh et al., 2015), (iii) there is neither functional mRNA nor ribonucleotide biosynthetic enzymes in dormant spores, and therefore synthesis of protein when spores germinate will require synthesis of new mRNA using ribonucleotides generated by degradation of some dormant spore RNA. Indeed, there was indirect evidence for this RNA degradation even 50 years ago (Setlow and Kornberg, 1970a,b), as *B. megaterium* spores germinating with glucose plus an inhibitor of RNA synthesis accumulated ATP levels >5-fold higher than did spores germinating with glucose alone, and (iv) spores contained ribosomes, the components of which appeared to be indistinguishable from their vegetative cell counterparts (Chambon et al., 1968; Deutscher et al., 1968). Moreover, spore ribosomes appeared to be translationally active, in the sense that when extracted, they were capable of synthesizing polyphenylalanine from a polyuracil template in cell-free reaction mixtures at comparable rates to ribosomes purified from log-phase vegetative cells (Chambon et al., 1968; Deutscher et al., 1968). Whether ribosome hibernation factors, such as the Hpf protein, have a role in spores in preserving the structural integrity of ribosomes during dormancy, analogous to that observed in stationary phase vegetative cells (Feaga et al., 2020), perhaps facilitating rapid resumption of activity upon germination, has not been established. Regardless, and most importantly, the thinking about spores until ~2006 was that they did not have "functional" mRNA.

The Omics Era Arrives

The idea that spores do not have mRNA, even if this was not quite what the data from the 1960s and 1970s showed, took a real body blow in 2006 with the publication of two reports of analyses of mRNAs in bacteria of two spore formers, *B. subtilis* and *Clostridium novyi*, using microarray hybridization technology (Bettegowda et al., 2006; Keijser et al., 2007). These reports examined not only cells of these two organisms, but also dormant spores, and found readily detectable levels of mRNAs in spores from a number of genes coding for proteins ~23 in *B. subtilis* spores and 960 in *C. novyi* spores. The difference in the numbers in these two species was likely simply because of what was deemed significant in the different studies. Notably, a large number of the genes encoding mRNAs in *B. subtilis* were ones that: (i) encoded proteins known to be present in dormant spores and (ii) were expressed only in the developing forespore late in sporulation, and using RNA polymerase with the forespore-specific σ (specificity) factors for RNA polymerase, σ^F and/or σ^G (**Figure 2**; Arrieta-Ortiz et al., 2015; Swarge et al., 2018). The *C. novyi* spore data were more difficult to analyze fully because of the lack of proteomics data for spores of this organism. However, at least three of the most abundant five mRNAs in these spores encoded small, SASP, which in other spore formers are known to be: (i) synthesized only in the developing spore under σ^G and/or σ^F control and (ii) abundant proteins in spores (Setlow, 2007).

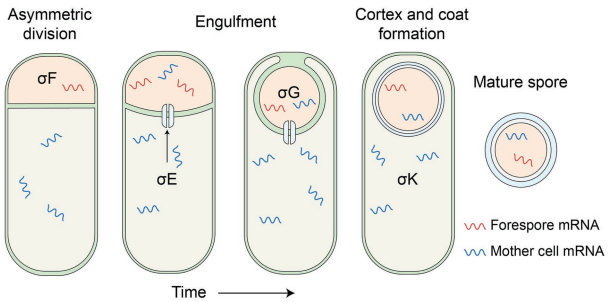


FIGURE 2 | Schematic of *B. subtilis* sporulation. The process starts with an asymmetric cell division, which results ultimately in the larger mother cell compartment engulfing the nascent forespore. Both compartments are characterized by the activity of RNA polymerase containing sequentially different sporulation-specific sigma factors, the products of which progressively assemble the spore prior to its release to the environment. Despite being metabolically dormant, spores contain significant amounts of mRNA, including a minority of σE and σK associated transcripts. Neither the source nor purpose, if any, of these mother-cell associated mRNA molecules has been established, although the SpoIIIA-IIQ channel that connects both cellular compartments at or around the engulfment stage of the process may have a contributing role in the former.

However, note that none of these spore mRNAs have been shown to direct the synthesis of a protein after spores germinate.

The publication of the two papers cited above applying the new transcriptomics technology to spores unleashed a flood of publications reporting the mRNAs in spores of multiple species. The list of these species includes a veritable who's who of spore formers and includes the spores of the two species noted above as well as spores of *Bacillus cereus*, *Bacillus anthracis*, *Bacillus thuringiensis*, *Clostridium acetobutylicum*, *Clostridium sporogenes*, and *Clostridioides difficile*, and several additional publications examining *B. subtilis* spores (Bergman et al., 2006; Bettegowda et al., 2006; Keijser et al., 2007; Jones et al., 2008; van Melis et al., 2011; Nicholson et al., 2012; Segev et al., 2012; Bassi et al., 2013; Dembek et al., 2013; Bate et al., 2014; Bassi et al., 2016; Nagler et al., 2016; Swarge et al., 2020a). All these studies reported generally similar findings as follows: (i) there was indeed hybridization of spore RNA to protein coding regions from genomes of these species, with these spore RNAs being almost certainly mRNAs, (ii) the variability in the numbers of these mRNAs was quite high, ranging from the low of 23 noted above to highs of >1,000, with most studies finding many 100s of mRNAs, (iii) as noted above, most of the very abundant *B. subtilis* spore mRNAs encoded proteins found in spores and their coding genes were expressed in the developing spore late in sporulation, but (iv) some mRNAs present at high levels in *B. subtilis* spores were not known to be expressed in the developing spore, and surprisingly, some were thought to be expressed only in the mother cell compartment of the sporulating cell (**Figure 2**). Notably, these putative mother cell mRNAs were not contaminants in outer layers of at least *B. subtilis* spores, since spores incubated in alkaline conditions which would hydrolyze RNA in spores' outer layers still had the mother cell-expressed mRNAs present in the spore core (Korza et al., 2019).

In addition to reports of the presence of multiple mRNAs in spores, there was also extensive analysis of when the various genes encoding these spore mRNAs were actually transcribed, and whether in growing cells or during sporulation. However, these analyses were hampered by the lack of specific information on where genes were transcribed in sporulation of almost all species, in particular whether in the mother cell or forespore compartment of the sporulating cell. *B. subtilis* and *B. anthracis* were the notable exceptions to the relative paucity of definitive information on the "where" question posed above, in large part because of the classification of what σ (specificity) factors for RNA polymerase directed the transcription of various genes in developing spores (Bergman et al., 2006; Arrieta-Ortiz et al., 2015). Indeed, as noted above, of the 23 mRNAs identified in the initial report on *B. subtilis* spore mRNAs, the transcription of almost all was dependent on σ^F and σ^G known to be responsible for specific transcription late in forespore development (Keijser et al., 2007). However, for studies in which many 100s of *B. subtilis* spore mRNAs were identified, this was not the case. The second major type of analysis of the many 100s of mRNAs found in spores of different species was using one of several programs to analyze the role of the proteins encoded by the spore mRNAs in metabolism, macromolecular synthesis, and stress responses etc. (Hosack et al., 2003; Kanehisa et al., 2017). This generated a huge amount of speculation on the precise functions for specific mRNAs in spores.

A second application of the new transcriptomics technology was looking at the fate of these mRNAs in spores, and one study with *B. anthracis* spores found that they almost completely disappeared when spores were stored for ~30 days at 37°C (Bergman et al., 2006). This was largely true for *B. subtilis* spore mRNAs as well, and in even less time (Segev et al., 2012). The latter report also found that even spore rRNA was fragmented when spores were stored for various lengths of time at 37–50°C; this fragmentation also appeared to be endonucleolytic, and it was suggested to be due to RNase Y. This finding of rRNA degradation in spores incubated at elevated temperatures was confirmed more recently, and in both *B. megaterium* and *B. subtilis* spores, although not the involvement of RNase Y, and this rRNA fragmentation did not decrease spore viability or germination rates (Korza et al., 2016). However, since the rRNA fragmentation was not associated with generation of mononucleotides, it certainly appeared to be endonucleolytic. The fact that the rate of this rRNA fragmentation increased drastically as spore incubation temperature increased up to ~80°C, even though spores remained viable, suggested that the rRNA fragmentation was not enzyme catalyzed. Notably, the mRNAs in *C. novyi* spores appeared to be more stable than in spores of *Bacillus* species (Bettegowda et al., 2006), but this finding has not been studied further.

Another extremely surprising observation made based on the new technology was a report that not only was rRNA fragmented in spores incubated at 37–50°C, but that there was actual synthesis of mRNAs in spores incubated at 4°C, as shown primarily by transcriptomics and qRT-PCR (Segev et al., 2012). This report was particularly surprising in that as noted in section The Early Years above, spores have very low, if any, levels of ATP, and generate no detectable steady state level of ATP on incubation at 37°C. In addition, it seems very likely that: (i) proteins needed for RNA synthesis would not work in the low water content in the spore core and (ii) spores' extremely low ratio of ATP/AMP would thermodynamically preclude phosphodiester bond formation, and if pyrophosphatase also does not work in the dormant spore core, this would make phosphodiester bond formation even more unfavorable. Thus, it is difficult to imagine how RNA synthesis would go on in all spores in the spore population at the same time, although it is of course impossible to prove that no spores in populations do this intermittently.

The overall picture that emerged from the work noted above in 2006–2016 was that: (i) spores appeared to have many 100s of different mRNA species, (ii) some of these spore mRNAs are synthesized specifically late in spore formation, and thus might be "left over" when the water content in the developing spore drops from 80% of wet wt to <35%, (iii) but many spore mRNAs are not known to be synthesized specifically in developing spores, and (iv) and surprisingly, some of these spore mRNAs are thought to be expressed in the mother cell compartment of the sporulating cell. So, the key questions raised but not answered by the transcriptomics work between 2006 and 2016 were: (i) where did all these various mRNAs come from and (ii) what is their function?

Some new insight into the answers to these questions posed above came from application of even newer transcriptomics technology, RNA-Seq, to analysis of *B. subtilis* spore mRNA (Nagler et al., 2016). This work found ~1,800 individual mRNAs in dormant *B. subtilis* spores, and also reported the relative abundance of these mRNAs, based on values for reads per kb of transcript per million mapped reads (RPKM). Notably, there was a huge variation in these RPKM values, ranging from $>10^6$ for the most abundant mRNA down to <10 for many others, with all else in between. Analysis of these many mRNAs found that while the ~50 most abundant ones were almost all expressed in the developing spore under σ^G control (and see below), many spore mRNAs were known to be expressed in the mother cell compartment of the sporulating cell (Korza et al., 2019), and how such mRNAs ended up in the spore was not clear. Equally, only two of the mother-cell associated mRNA molecules were present in the top 50 or so most abundant transcripts, averaging roughly one molecule per spore (in contrast, the most abundant mRNA molecules are present at ~100 copies per spore; Korza et al., 2019). It is possible also that mother-cell associated mRNA is fragmented and may not serve as a template for translation, as discussed further below. Regardless, the results from the use of this new technology, in some ways, raised more questions than it answered. However, the improved quantitation that RNA-Seq provided over microarray hybridization, especially in terms of the whole dynamic range, certainly made the questions about these mRNAs more specific, particularly in terms of the quantitation of spore mRNAs.

The Smoke Begins to Clear

While the reports of transcriptomics work that appeared starting in 2006 and up to 2016 provided much new information on spore mRNAs, the role of all these mRNAs was not clear.

It was also not clear how all these mRNAs got into the spore core, since many were known to be synthesized in the mother cell compartment of the sporulating cell. However, the shift to analyses by RNA-Seq in the 2016 report (Nagler et al., 2016) provided important quantitative information, since the RPKM values for all ~1,800 dormant spore mRNAs identified varied over $>10^5$-fold, a dynamic range that could not easily be achieved with microarray technology. Notably, of the most abundant ~50 mRNAs detected in this RNA-Seq study almost all: (i) were expressed in the developing spore late in sporulation, and under the control of σ^F and/or σ^G and (ii) encoded proteins present in the dormant spore, with some, in particular the α/β- and γ-type SASP, being the most abundant spore proteins (Setlow, 2007; Zhu and Stülke, 2018). These two types of SASP are degraded rapidly when spores germinate to provide amino acids for new protein synthesis. Notably if α/β-type SASP is not degraded, this causes the death of the outgrowing spore, as these proteins saturate spore DNA and if not degraded after germination, this will likely block transcription of much of the genome (Hayes and Setlow, 2001). Thus, it seems extremely unlikely that there will be any need for synthesis of proteins from these genes' abundant mRNAs soon after spore germination is triggered. Indeed, recent work has shown that almost all abundant mRNAs are rapidly degraded following initiation of spore germination (Swarge et al., 2018). However, it is certainly possible that one or more spore mRNAs, in particular one of the most abundant ones, could direct synthesis of a protein or proteins that are essential for the development of the germinated spore. Indeed, there was one report that the malic enzyme encoded by *malS* was synthesized very early in *B. subtilis* spore germination, and even before the phase bright dormant spore turned phase dark due to excretion of CaDPA and cortex hydrolysis (Sinai et al., 2015). Evidence to support protein synthesis prior to the phase transition was principally in the form of an unexpected increase in fluorescence associated with a MalS-GFP fusion protein, coupled with Western blot data showing an apparent increase in abundance of the protein during this period. The same authors suggested that the concentration of malate in dormant spores was relatively high, in contrast to earlier studies and a more recent one (Setlow et al., 1977; Scott and Ellar, 1978; Korza et al., 2017), providing a readily available substrate with which to generate ATP for the germinating spore. A more recent study, employing analogous strains and a range of sporulation conditions, reported a similar increase in fluorescence associated with MalS-GFP during the early stages of germination (Swarge et al., 2020b). However, in contrast to results of Sinai et al. (2015), Western blot data in the Swarge et al. report indicated that MalS-GFP abundance remained stationary throughout germination. Indeed, mass spectrometry data, using a more sensitive approach than that employed by Sinai et al. (2015), convincingly demonstrates that synthesis of MalS does not begin until ~80 min after the completion of germination. The same authors suggest that the observed increase in MalS-GFP fluorescence reflects a change in the environment of the germinating spore core – most likely partial ingress of water from the environment promoting enhanced fluorescence from existing MalS-GFP – and is not associated with *de novo* protein synthesis. The observation that *malS* mRNA sits outside the top 1,000 most abundant transcripts detected within spores, at levels indicative of being present in only a small sub-set of the population and with a copy number averaging much less than 1/spore (Nagler et al., 2016; Korza et al., 2019), further supports the conclusions reported by Swarge et al. (2020b). Indeed, very recently published work conducted by the same group, this time combining an integrated transcriptomic and multi-faceted proteomic approach, permitting a hitherto unsurpassed level of resolution to these analyses in spores, robustly supports the idea that both transcription and protein synthesis start after the completion of germination (Swarge et al., 2020a).

There is also the question of when ATP becomes available in the developing spore, and as noted in section The Early Years, this appears to begin at most very late in spore germination, and perhaps not until this process is complete and spore core water content rises to the 80% of wet wt in growing cells. While all this evidence by no means conclusively proves that spore mRNAs are not saved to direct synthesis of some proteins when spores return to life in germination, these data certainly raise significant concerns about whether this could be the case.

It was the large dynamic range of the RNA-Seq data that led to a clearer picture of what spore mRNAs might do. Work done many years ago had indicated that as in most organisms, in growing cells of *B. subtilis* and a very closely related species, approximately 3% of their RNA is mRNA (Midgley, 1969; Brown and Coleman, 1975), and notably the uncharacterized RNA in *B. subtilis* spores termed "mRNA-like" was also ~3% of total RNA (Jeng and Doi, 1974). This number plus the RPKM values from RNA-Seq analysis of spore mRNAs in rRNA depleted spore RNA samples then allowed assignment of relative levels of all spore mRNAs. The other crucial number was the exact amount of RNA present in a single *B. subtilis* spore. This number, as well as the amount of spore DNA, was determined ~50 years ago in ^{32}P-labeled *B. subtilis* and *B. megaterium* spores (Nelson and Kornberg, 1970a). Notably, although not known at that time, the value of *B. subtilis* spore DNA bp/spore determined in this report was almost identical to the length of the sequenced *B. subtilis* chromosome. In contrast, *B. megaterium* spores had slightly more than twice the DNA/spore than *B. subtilis* spores, as is now known to be consistent with *B. megaterium* making digenomic spores (Hauser and Karamata, 1992) and with a slightly larger chromosome than *B. subtilis*. Most importantly, the average values for RNA nt in *B. subtilis* spores, coupled with the assumption that 3% of spore RNA was mRNA as it is in growing cells, indicated that mRNAs in *B. subtilis* spores contributed an average of ~10^6 nt/spore. Combining this latter number with the relative levels of spore mRNAs determined from RNA-Seq, and these mRNAs' lengths adjusted for extra nt at the 5' and 3' ends, allowed calculation of the relative levels of these mRNAs needed to give the 10^6 nt in *B. subtilis* spores (Korza et al., 2019). This analysis indicated that in *B. subtilis* there were only ~50 mRNAs present at ≥1 molecule/spore in populations, with many 100s present in <10% of the population. Notably, almost all of the abundant spore

mRNAs are transcribed by RNA polymerase with σ^G or σ^F and encode proteins in spores. This suggests these mRNAs were those that were being made when the developing spore shut down, and they have been in suspended animation through dormancy.

While the above work certainly clarifies where the abundant and likely most significant spore mRNAs came from, they do not completely address the function of these mRNAs, as well as that of the many mRNAs present in only small fractions of the spore population. Might these be important in directing protein synthesis early in germination even if only for some spores? Ruling this out is most likely impossible. However, a simple experiment of leaving spores to sporulate for different times, and then analyzing their mRNA levels provided strong evidence that spore mRNAs, more specifically intact spore mRNAs, are not essential for at least *B. subtilis* spore germination and outgrowth (Camilleri et al., 2019). Thus RNA-Seq analysis on *B. subtilis* spores harvested after 2 days of sporulation gave approximately the same ~50 spore mRNAs at ≥1 molecule/spore, but spores allowed to sporulate on plates for 45–90 days at 37°C gave RNA from which levels of all spore mRNAs were decreased ≥95-fold, and rRNAs were significantly fragmented. Yet despite this loss of all intact spore mRNA and significant damage to rRNA, spore germination and return to growth were not appreciably affected. This result certainly rules out spore mRNAs as playing any role in directing protein synthesis when spores germinate, and suggests that these mRNAs main function is to serve as a reservoir of ribonucleotides for new RNA synthesis when spores germinate fully. As noted above, almost all the abundant spore mRNAs are rapidly degraded when spores germinate, but how this degradation takes place in the dormant spore is not clear, but it is not to ribonucleotides (Korza et al., 2016). Indeed, rRNA was also fragmented in spores after 45–90 days of sporulation, and at specific regions and into rather large fragments. Almost all mRNA from these aged spores is thus also almost certainly fragmented into pieces too small to be detected by RNA-Seq, but this will require further work to determine the fragment sizes of the degraded spore mRNAs. It is also possible that some small mRNA fragments play a regulatory role in gene expression, although there is currently no evidence relevant to this suggestion.

Probably the biggest question not answered to date is how so much mRNA from genes expressed in the mother cell shows up in dormant spores. Certainly, contamination in dormant spores' outer layers has been thoroughly ruled out (Korza et al., 2019). However, there is known to be a feeding tube between the mother cell and forespore through which small molecules can pass to "feed" the developing spore (**Figure 2**). Perhaps fragments of mother cell mRNA can also pass into the developing spore as was previously suggested (Segev et al., 2012). Indeed, analysis of the coverage of RNA-Seq data for high and low abundance mRNAs in *B. subtilis* spores indicates that while coverage of reads of high abundance mRNAs is reasonably uniform, coverage of reads for low abundance mRNAs is somewhat less uniform, consistent with at least some of these mRNAs being in fragments. Indeed, direct evidence for the fragmented nature of at least some spore mRNAs was obtained recently by examining spore mRNAs for an NAD residue incorporated at mRNAs 5'-termini (Craft et al., 2020).

ISSUES FOR THE FUTURE

Given all the material above, the most likely assumptions now about spore mRNA seem to be as follows: (1) this mRNA is not important for directing synthesis of proteins needed for spore germination or outgrowth and (2) the role of this mRNA is to serve as a ribonucleotide reserve for new RNA synthesis when spores resume macromolecular synthesis. The available data on spore ATP levels also make it extremely unlikely that there can be significant levels of new RNA synthesis in dormant spore populations. However, given the well-known heterogeneity in spore populations (Setlow et al., 2012, 2017), it is certainly possible that a small percentage of dormant spores could be active in RNA synthesis from time to time, and disproving such possible heterogeneity in dormant spore populations will be extremely difficult. Perhaps going forward it may become possible to conduct transcriptomic analyses on individual dormant spores. Alternatively, small fractions of spores within dormant populations have been observed to initiate germination in a seemingly stochastic manner that reflects phenotypic rather than genetic diversity (Paidhungat and Setlow, 2000; Sturm and Dworkin, 2015). It is plausible then that these "spontaneously" germinating spores account for detectable mRNA synthesis in otherwise dormant spore populations, particularly if sensitive assays are employed.

Another topic touched upon in this review is whether RNA and/or protein synthesis are needed for spore germination. Most, but not all evidence supports an answer of no to this question, with the lack of the inhibition of spore germination by inhibitors of RNA or protein synthesis seen by most workers being potentially definitive results. However, it is well-known that the dormant spore inner membrane (IM), at least of *B. subtilis*, is impermeable to charged compounds, and exhibits only very slow permeability to neutral lipophilic compounds and even water (Gerhardt and Black, 1961; Knudsen et al., 2016). Thus, it seems most likely that charged and hydrophilic inhibitors of macromolecular synthesis would not get into the spore core until after germination is complete, including the hydrolysis of the spore peptidoglycan cortex, when IM lipid mobility and permeability return to that of growing cells. Indeed, there is no direct evidence for uptake of inhibitors of RNA or protein synthesis into an intact dormant spore, including when various treatments have been used in hopes of increasing spore permeability (Tanimoto et al., 1996; Sinai et al., 2015). Thus, to make this inhibitor experiment definitive, new ways must be developed to ensure that inhibitors indeed are able to freely enter the spore core. The recent demonstration of the high apparent IM permeability of decoated, CaDPA-less spores which retain characteristics of dormant spores (Mokashi et al., 2020) suggests that these spores could be useful for this type of experiment.

The second major evidence against any essential involvement of macromolecular synthesis in spore germination is that ATP

levels are almost non-existent in dormant spores, and appear to increase only after germination is complete (Setlow and Kornberg, 1970a; Scott and Ellar, 1978). While this result is potentially definitive in ruling out macromolecular synthesis' involvement in spore germination, this experiment can be confounded to some extent by at least three factors – (i) the heterogeneity in germination rates between individual spores in populations (Setlow et al., 2017), so that curves for both germination and ATP accumulation are broadened significantly, (ii) completion of germination after CaDPA release requires hydrolysis of the spore cortex peptidoglycan, which leads to an increase in core water content to 80% of wet wt allowing full protein mobility in spores and thus metabolic activity; however, the latter increase in core water is not rapid and takes 10–15 min in *B. subtilis* (Camilleri et al., 2019), and perhaps there could be metabolic activity when core water content reaches 60% of wet wt, and (iii) analyses of the accumulation of ATP by germinating spores has always been measured by extraction of small molecules and subsequent analyses and it would be ideal to measure molecules such as ATP, NADH, etc., in individual germinating spores, as this would give much higher resolution for analyses of when ATP and other high energy compounds were generated in spore germination. Indeed, certainly our understanding of spore germination has been significantly increased by examining the germination of multiple individual spores (Setlow et al., 2017). It certainly seems possible that use of new technology to examine individual germinating spore properties could give further insight into events early in germination and with much better time resolution than we currently have.

Finally, one of the most puzzling findings to come from detailed analysis of mRNAs in dormant spores is that a number of these mRNAs are thought to be expressed only in the mother cell compartment of this sporulating cell under the control of the mother cell compartment-specific σ^E or σ^K factors for RNA polymerase. It is certainly possible that there is some small amount of either or both of these σ-factors in forespores, although expression of such mother cell proteins has not been observed, and indeed would seem to serve no purpose, especially for mRNAs that encode spore coat or exosporium proteins. Readthrough transcription from an upstream gene under control of a forespore promoter may also be a contributing factor here although it seems unlikely that such a mechanism would be relied upon to confer molecules with functional significance. Another possibility suggested ~8 years ago (Segev et al., 2012) is that mother cell mRNAs might migrate into the developing spore *via* the "feeding tube" through which the mother cell nurtures the forespore (Camp and Losick, 2008; Meisner et al., 2008; Camp and Losick, 2009; Crawshaw et al., 2014; Zeytuni and Strynadka, 2019). However, if these mRNAs were intact, this would again seem to raise the concern of expressing mother cell protein in the spore core. An obvious alternative is that the mother cell mRNAs are only fragments, which might even facilitate their passage through the feeding tube. Indeed, many spore mRNAs appear to be significantly fragmented as determined by both analysis of the coverage of reads along a gene (Camilleri et al., 2019; Korza et al., 2019), as well as isolation of 5'-modified regions of a few *B. subtilis* spore mRNAs and demonstration that 3'-regions of these mRNAs were not isolated with 5' ends (Craft et al., 2020). This analysis, as well as the mechanism of mRNA (as well as rRNA) hydrolysis in spores and which does not give rise to mononucleotides (Korza et al., 2016) are also subjects for further work, and will almost certainly give new and interesting information on what is going on inside developing spores, and even dormant spores themselves.

AUTHOR CONTRIBUTIONS

PS and GC contributed equally to the writing and preparation of this article. Both the authors contributed to the article and approved the submitted version.

FUNDING

Work in the Setlow lab has been supported by grants and contracts from agencies in the USA including the Army Research Office, Defense Threat Reduction Agency, Defense Advanced Products Research Agency, Department of Agriculture and the National Institutes of Health. Work in the Christie lab has been supported by various UK Research Council and industry partners, including the Biotechnology and Biological Sciences Research Council, the Engineering and Physical Sciences Research Council, the Defence Science and Technology Laboratory, The Royal Society, and the UK Home Office.

ACKNOWLEDGMENTS

We are grateful to Chris Doona, Wendy Mok, Melissa Caimano, Ralf Moeller, Barbara Setlow, and Stanley Brul for comments on the manuscript.

REFERENCES

Arrieta-Ortiz, M. L., Hafemeister, C., Bate, A. R., Chu, T., Greenfield, A., Shuster, B., et al. (2015). An experimentally supported model of the *Bacillus subtilis* global transcriptional regulatory network. *Mol. Syst. Biol.* 11:839. doi: 10.15252/msb.20156236

Bassi, D., Cappa, F., and Cocconcelli, P. S. (2013). Array-based transcriptional analysis of *Clostridium sporogenes* UC9000 during germination, cell outgrowth and vegetative life. *Food Microbiol.* 33, 11–23. doi: 10.1016/j.fm.2012.08.004

Bassi, D., Colla, F., Gazzola, S., Puglisi, E., Delledonne, M., and Cocconcelli, P. S. (2016). Transcriptome analysis of *Bacillus thuringiensis* spore life, germination and cell outgrowth in a vegetable-based food model. *Food Microbiol.* 55, 73–85. doi: 10.1016/j.fm.2015.11.006

Bate, A. R., Bonneau, R., and Eichenberger, P. (2014). *The bacterial spore: From molecules to systems.* eds. A. Driks, and P. Eichenberger (Washington, DC: American Society for Microbiology), 397.

Bergman, N. H., Anderson, E. C., Swenson, E. E., Niemeyer, M. M., Miyoshi, A. D., and Hanna, P. C. (2006). Transcriptional profiling of the *Bacillus anthracis* life cycle in vitro and an implied model for regulation of spore formation. *J. Bacteriol.* 188, 6092–6100. doi: 10.1128/JB.00723-06

Bettegowda, C., Huang, X., Lin, J., Cheong, I., Kohli, M., Szabo, S. A., et al. (2006). The genome and transcriptomes of the anti-tumor agent

Clostridium novyi-NT. *Nat. Biotechnol.* 24, 1573–1580. doi: 10.1038/nbt1256

Brown, S., and Coleman, G. (1975). Messenger ribonucleic acid content of *Bacillus amyloliquefaciens* throughout its growth cycle compared with *Bacillus subtilis* 168. *J. Mol. Biol.* 96, 345–352.

Buhr, T. L., Minter, Z. A., Kennihan, N. L., Young, A. A., Borgers-Klonkowski, E. L., Osborn, E. B., et al. (2020). Combining spore germination and heat inactivation to decontaminate materials contaminated with *Bacillus anthracis* spores. *J. Appl. Microbiol.* 128, 124–137. doi: 10.1111/jam.14474

Camilleri, E., Korza, G., Green, J., Yuan, J., Li, Y. Q., Caimano, M. J., et al. (2019). Properties of aged spores of *Bacillus subtilis*. *J. Bacteriol.* 201, 1–17. doi: 10.1128/JB.00231-19

Camp, A. H., and Losick, R. (2008). A novel pathway of intercellular signalling in *Bacillus subtilis* involves a protein with similarity to a component of type III secretion channels. *Mol. Microbiol.* 69, 402–417. doi: 10.1111/j.1365-2958.2008.06289.x

Camp, A. H., and Losick, R. (2009). A feeding tube model for activation of a cell-specific transcription factor during sporulation in *Bacillus subtilis*. *Genes Dev.* 23, 1014–1024. doi: 10.1101/gad.1781709

Chambon, P., Deutscher, M. P., and Kornberg, A. (1968). Biochemical studies of bacterial sporulation and germination. X. Ribosomes and nucleic acids of vegetative cells and spores of *Bacillus megaterium*. *J. Biol. Chem.* 243, 5110–5116.

Chapman, A. G., Fall, L., and Atkinson, D. E. (1971). Adenylate energy charge in *Escherichia coli* during growth and starvation. *J. Bacteriol.* 108, 1072–1086.

Cowan, A. E., Olivastro, E. M., Koppel, D. E., Loshon, C. A., Setlow, B., and Setlow, P. (2004). Lipids in the inner membrane of dormant spores of *Bacillus* species are largely immobile. *Proc. Natl. Acad. Sci. U. S. A.* 101, 7733–7738. doi: 10.1073/pnas.0306859101

Craft, D. L., Korza, G., Zhang, Y., Frindert, J., Jäschke, A., Caimano, M. J., et al. (2020). Analysis of 5'-NAD capping of abundant mRNAs in dormant spores of *Bacillus subtilis*. *FEMS Microbiol. Lett.* 367:fnaa143. doi: 10.1093/femsle/fnaa143

Crawshaw, A. D., Serrano, M., Stanley, W. A., Henriques, A. O., and Salgado, P. S. (2014). A mother cell-to-forespore channel: current understanding and future challenges. *FEMS Microbiol. Lett.* 358, 129–136. doi: 10.1111/1574-6968.12554

Dembek, M., Stabler, R. A., Witney, A. A., Wren, B. W., and Fairweather, N. F. (2013). Transcriptional analysis of temporal gene expression in germinating *Clostridium difficile* 630 endospores. *PLoS One* 8:e64011. doi: 10.1371/journal.pone.0064011

Deutscher, M. P., Chambon, P., and Kornberg, A. (1968). Biochemical studies of bacterial sporulation and germination. XI. Protein-synthesizing systems from vegetative cells and spores of *Bacillus megaterium*. *J. Biol. Chem.* 243, 5117–5125.

Feaga, H. A., Kopylov, M., Kim, J. K., Jovanovic, M., and Dworkin, J. (2020). Ribosome dimerization protects the small subunit. *J. Bacteriol.* 202, e00009–e00020. doi: 10.1128/JB.00009-20

Gerhardt, P., and Black, S. H. (1961). Permeability of bacterial spores. II. Molecular variables affecting solute permeation. *J. Bacteriol.* 82, 750–760.

Gerhardt, P., and Marquis, R. E. (1989). "Spore thermoresistance mechanisms" in *Regulation of prokaryotic development*. eds. I. Smith, R. A. Slepecky and P. Setlow (Washington, DC: American Society of Microbiology), 43–63.

Ghosh, S., Korza, G., Maciejewski, M., and Setlow, P. (2015). Analysis of metabolism in dormant spores of *Bacillus* species by 31P nuclear magnetic resonance analysis of low-molecular-weight compounds. *J. Bacteriol.* 197, 992–1001. doi: 10.1128/JB.02520-14

Hauser, P. M., and Karamata, D. (1992). A method for the determination of bacterial spore DNA content based on isotopic labelling, spore germination and diphenylamine assay; ploidy of spores of several *Bacillus* species. *Biochimie* 74, 723–733.

Hayes, C. S., and Setlow, P. (2001). An alpha/beta-type, small, acid-soluble spore protein which has very high affinity for DNA prevents outgrowth of *Bacillus subtilis* spores. *J. Bacteriol.* 183, 2662–2666. doi: 10.1128/JB.183.8.2662-2666.2001

Hosack, D. A., Dennis, G. Jr., Sherman, B. T., Lane, H. C., and Lempicki, R. A. (2003). Identifying biological themes within lists of genes with EASE. *Genome Biol.* 4:R70. doi: 10.1186/gb-2003-4-10-r70

Jeng, Y. H., and Doi, R. H. (1974). Messenger ribonucleic acid of dormant spores of *Bacillus subtilis*. *J. Bacteriol.* 119, 514–521.

Jones, S. W., Paredes, C. J., Tracy, B., Cheng, N., Sillers, R., Senger, R. S., et al. (2008). The transcriptional program underlying the physiology of clostridial sporulation. *Genome Biol.* 9:R114. doi: 10.1186/gb-2008-9-7-r114

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* 45, D353–d361. doi: 10.1093/nar/gkw1092

Keijser, B. J., Ter Beek, A., Rauwerda, H., Schuren, F., Montijn, R., Van Der Spek, H., et al. (2007). Analysis of temporal gene expression during Bacillus subtilis spore germination and outgrowth. *J. Bacteriol.* 189, 3624–3634. doi: 10.1128/JB.01736-06

Knudsen, S. M., Cermak, N., Delgado, F. F., Setlow, B., Setlow, P., and Manalis, S. R. (2016). Water and small-molecule permeation of dormant *Bacillus subtilis* spores. *J. Bacteriol.* 198, 168–177. doi: 10.1128/JB.00435-15

Korza, G., Abini-Agbomson, S., Setlow, B., Shen, A., and Setlow, P. (2017). Levels of L-malate and other low molecular weight metabolites in spores of *Bacillus* species and *Clostridium difficile*. *PLoS One* 12:e0182656. doi: 10.1371/journal.pone.0182656

Korza, G., Camilleri, E., Green, J., Robinson, J., Nagler, K., Moeller, R., et al. (2019). Analysis of the mRNAs in spores of *Bacillus subtilis*. *J. Bacteriol.* 201, e00007–e00019. doi: 10.1128/JB.00007-19

Korza, G., Setlow, B., Rao, L., Li, Q., and Setlow, P. (2016). Changes in *Bacillus* spore small molecules, rRNA, germination, and outgrowth after extended sublethal exposure to various temperatures: evidence that protein synthesis is not essential for spore germination. *J. Bacteriol.* 198, 3254–3264. doi: 10.1128/jb.00583-16

Levinson, H. S., and Hyatt, M. T. (1966). Sequence of events during *Bacillus megaterium* spore germination. *J. Bacteriol.* 91, 1811–1818.

Meisner, J., Wang, X., Serrano, M., Henriques, A. O., and Moran, C. P. Jr. (2008). A channel connecting the mother cell and forespore during bacterial endospore formation. *Proc. Natl. Acad. Sci. U. S. A.* 105, 15100–15105. doi: 10.1073/pnas.0806301105

Midgley, J. E. (1969). The messenger ribonucleic acid content of *Bacillus subtilis* 168. *Biochem. J.* 115, 171–181.

Mokashi, S., Kanaan, J., Craft, D. L., Byrd, B., Zenick, B., Laue, M., et al. (2020). Killing of bacterial spores by dodecylamine and its effects on spore inner membrane properties. *J. Appl. Microbiol.* doi: 10.1111/jam.14732 [Epub ahead of print]

Nagler, K., Krawczyk, A. O., De Jong, A., Madela, K., Hoffmann, T., Laue, M., et al. (2016). Identification of differentially expressed genes during *Bacillus subtilis* spore outgrowth in high-salinity environments using RNA sequencing. *Front. Microbiol.* 7:1564. doi: 10.3389/fmicb.2016.01564

Nelson, D. L., and Kornberg, A. (1970a). Biochemical studies of bacterial sporulation and germination. XIX. Phosphate metabolism during sporulation. *J. Biol. Chem.* 245, 1137–1145.

Nelson, D. L., and Kornberg, A. (1970b). Biochemical studies of bacterial sporulation and germination. XX. Phosphate metabolism during germination. *J. Biol. Chem.* 245, 1146–1155.

Nicholson, W. L., Moeller, R., and Horneck, G. (2012). Transcriptomic responses of germinating *Bacillus subtilis* spores exposed to 1.5 years of space and simulated martian conditions on the EXPOSE-E experiment PROTECT. *Astrobiology* 12, 469–486. doi: 10.1089/ast.2011.0748

Paidhungat, M., and Setlow, P. (2000). Role of Ger proteins in nutrient and nonnutrient triggering of spore germination in *Bacillus subtilis*. *J. Bacteriol.* 182, 2513–2519. doi: 10.1128/jb.182.9.2513-2519.2000

Ramírez-Guadiana, F. H., Meeske, A. J., Rodrigues, C. D. A., Barajas-Ornelas, R. D. C., Kruse, A. C., and Rudner, D. Z. (2017). A two-step transport pathway allows the mother cell to nurture the developing spore in *Bacillus subtilis*. *PLoS Genet.* 13:e1007015. doi: 10.1371/journal.pgen.1007015

Scott, I. R., and Ellar, D. J. (1978). Metabolism and the triggering of germination of *Bacillus megaterium*. Concentrations of amino acids, organic acids, adenine nucleotides and nicotinamide nucleotides during germination. *Biochem. J.* 174, 627–634.

Segev, E., Smith, Y., and Ben-Yehuda, S. (2012). RNA dynamics in aging bacterial spores. *Cell* 148, 139–149. doi: 10.1016/j.cell.2011.11.059

Setlow, P. (1974). Percent charging of transfer ribonucleic acid and levels of ppGpp and pppGpp in dormant and germinated spores of *Bacillus megaterium*. *J. Bacteriol.* 118, 1067–1074.

Setlow, P. (2006). Spores of *Bacillus subtilis*: their resistance to and killing by radiation, heat and chemicals. *J. Appl. Microbiol.* 101, 514–525. doi: 10.1111/j.1365-2672.2005.02736.x

Setlow, P. (2007). I will survive: DNA protection in bacterial spores. *Trends Microbiol.* 15, 172–180. doi: 10.1016/j.tim.2007.02.004

Setlow, P., and Johnson, E. A. (2019). "Spores and their significance" in *Food microbiology, fundamentals and frontiers. 5th Edn.* eds. M. P. Doyle, F. Diez-Gonzalez and C. Hill. (Washington, DC: ASM Press), 23–64.

Setlow, P., and Kornberg, A. (1970a). Biochemical studies of bacterial sporulation and germination. XXIII. Energy metabolism during early stages of germination of spores of *Bacillus megaterium*. *J. Biol. Chem.* 245, 3637–3644.

Setlow, P., and Kornberg, A. (1970b). Biochemical studies of bacterial sporulation and germination. XXIII. Nucleotide metabolism during spore germination. *J. Biol. Chem.* 245, 3645–3652.

Setlow, P., Liu, J., and Faeder, J. R. (2012). "Heterogeneity in bacterial spore populations" in *Bacterial spores: Current research and applications*. ed. E. Abel-Santos (Norwich, UK: Horizon Scientific Press), 201–216.

Setlow, P., and Primus, G. (1975). Protein metabolism during germination of *Bacillus megaterium* spores. I. Protein synthesis and amino acid metabolism. *J. Biol. Chem.* 250, 623–630.

Setlow, P., Primus, G., and Deutscher, M. P. (1974). Absence of 3'-terminal residues from transfer ribonucleic acid of dormant spores of *Bacillus megaterium*. *J. Bacteriol.* 117, 126–132.

Setlow, B., and Setlow, P. (1977). Levels of oxidized and reduced pyridine nucleotides in dormant spores and during growth, sporulation, and spore germination of *Bacillus megaterium*. *J. Bacteriol.* 129, 857–865.

Setlow, B., Shay, L. K., Vary, J. C., and Setlow, P. (1977). Production of large amounts of acetate during germination of *Bacillus megaterium* spores in the absence of exogenous carbon sources. *J. Bacteriol.* 132, 744–746.

Setlow, P., Wang, S., and Li, Y. Q. (2017). Germination of spores of the orders Bacillales and Clostridiales. *Annu. Rev. Microbiol.* 71, 459–477. doi: 10.1146/annurev-micro-090816-093558

Shen, A., Edwards, A. N., Sarker, M. R., and Paredes-Sabja, D. (2019). Sporulation and germination in *Clostridial pathogens*. *Microbiol. Spectr.* 7. doi: 10.1128/microbiolspec.GPP3-0017-2018

Sinai, L., Rosenberg, A., Smith, Y., Segev, E., and Ben-Yehuda, S. (2015). The molecular timeline of a reviving bacterial spore. *Mol. Cell* 57, 695–707. doi: 10.1016/j.molcel.2014.12.019

Steinberg, W., Halvorson, H. O., Keynan, A., and Weinberg, E. (1965). Timing of protein synthesis during germination and outgrowth of spores of *Bacillus cereus* strain T. *Nature* 208, 710–711.

Sturm, A., and Dworkin, J. (2015). Phenotypic diversity as a mechanism to exit cellular dormancy. *Curr. Biol.* 25, 2272–2277. doi: 10.1016/j.cub.2015.07.018

Swarge, B. N., Abhyankar, W. R., Jonker, M., Hoefsloot, H., Kramer, G., Setlow, P., et al. (2020a). Integrative analysis of proteome and transcriptome dynamics during *Bacillus subtilis* spore revival. *mSphere* 5, e00463–e00520. doi: 10.1128/mSphere.00463-20

Swarge, B. N., Nafid, C., Vischer, N., Kramer, G., Setlow, P., and Brul, S. (2020b). Investigating synthesis of the MalS malic enzyme during *Bacillus subtilis* spore germination and outgrowth and the influence of spore maturation and sporulation conditions. *mSphere* 5, e00464–e00520. doi: 10.1128/mSphere.00464-20

Swarge, B. N., Roseboom, W., Zheng, L., Abhyankar, W. R., Brul, S., De Koster, C. G., et al. (2018). "One-Pot" sample processing method for proteome-wide analysis of microbial cells and spores. *Proteomics Clin. Appl.* 12:e1700169. doi: 10.1002/prca.201700169

Tan, I. S., and Ramamurthi, K. S. (2014). Spore formation in *Bacillus subtilis*. *Environ. Microbiol. Rep.* 6, 212–225. doi: 10.1111/1758-2229.12130

Tanimoto, Y., Ichikawa, Y., Yasuda, Y., and Tochikubo, K. (1996). Permeability of dormant spores of *Bacillus subtilis* to gramicidin S. *FEMS Microbiol. Lett.* 136, 151–156.

Torriani, A., and Levinthal, C. (1967). Ordered synthesis of proteins during outgrowth of spores of *Bacillus cereus*. *J. Bacteriol.* 94, 176–183.

Ulrich, N., Nagler, K., Laue, M., Cockell, C. S., Setlow, P., and Moeller, R. (2018). Experimental studies addressing the longevity of *Bacillus subtilis* spores – The first data from a 500-year experiment. *PLoS One* 13:e0208425. doi: 10.1371/journal.pone.0208425

van Melis, C. C., Nierop Groot, M. N., Tempelaars, M. H., Moezelaar, R., and Abee, T. (2011). Characterization of germination and outgrowth of sorbic acid-stressed *Bacillus cereus* ATCC 14579 spores: phenotype and transcriptome analysis. *Food Microbiol.* 28, 275–283. doi: 10.1016/j.fm.2010.04.005

Zeytuni, N., and Strynadka, N. C. J. (2019). A hybrid secretion system facilitates bacterial sporulation: a structural perspective. *Microbiol. Spectr.* 7, 1–9. doi: 10.1128/microbiolspec.PSIB-0013-2018

Zhu, B., and Stülke, J. (2018). SubtiWiki in 2018: from genes and proteins to functional network annotation of the model organism *Bacillus subtilis*. *Nucleic Acids Res.* 46, D743–D748. doi: 10.1093/nar/gkx908

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 26 November 2020
doi: 10.3389/fmicb.2020.604458

6S-1 RNA Contributes to Sporulation and Parasporal Crystal Formation in *Bacillus thuringiensis*

Zhou Li[1], Li Zhu[1], Zhaoqing Yu[1], Lu Liu[1], Shan-Ho Chou[1], Jieping Wang[2] and Jin He[1]**

[1]State Key Laboratory of Agricultural Microbiology, College of Life Science and Technology, Huazhong Agricultural University, Wuhan, China, [2]Agricultural BioResources Institute, Fujian Academy of Agricultural Sciences, Fuzhou, China

6S RNA is a kind of high-abundance non-coding RNA that globally regulates bacterial transcription by interacting with RNA polymerase holoenzyme. Through bioinformatics analysis, we found that there are two tandem 6S RNA-encoding genes in the genomes of *Bacillus cereus* group bacteria. Using *Bacillus thuringiensis* BMB171 as the starting strain, we have explored the physiological functions of 6S RNAs, and found that the genes *ssrSA* and *ssrSB* encoding 6S-1 and 6S-2 RNAs were located in the same operon and are co-transcribed as a precursor that might be processed by specific ribonucleases to form mature 6S-1 and 6S-2 RNAs. We also constructed two single-gene deletion mutant strains Δ*ssrSA* and Δ*ssrSB* and a double-gene deletion mutant strain Δ*ssrSAB* by means of the markerless gene knockout method. Our data show that deletion of 6S-1 RNA inhibited the growth of *B. thuringiensis* in the stationary phase, leading to lysis of some bacterial cells. Furthermore, deletion of 6S-1 RNA also significantly reduced the spore number and parasporal crystal content. Our work reveals that *B. thuringiensis* 6S RNA played an important regulatory role in ensuring the sporulation and parasporal crystal formation.

Keywords: non-coding RNA, 6S RNA, *Bacillus thuringiensis*, sporulation, parasporal crystal formation

OPEN ACCESS

Edited by:
Jiqiang Ling,
University of Maryland, College Park, United States

Reviewed by:
Ma. Cristina Del Rincón-Castro,
University of Guanajuato, Mexico
Jia Yin,
Hunan Normal University, China

****Correspondence:***
Jin He
hejin@mail.hzau.edu.cn
Jieping Wang
wangjpfaas@foxmail.com

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *09 September 2020*
Accepted: *04 November 2020*
Published: *26 November 2020*

Citation:
Li Z, Zhu L, Yu Z, Liu L, Chou S-H, Wang J and He J (2020) 6S-1 RNA Contributes to Sporulation and Parasporal Crystal Formation in Bacillus thuringiensis. Front. Microbiol. 11:604458. doi: 10.3389/fmicb.2020.604458

INTRODUCTION

In 1967, Hindley (1967) first discovered a highly abundant non-coding RNA in *Escherichia coli*. Because it exhibits a sedimentation coefficient of 6S, it was thus named 6S RNA, with its nucleotide sequence further determined by Brownlee (1971) afterwards. However, the study of 6S RNA remains silent for a long time until 2000, when Wassarman and Storz (2000) found that *E. coli* 6S RNA exists in high abundance throughout all bacterial growth phases, and exhibits a concentration as high as 10,000 molecules/cell in *E. coli* cells in the stationary phase. High abundant 6S RNA is found to regulate gene transcription in the stationary phase by combining with σ^{70}-denpendent RNA polymerase holoenzyme, thus leading to the response regulation of bacterial to stresses such as starvation. Infact, all previous studies also seem to indicate that 6S RNA is a global regulatory factor (Wassarman and Storz, 2000; Cavanagh and Wassarman, 2014; Steuten et al., 2014; Wassarman, 2018). Later on, 6S RNA is found to exist in even more bacteria. Wehner et al. (2014) identified 1,750 6S RNA-encoding genes in 1,611 bacterial genomes, and found that the sizes of 6S RNA-encoding

genes ranging from 153 to 237 base pairs (with an average of 183 bp). While most bacteria seem to contain only one 6S RNA-encoding gene in their genomes, some bacteria, such as *Bacillus* genus in *Firmicutes* phylum, do contain two 6S RNA-encoding genes. Interestingly, even four such 6S RNA-encoding genes are found in the genome of *Magnetococcus* sp. MC-1 of α-*Proteobacteria*. In fact, 6S RNA-encoding genes are now found to be widely distributed in various bacteria, even in bacteria from the phyla *Chloroflexi* and *Aquificae* at the root of the phylogenetic tree (Wehner et al., 2014), indicating that 6S RNAs are originated quite early and are widely distributed.

Besides, Trotochaud and Wassarman (2004) also reveal that the survival ability of *E. coli* is distinctively reduced in the stationary phase when the 6S RNA-encoding gene *ssrS1* is deleted, indicating that 6S RNA is beneficial to the normal growth and survival of *E. coli* in the stationary phase. Indeed, more and more regulatory functions of 6S RNA are gradually unveiled after many more in-depth researches. For example, it can now delay the sporulation of *Bacillus subtilis* (Cavanagh and Wassarman, 2013), promote the photosynthesis of *Synechocystis* (Heilmann et al., 2017), enhance the gene expression of pathogenicity island in *Salmonella enterica* serovar Typhimurium (Ren et al., 2017), increase the synthesis of antibiotics in *Streptomyces coelicolor* (Mikulík et al., 2014), optimize the symbiosis of *Bradyrhizobium* with leguminous plants (Madhugiri et al., 2012), and so on.

B. subtilis contains two kinds of 6S RNAs, called 6S-1 and 6S-2 RNAs, with their encoding genes *bsrA* and *bsrB* located in different regions of the genome. While deletion of 6S-1 RNA inhibits the growth of *B. subtilis* in the stationary phase (Hoch et al., 2015) and promotes earlier sporulation initiation by accelerating the utilization of nutrients (Cavanagh and Wassarman, 2013), the lack of 6S-2 RNA does not, however, seem to affect the growth and sporulation in the stationary phase (Cavanagh and Wassarman, 2013; Hoch et al., 2015). Until now, its physiological function remains unclear.

The *Bacillus cereus* group is a taxonomic group comprising closely related species of the *Bacillus genus* that includes more than 20 different species (Miller et al., 2016; Liu et al., 2017). Among them, the insect pathogen *Bacillus thuringiensis*, the anthracnose pathogen *Bacillus anthracis*, and the food-borne opportunistic pathogen *B. cereus* (Wang et al., 2019), have attracted extensive attention. Different from *B. subtilis* that harbors two 6S RNA-encoding genes located in different regions, *B. cereus* group bacteria contain the 6S RNA-encoding genes arranged in tandem in their genomes (Wehner et al., 2014). In this study, taking the *B. thuringiensis* BMB171 as a model to study the physiological function of 6S RNAs in *B. cereus* group bacteria, we found that the encoding genes *ssrSA* and *ssrSB* of 6S-1 and 6S-2 RNAs in BMB171 were located in the same operon and were co-transcribed as a precursor, which might be processed by ribonucleases to form mature 6S-1 and 6S-2 RNAs. Furthermore, deletion of the *ssrSA* gene inhibited the growth of *B. thuringiensis* in the stationary phase and decreased the sporulation and parasporal crystal formation. Besides, we found that diminished sporulation is primarily due to the decreased growth rate of the *ssrSA* deletion mutant in the stationary phase.

MATERIALS AND METHODS

Bacterial Strains and Growth Conditions

The plasmids and strains used in this study are listed in **Tables 1, 2**, with primers used listed in **Supplementary Table S1**, respectively. *Escherichia coli* DH5α was used for cloning experiment and was cultured at 37°C in lysogeny broth (LB) medium (g/L: tryptone, 10; yeast extract, 5; and NaCl, 10). The medium was adjusted to pH 7.0 before autoclaving at 121°C for 15 min. *Bacillus thuringiensis* BMB171 and its derivative strains were cultured at 28°C in the GYS medium (g/L: glucose, 1.00; yeast extract, 2.00; $K_2HPO_4{\cdot}3H_2O$, 0.66; $(NH_4)_2SO_4$, 2.00; $MgSO_4{\cdot}7H_2O$, 0.04; $MnSO_4{\cdot}H_2O$, 0.04; and $CaCl_2$, 0.08), with the medium autoclaved at 115°C for 30 min after pH value being adjusted to 7.8. When necessary, relevant antibiotics were added to the cultures with the following final concentration: 50 μg/ml kanamycin, 25 μg/ml erythromycin, 100 μg/ml ampicillin, 300 μg/ml spectinomycin, or 60 U polymyxin (Wang et al., 2019). For the decoyinine-added experiment, cells were grown to an OD_{600} of 0.5 in GYS medium, followed by the addition of decoyinine to 0.4 mg/ml as previously described (Mitani et al., 1977; Ikehara et al., 1982; Cavanagh and Wassarman, 2013).

RNA Extraction and RT-qPCR

The samples of 30 ml each from BMB171 and its derivative strains were cultured in GYS medium for 11 h before being centrifuged, followed by total RNA extraction and RT-qPCR experiments as previously described (Zheng et al., 2015, 2020; Fu et al., 2018). In these experiments, the *gapdh* gene was used as an internal control.

Identification of Transcription Start Site

The 5'-rapid amplification of complementary DNA (cDNA) ends (5'-RACE) experiment was performed to identity the trancription start site (TSS) as described previously with some modifications (Ali et al., 2017). RNA was first extracted from BMB171 cells that were grown in GYS, followed by reverse transcription to cDNA. The 3'-end of cDNA was then labeled by poly(dA) using terminal deoxynucleotidyl transferase (Takara, Japan). The cDNA was then PCR amplified using primers of Primer-8 and *ssrSB*-R as listed in **Supplementary Table S1**. The PCR products were then cloned to the pMD19-T vector (Takara, Japan) and sequenced (Wang et al., 2019).

Determination of β-Galactosidase Activity

BMB171/pHT1K, BMB171/P1-*lacZ*, and BMB171/P2-*lacZ* strains were grown at 28°C in a shaking incubator at 200 rpm in 100 ml GYS with 25 μg/ml erythromycin. Two milliliters of each culture was separately collected at indicated time and assayed for β-galactosidase activity as described previously (Zhou et al., 2017; Wang et al., 2019).

Spore Count by Spread-Plate Method

BMB171 and its single deletion 6S-1 RNA mutant Δ*ssrSA*, single deletion 6S-2 RNA mutant Δ*ssrSB*, and the double deletion mutant

TABLE 1 | Plasmids used in this study.

Plasmids	Relevant characteristics	Purposes	Origins
pHT1K	*B. thuringiensis-E. coli* shuttle plasmid; Amp^R Erm^R	For β-galactosidase assays and gene complementation	Wang et al., 2013a
pHT1K-*lacZ*	pHT1K plasmid carrying the promoter-less *lacZ* gene, for β-galactosidase activity assay	β-galactosidase assays	Wang et al., 2013a
pHT1K-P1-*lacZ*	pHT1K carrying the promoter region (−190 to +10) of *ssrS* operon fused with *lacZ*	β-galactosidase assays	This work
pHT1K-P2-*lacZ*	pHT1K carrying the upstream region (+10 to +214) of *ssrSB* fused with *lacZ*	β-galactosidase assays	This work
pHT1K-P1-*ssrSA*	pHT1K carrying the promoter and *ssrSA* encoding region of *ssrS* operon fused with terminator region of *ssrS* operon	Gene complementation	This work
pHT1K-P1-*ssrSAB*	pHT1K carrying the promoter and encoding region of the *ssrS* operon	Gene complementation	This work
pRP1028	*B. thuringiensis-E. coli* shuttle plasmid; Amp^R Erm^R; carrying *turbo-rfp* gene and an I-*Sce*I recognition site	Gene deletion	Janes and Stibitz, 2006
pSS4332	*B. thuringiensis-E. coli* shuttle plasmid; Km^R; carrying *gfp* and I-*Sce*I restriction enzyme encoding gene	Gene deletion	Janes and Stibitz, 2006
pSS1827	The helper plasmid for conjugative transfer; Amp^R	Gene deletion	Janes and Stibitz, 2006
pRP1028-*ssrSA*-UD	pRP1028 with the upstream and downstream regions of *ssrSA*, an intermediate plasmid in gene deletion experiment	Gene deletion	This work
pRP1028-*ssrSB*-UD	pRP1028 with the upstream and downstream regions of *ssrSB*, an intermediate plasmid in gene deletion experiments	Gene deletion	This work
pRP1028-*ssrSAB*-UD	pRP1028 with the upstream region of *ssrSA* and downstream region of *ssrSB*, an intermediate plasmid in gene deletion experiments	Gene deletion	This work
pBMB43-304	*B. thuringiensis-E. coli* shuttle plasmid; $Amp^R Erm^R$; carrying ORF of *cry1Ac10*;	Determination of the parasporal crystal protein Cry1Ac 10	Qi et al., 2015

TABLE 2 | Strains used in this study.

Strains	Relevant characteristics	Origins
E. coli DH5α	F-Φ80*lacZ*ΔM15 Δ(*lacZYA-argF*) U169 *recA*1 *endA*1 *hsdR*17 (rk-, mk+) *phoA supE*44 *thi*-1 *gyrA*96 *relA*1 λ-	Beijing TransGen Biotech Co., Ltd.
BMB171	*B. thuringiensis* strain BMB171; an acrystalliferous mutant strain; high transformation frequency	He et al., 2010; Wang et al., 2013a
Δ*ssrSA*	Markerless deletion of *ssrSA* in BMB171	This work
Δ*ssrSB*	Markerless deletion of *ssrSB* in BMB171	This work
Δ*ssrSAB*	Markerless deletion of *ssrSA* and *ssrSB* in BMB171	This work
BMB171/pHT1K	BMB171 strain harboring plasmid pHT1K	This work
Δ*ssrSA*/P1-*ssrSA*	Gene complemented strain: Δ*ssrSA* strain harboring plasmids pHT1K-P1-*ssrSA*	This work
Δ*ssrSAB*/P1-*ssrSAB*	Gene complemented strain: Δ*ssrSAB* strain harboring plasmid pHT1K-P1-*ssrSAB*	This work
BMB171/*lacZ*	BMB171 strain harboring pHT1K carrying the promoter-less *lacZ* gene	This work
BMB171/P1-*lacZ*	BMB171 strain harboring pHT1K carrying the promoter region of *ssrSA* fused with *lacZ*	This work
BMB171/P2-*lacZ*	BMB171 strain harboring pHT1K carrying the upstream region of *ssrSB* fused with *lacZ*	This work
BMB171-*cry*	BMB171 strain harboring plasmid pBMB43-304	This work
Δ*ssrSA-cry*	Δ*ssrSA* strain harboring plasmid pBMB43-304	This work
Δ*ssrSB-cry*	Δ*ssrSB* strain harboring plasmid pBMB43-304	This work
Δ*ssrSAB-cry*	Δ*ssrSAB* strain harboring plasmid pBMB43-304	This work

Δ*ssrSAB* were cultured in GYS medium at 28°C for 24 h, which were then heated to 65°C for 30 min, followed by gradient dilution (10 times) with M9 minimum medium. Around 100 μl of each diluent was then spread onto LB plates. The colony-forming units (CFUs) per ml were then counted (Wang et al., 2016).

Phase-Contrast Microscopic Analysis of Sporulation

BMB171 and its derivative strains were cultured at 28°C in GYS medium. To observe the morphology of vegetative cells and spores, 5 μl of each cell sample was collected at indicated time points, spotted onto the center of a glass slide, and covered with a coverslip. Spores were then observed with a phase-contrast microscope (Olympus, Japan; Zheng et al., 2020).

Transmission Electron Microscope

To observe the morphology of BMB171 cells and its derivative strains, 4 ml of each sample was harvested by centrifugation at 17 h, with the cell pellets resuspended in 2.5% glutaraldehyde, and stored at 4°C overnight. Ultra-thin sections were finally prepared and stained as described (Craig et al., 1997). A Hitachi H-7000 FA transmission electron microscope (Hitachi, Japan) was then used for observation.

Construction of Markerless Gene Deletion Strains

The markerless gene deletion mediated by homing endonuclease I-*Sce*I was performed in *B. thuringiensis* as previously reported (Zheng et al., 2015; Tang et al., 2016; Wang et al., 2016).

Intermediate plasmids (pRP1028-*ssrSA*-UD, pRP1028-*ssrSB*-UD, and pRP1028-*ssrSAB*-UD) used in this study for gene deletion experiments were listed in **Table 1**. The helper plasmid pSS1827 for conjugational transfer and *B. thuringiensis-E. coli* shuttle plasmid pRP1028 carrying an I-*Sce*I recognition site and pSS4332 carrying I-*Sce*I restriction enzyme encoding gene was listed in **Table 1**.

Observation of Parasporal Crystal and Determination of Parasporal Crystal Protein

Crystalliferous strains of BMB171-*cry*, Δ*ssrSA-cry*, Δ*ssrSB-cry*, and Δ*ssrSAB-cry* were obtained by transformation of the *cry1Ac10* gene with its original promoter in the plasmid pBMB43-304 (Qi et al., 2015) into BMB171, Δ*ssrSA*, Δ*ssrSB*, and Δ*ssrSAB* strains, respectively. To observe parasporal crystals, the crystalliferous strains were grown at 28°C and 200 rpm for 24 h in GYS medium supplemented with 25 mg/ml erythromycin. One drop from each culture was spotted onto the center of a glass slide, and covered with a coverslip. Parasporal crystals were then observed with a phase-contrast microscope (Olympus, Japan). To extract Cry1Ac10 protein, each culture was collected separately by centrifugation at 6,000 *g* for 15 min (AG Eppendorf, Hamburg, Germany). Procedure for the separation of Cry1Ac10 protein was carried out according to a previous study (Wang et al., 2013b). Finally, the Cry1Ac10 protein was visualized by SDS-PAGE, with its concentration measured by the Bradford method (Wang et al., 2016).

RESULTS

The *ssrSA and ssrSB* Genes Were Located in the Same Operon and Were Co-Transcribed

The BMB171 strain was found to possess two 6S RNAs, with their encoding genes *ssrSA* (*BMB171_RS29145*) and *ssrSB* (*BMB171_RS291506*) located in tandem in the genome (**Figure 1A**). To explore whether the *ssrSA* and *ssrSB* genes are within the same operon, we first verified the co-transcription of the *ssrSA* and *ssrSB* genes by the semi-quantitative reverse transcription PCR (SqRT-PCR; **Supplementary Figure S1**), and found they were indeed co-transcribed. To further confirm that the *ssrSB* gene is not transcribed individually, we carried out the β-galactosidase assays to detect the promoter activities of the upstream regions of *ssrSA* and *ssrSB* genes in the BMB171 genome, respectively. We found that the upstream sequence of the *ssrSA* gene (P1 region) exhibited strong transcription initiation activities in the logarithmic phase (5 h), transition phase (11 h), and stationary phase (17 h), while those of the *ssrSB* gene (P2 region) basically exhibited no such transcription initiation activities at all when compared to the no promoter control (**Figure 1B**). These data indicate that the *ssrSB* gene was co-transcribed along, but not individually, with *ssrSA*. Finally, through the 5'-RACE experiment (**Figure 2A**), we identified a TSS as well as the canonical −35 and −10 regions (**Figure 2B**) upstream of the *ssrSA* gene, but found no independent TSS upstream of the *ssrSB* gene. These results thus clearly confirmed that the 6S RNA-encoding genes *ssrSA* and *ssrSB* were co-transcribed in BMB171.

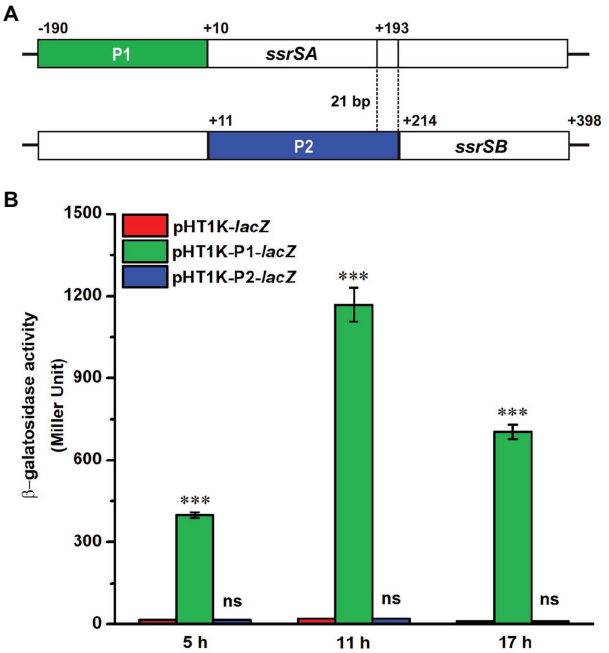


FIGURE 1 | Promoter activity assays of the *ssrSA* and *ssrSB* genes in BMB171. **(A)** Schematic diagram of *ssrSA* and *ssrSB* genes and their putative promoter regions of P1 (marked in green) and P2 (marked in blue). The P1 and P2 regions were each fused with a β-galactosidase gene (*lacZ*) and transformed into the plasmid pHT1K, respectively. **(B)** The β-galactosidase activity assay of strains BMB171/P1-*lacZ* and BMB171/P2-*lacZ* harboring plasmids pHT1K-P1-*lacZ* and pHT1K-P2-*lacZ* with the BMB171/-*lacZ* harboring pHT1K carrying a promoter-less *lacZ* as the control. The strains were cultured at 28°C in GYS medium and were determined in the logarithmic phase (5 h), transition phase (11 h), and stationary phase (17 h). The values were means ± SDs for triplicate assays. Significances of differences by Student's *t*-test are indicated. $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; ns, $p > 0.05$.

Deletion of 6S-1 RNA Inhibited the Growth of *B. thuringiensis* in the Stationary Phase

It has been reported that deletion of 6S-1 RNA results in inhibition of the growth of *B. subtilis* in the stationary phase (Hoch et al., 2015). We thus wonder whether deletion of 6S RNAs also exhibits the similar effect on the growth of *B. thuringiensis*. We have thus used the markerless gene knockout technology to precisely delete the two 6S RNA-encoding genes in the starting strain BMB171, and constructed the single deletion 6S-1 RNA mutant Δ*ssrSA*, single deletion 6S-2 RNA mutant Δ*ssrSB*, and the double deletion mutant Δ*ssrSAB*. We first determined the growth curves of BMB171 and its mutants, and found that the growth rates of mutants Δ*ssrSA* and Δ*ssrSAB* declined much more rapidly than that of starting strain BMB171 in the stationary phase, while Δ*ssrSB* exhibited no significant change compared to BMB171 (**Figure 3A**). Next, we checked the cell morphologies of BMB171, Δ*ssrSA*, Δ*ssrSB*, and Δ*ssrSAB* at 17 h with transmission electron

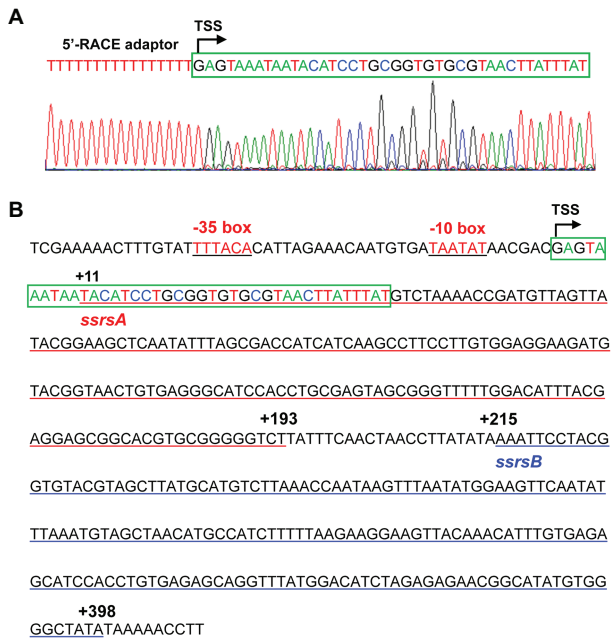


FIGURE 2 | Identification of the transcription start site (TSS) and promoter regions of the *ssrSAB* operon. **(A)** TSS is identified by 5'-rapid amplification of complementary DNA (cDNA) ends (5'-RACE), with the 5'-RACE aptamer sequence (a tandem of 16 thymine nucleotides) shown in red, and the TSS marked with a curved arrow. The 5'-end 40 bp sequence of the *ssrSAB* operon was highlighted with a green box, with the colored nucleotides consistently as those in the sequencing diagram (below). **(B)** The promoter (−51 to −1) and encoding region (+1 to +398) sequences of the *ssrSAB* operon. The −35 and −10 regions were underlined in black, and the *ssrSA* (+1 to +193) and *ssrSB* (+215 to +398) encoding regions underlined in red and blue, respectively. The TSS was marked with a curved arrow with the 5'-end 40 bp sequence enclosed by a green box and colored as those in **(A)**.

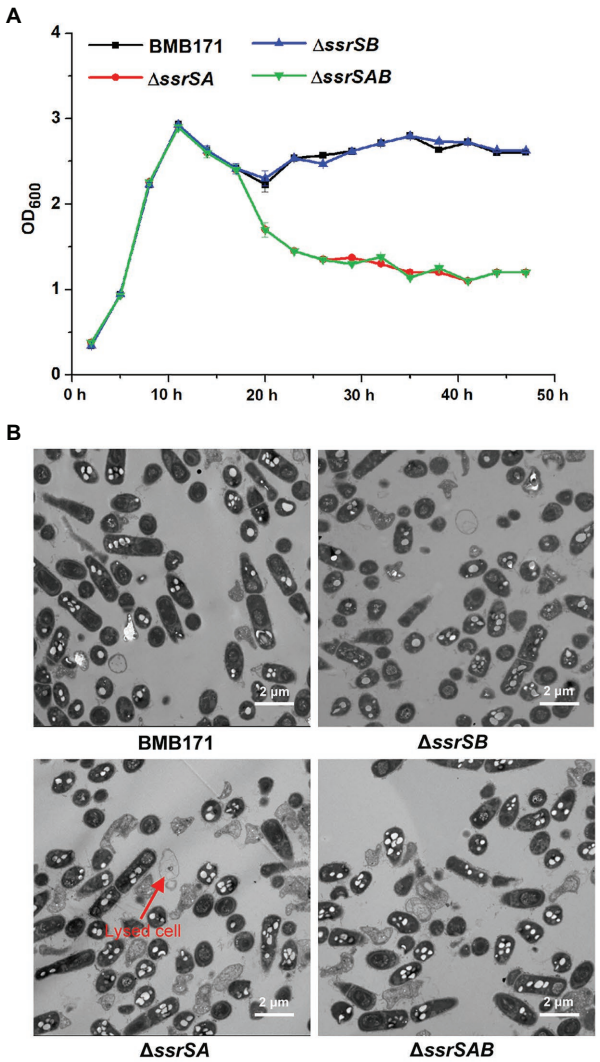


FIGURE 3 | The effect of deleting of 6S RNA-encoding genes on the growths of *Bacillus thuringiensis*. **(A)** Determination of the growth curves of BMB171 and its 6S RNA-encoding genes deletion mutants Δ*ssrSA*, Δ*ssrSB*, and Δ*ssrSAB*. **(B)** Transmission electron microscope images of the cell morphologies of BMB171 and its deletion mutants Δ*ssrSA*, Δ*ssrSB*, and Δ*ssrSAB* at 17 h (stationary phase). Lysed cells are indicated by red arrows. The above-mentioned strains were cultured at 28°C in GYS medium. The values were means ± SDs for triplicate assays.

microscope, and found that deletion of 6S-1 RNA caused more cells to lyse (**Figure 3B**), while the complemented strains Δ*ssrSA*/P1-*ssrSA* and Δ*ssrSAB*/P1-*ssrSAB* showed no significant differences in bacterial growth and morphology (**Supplementary Figure S2**) compared to the control strain BMB171/pHT1K. Taken together, these results indicate that deletion of 6S-1 RNA but not 6S-2 RNA inhibited the growth of *B. thuringiensis* in the stationary phase.

Deletion of 6S-1 RNA Inhibited the Sporulation and Parasporal Crystal Formation

To investigate the roles of 6S-1 and 6S-2 RNAs in the process of sporulation, we first checked the spores of BMB171 and its mutants in the stationary phase through a phase-contrast microscope, and found that the spores formed by mutants Δ*ssrSA* and Δ*ssrSAB* were much fewer than those of the starting strain BMB171. However, the spore amount of Δ*ssrSB* was similar to that of BMB171 (**Figure 4A**). Second, we counted the spores by spread-plate method, and found that the spore numbers formed by Δ*ssrSA* and Δ*ssrSAB* were much fewer than those of BMB171, whereas the spore number in Δ*ssrSB* did not differ much from that in BMB171 (**Figure 4B**). Meanwhile, we also complemented back the *ssrSA* and *ssrSAB* genes into the mutants Δ*ssrSA* and Δ*ssrSAB*, respectively. Phase-contrast microscopy observation and spore count did show that the sporulation capabilities of the complemented strains Δ*ssrSA*/P1-*ssrSA* and Δ*ssrSAB*/P1-*ssrSAB* have returned to the original level of the control strain BMB171/pHT1K (**Supplementary Figure S3**). These results indicate that deletion of the *ssrSA* gene, but not *ssrSB*, inhibited sporulation; namely, the presence of 6S-1 RNA was required for the normal sporulation.

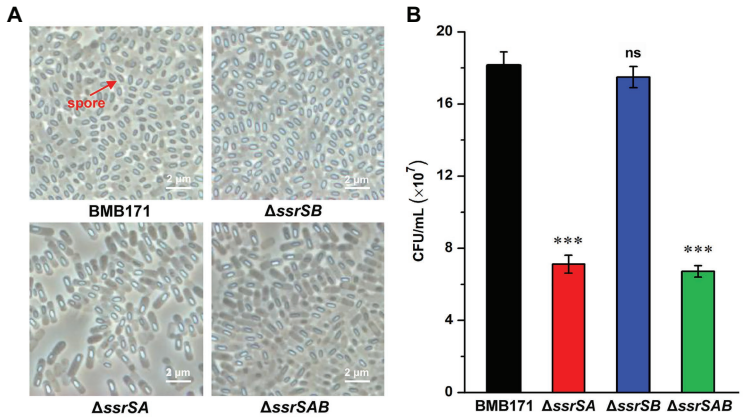


FIGURE 4 | The effect of deleting of 6S RNA-encoding genes on the *B. thuringiensis* sporulation. **(A)** Observation of the spore formed by BMB171 and its deletion mutants Δ*ssrSA*, Δ*ssrSB*, and Δ*ssrSAB* at 24 h (stationary phase) using a phase contrast microscope. Oval-shaped spores are indicated by red arrows. **(B)** The spore counts of BMB171 and its mutants Δ*ssrSA*, Δ*ssrSB*, and Δ*ssrSAB* at 24 h (stationary phase). The strains were cultured at 28°C in GYS medium. The values were means ± SDs for triplicate assays. Significances of differences by Student's *t*-test are indicated. $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; ns, $p > 0.05$.

BMB171 is an acrystalliferous mutant of wild-type YBT-1463 (Li et al., 2000; Qi et al., 2015). To explore whether 6S RNAs affect the parasporal crystal formation, we introduced a pBMB43-304 plasmid containing the parasporal crystal protein encoding gene *cry1Ac10* and its original promoter sequence (Qi et al., 2015) into the BMB171 and its mutant strains Δ*ssrSA*, Δ*ssrSB*, and Δ*ssrSAB*, respectively, to obtain the crystalliferous strains of BMB171-*cry*, Δ*ssrSA-cry*, Δ*ssrSB-cry*, and Δ*ssrSAB-cry*. We first checked the parasporal crystal formation of BMB171 and its mutants in the stationary phase through a phase-contrast microscope, and found that the amounts of parasporal crystals formed by the Δ*ssrSA-cry* and Δ*ssrSAB-cry* strains were far less than that of BMB171-*cry*, while the amount of crystals in the Δ*ssrSB-cry* was similar to that of BMB171-*cry* (**Figure 5A**). Subsequently, we measured the content of parasporal crystal protein Cry1Ac10, and found that the concentrations of Cry1Ac10 in Δ*ssrSA-cry* and Δ*ssrSAB-cry* were remarkably reduced compared to BMB171-*cry*, while the Δ*ssrSB-cry* exhibited no substantial difference (**Figures 5B,C**). These experiments show that deletion of 6S-1 RNA not only inhibited the sporulation of bacterial cells, but also repressed the parasporal crystal formation.

Reduction in Sporulation Capability Was Due to Inhibition of Bacterial Growth in the Stationary Phase

To figure out the reasons accounting for inhibition of sporulation in BMB171 after 6S-1 RNA deletion, we used the real-time quantitative PCR (RT-qPCR) to examine the transcription levels of several key early sporulation-related genes in BMB171 and its mutant's strains Δ*ssrSA*, Δ*ssrSB*, and Δ*ssrSAB*. These relevant genes include *spo0A*, which encodes an essential transcriptional regulatory factor in the initial stage of sporulation (Lopez et al., 2009); *kinA*, which encodes a phosphorylation kinase that is mainly responsible for the phosphorylation of Spo0A (Jiang et al., 2000a); and *spo0H*, which encodes σ^H to promote the transcription of *spo0A* and *kinA* (Predich et al., 1992). Taking gene *gapdh* encoding glyceraldehyde-3-phosphate dehydrogenase as an internal control, we demonstrated that the transcription levels of *spo0A*, *kinA*, and *spo0H* from mutants Δ*ssrSA*, Δ*ssrSB*, and Δ*ssrSAB* at 0.5, 1.0, 1.5, and 2.0 h showed almost no difference to those of BMB171 (**Supplementary Figure S4**). Besides, we further determined the transcriptomes of BMB171 and Δ*ssrSAB* in the stationary phase and found that the transcription levels of sporulation-related genes did not change significantly between BMB171 and Δ*ssrSAB* (**Supplementary Table S2**). These data demonstrate that deletion of 6S-1 RNA did not change the start time of sporulation at the molecular level.

Since deletion of 6S-1 RNA did inhibit the growth of *B. thuringiensis* in the stationary phase, we wonder whether such inhibition also affects their sporulation? Because BMB171 and its various mutants Δ*ssrSA*, Δ*ssrSB*, and Δ*ssrSAB* exhibited no obvious difference in the growth curves in the logarithmic phase, we wonder whether there is difference in sporulation efficacy when these strains were induced to produce spores at this time? Given that decoyinine can induce *Bacillus* cells to produce spores in advance (Cavanagh and Wassarman, 2013; Wang et al., 2016; Zhou et al., 2017), we thus conducted an experiment by adding decoyinine to the GYS medium in the logarithmic phase. The results showed that the spore numbers formed by Δ*ssrSA*, Δ*ssrSB*, and Δ*ssrSAB* did not differ much from that of BMB171 (**Figure 6**). This experiment proves that the reduced sporulation efficacy was mainly caused by inhibition of the growth of *B. thuringiensis* in the stationary phase, not in the logarithmic phase. Further, we found that the transcription levels of genes involved in carbohydrate transport and metabolism, nucleotide transport and metabolism, protein

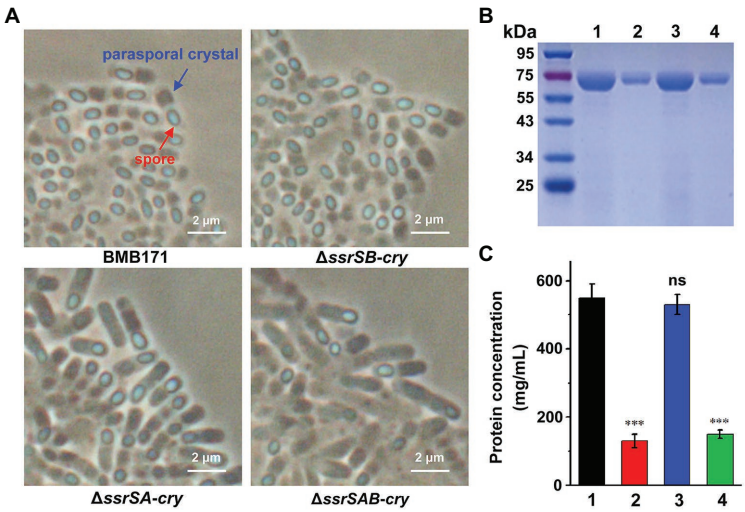


FIGURE 5 | The effect of deleting 6S RNA-encoding genes on the parasporal crystal formation of *B. thuringiensis* **(A)** Phase contrast microscopy image of the parasporal crystal formation in crystalliferous strains BMB171-*cry*, Δ*ssrSA-cry*, Δ*ssrSB-cry*, and Δ*ssrSAB-cry* at 24 h (stationary phase). Cells with parasporal crystals and spores are indicated by blue and red arrows, respectively. **(B)** Separation of Cry1Ac10 from the four strains by SDS-PAGE. Lines 1–4 represent BMB171-*cry*, Δ*ssrSA-cry*, Δ*ssrSB-cry*, and Δ*ssrSAB-cry*, respectively. **(C)** The concentrations of Cry1Ac10 in BMB171-*cry* (1), Δ*ssrSA-cry* (2), Δ*ssrSB-cry* (3), and Δ*ssrSAB-cry* (4) at 24 h (stationary phase). The above-mentioned strains were cultured at 28°C in GYS medium. The values were means ± SDs for triplicate assays. Significances of differences by Student's *t*-test are indicated. $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; ns, $p > 0.05$.

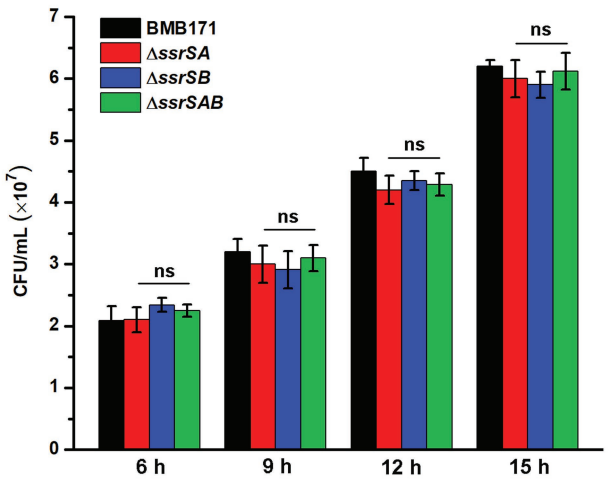


FIGURE 6 | The effect of deleting 6S RNA-encoding genes on the sporulation of BMB171 and its mutants in the logarithmic phase. After adding the sporulation inducer decoyinine (0.4 mg/ml) to the GYS cultures of above-mentioned strains in the logarithmic phase (3 h, $OD_{600} \approx 0.5$), the spore numbers were counted at 6, 9, 12, and 15 h after induction. The strains were cultured at 28°C and the numbers are revealed as means ± SDs in triplicate assays. Significances of differences by Student's *t*-test are indicated. $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; ns, $p > 0.05$.

translation, and energy production and conversion decreased significantly in Δ*ssrSAB* (**Supplementary Table S3**). Therefore, the decreased growth rate of the *ssrSA* deletion mutant in the stationary phase is the main reason for the diminished sporulation of *B. thuringiensis*.

DISCUSSION

In this study, we verified that the 6S-1 and 6S-2 RNAs-encoding genes *ssrSA* and *ssrSB* were located in the same operon and co-transcribed in *B. thuringiensis*. Through the phenotypic study of BMB171 and its deletion mutants of Δ*ssrSA*, Δ*ssrSB*, and Δ*ssrSAB*, we found that deletion of 6S-1 RNA not only inhibited the growth of *B. thuringiensis* in the stationary phase, but also decreased the sporulation and parasporal crystal formation. We further confirmed that inhibition of the growth in the stationary phase is likely the primary reason for the reduced sporulation efficacy.

In BMB171, we validated that the 6S-1 and 6S-2 RNAs-encoding genes *ssrSA* and *ssrSB* were arranged in tandem and co-transcribed, and verified that the upstream sequence of the *ssrSA* gene exhibited the transcription initiation activity, but not the upstream sequence of the *ssrSB* gene (**Figure 1**). These results indicate that the transcriptions of *ssrSA* and *ssrSB* genes were initiated through the upstream promoter region of the *ssrSA* gene, and were jointly transcribed into a long RNA precursor. In *E. coli*, the 6S RNA-encoding gene is first transcribed into a long RNA precursor, which is then cleaved by RNase E and RNase G into mature 6S RNA (Kim and Lee, 2004). Given the fact that RNase E and RNase G prefer to cut the regions enriched in nucleotides A and U on single-stranded RNA (Mackie, 1998; Jiang et al., 2000b; Tock et al., 2000), so we analyzed the RNA precursor sequence in BMB171 and found that, similar to *E. coli*, the 5'- and 3'-ends of the precursor RNA of 6S-1 RNA and 6S-2 RNA in BMB171 are also rich in A and U (**Supplementary Figure S5**). Moreover, transcriptomic data showed that 6S RNAs existed mainly as two separate mature 6S-1 and

6S-2 RNAs (**Supplementary Figure S6**). Combining the above results (**Supplementary Figures S1**, **S5**, **S6**), we suggest that a long precursor 6S RNA is first transcribed, followed by processing with RNase E, RNase G, or other ribonucleases to form mature and functional 6S-1 and 6S-2 RNAs.

In the present manuscript, we have analyzed the 6S RNA-encoding genes of *Bacillus genus* bacteria and their gene organization. *Bacillus* genus bacteria usually contain two different 6S RNA-encoding genes of *ssrSA* and *ssrSB* (**Supplementary Tables S4–S6**). However, unlike other *Bacillus* genus bacteria that have these two genes located in different regions of the genome, the two 6S RNA-encoding genes of *B. cereus* group bacteria were located in tandem positions. Furthermore, we also verified that they were in the same operon and are transcribed together. The *ssrSA* and *ssrSB* genes of *B. cereus* group bacteria are located between the genes encoding thiocyanase and glutathione spermidine synthase. The former is related to the detoxification of exogenous toxins while the latter involved in the resistance to stress such as oxidation (**Supplementary Table S4**). Other than the *B. cereus* group, there is also a *B. subtilis* group bacteria composed of more than 20 important bacteria species, such as, *B. subtilis*, *Bacillus licheniformis*, *Bacillus pumilus*, *Bacillus velezensis*, and *Bacillus amyloliquefaciens* in *Bacillus* genus (Caulier et al., 2019; Fayad et al., 2019). Unlike the *B. cereus* group bacteria, the *ssrSA* gene of *B. subtilis* group bacteria was mainly located between the FMN-dependent NADH-azoreductase and DNA helicase RecQ encoding genes. The former is related to the detoxification of exogenous toxins, while the latter is linked with DNA repair, recombination, and replication. The *ssrSB* gene, on the other hand, was mainly located between the aspartate-tRNA ligase and tRNA threonylcarbamoyladenosine dehydratase, both of which were associated with protein translation (**Supplementary Table S5**). Among *Bacillus* genus bacteria excluding *B. cereus* group and *B. subtilis* group, only the aspartate-tRNA ligase encoding gene locus upstream of the *ssrSB* gene was relatively conservative (**Supplementary Table S6**).

The 6S RNAs in *B. cereus* group bacteria were quite different to those of *B. subtilis* group bacteria regarding their gene organization, which might lead to their differences in sporulation. For example, in *B. subtilis*, deletion of 6S-1 RNA accelerates the utilization of nutrients, leading to the early arrival of nutrient deprivation conditions, which in turn induces the earlier expression of key early sporulation-related genes, ultimately promoting the earlier sporulation (Wassarman, 2018), while deletion of 6S-1 RNA in *B. thuringiensis* resulted in suppression of sporulation. In order to explore its possible regulatory mechanism, we have determined the transcriptomes of BMB171 and Δ*ssrSAB* in the stationary phase. After analyses, we found that the transcription levels of sporulation-related genes did not change significantly between BMB171 and Δ*ssrSAB* (**Supplementary Table S2**), indicating that 6S RNA does not affect sporulation by directly regulating the temporal expression of sporulation-related genes. Moreover, the experiments of decoyinine-induced sporulation in the logarithmic phase (**Figure 6**) and RT-qPCR detection of the transcription levels of key early sporulation-related genes (**Supplementary Figure S4**) further proved this conclusion. Since there is no available transcriptomic data from the 6S-1 RNA deletion in *B. subtilis*, we are currently unable to compare the 6S-1 RNA regulatory mechanisms between the two bacteria in a more comprehensive way. Altogether, the 6S RNAs of *B. subtilis* group and *B. cereus* group bacteria exhibited different gene organization and physiological functions, indicating that 6S RNAs might regulate the biological functions of *B. subtilis* group and *B. cereus* group bacteria *via* different mechanisms.

Then, how does deletion of 6S-1 RNA inhibit sporulation of *B. thuringiensis*? After further analyzing the transcriptomic data, we found that the transcription levels of genes involved in carbohydrate transport and metabolism, nucleotide transport and metabolism, protein translation, and energy production and conversion decreased significantly in Δ*ssrSAB* (**Supplementary Table S3**). According to our previous report that sporulation and parasporal crystal formation do require a lot of material and energy supply in *B. thuringiensis* (Wang et al., 2013c), we therefore speculate that the insufficient supply of material and energy in Δ*ssrSAB* was the main reason for inhibition of sporulation.

How does deletion of 6S-1 RNA inhibits the parasporal crystal formation of *B. thuringiensis*? In crystalliferous strain BMB171-*cry*, *cry1Ac10* with the original promoter was regulated by a sporulation-specific sigma factor SigE. In the transcriptomic data, we did not find a significant expression difference of *sigE* between BMB171 and Δ*ssrSAB* (**Supplementary Table S2**). Meanwhile, RT-qPCR assays confirmed that *cry1Ac10* had no expression difference between BMB171 and Δ*ssrSAB* (**Supplementary Figure S7**). However, the content of parasporal crystals was significantly reduced in Δ*ssrSAB* (**Figure 5B**). This indicates that deletion of 6S RNA may inhibit the translation of Cry1Ac10, further reducing the parasporal crystal formation.

Like 6S RNA, CsrA, CarD, and (p)ppGpp are all global regulatory factors that can respond to starvation stress in the stationary phase (Kalia et al., 2013; Romeo et al., 2013; Flentie et al., 2016). CsrA is believed to inhibit bacterial translation by interacting with conserved sequences on target mRNA under carbon starvation conditions (Romeo et al., 2013). Both CarD and (p)ppGpp, on the other hand, seem to regulate downstream gene transcription by interacting with RNA polymerase; in addition, the binding of CarD with RNA polymerase has been found to stabilize the transcription initiation complex to initiate the transcription of downstream genes (Flentie et al., 2016). Under starvation conditions, high levels of intracellular (p)ppGpp can also inhibit translation, which not only regulates the transcription of tRNA, rRNA, and ribosomal protein genes by binding to RNA polymerase, but also directly inhibits the bacterial translation activity through combining translation initiation factor IF2 with the translation elongation factors EF-Tu and EF-G (Kalia et al., 2013). Like CarD and (p)ppGpp, 6S RNA is also a global regulatory factor that can regulate gene transcription by binding to RNA polymerase. The consistent function of 6S RNA, CsrA, CarD, and (p)ppGpp may be to maintain and optimize the survival rates of bacteria under different stress conditions, which may be manifested in rather complex metabolic regulation networks. Currently we cannot

figure out what are the specific roles 6S RNA plays in these networks and also how they co-regulate the bacterial response to starvation stress in the stationary phase, which are issues deserve further exploration.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

JH and ZL designed the experiments. ZL, LZ, ZY, and LL did the experiments. JH, S-HC, JW, and ZL wrote and revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This paper is supported by the National Natural Science Foundation of China (grants 31770087 and 31970074), and the Cultivation fund project of Fujian Academy of Agricultural Sciences (AGP2018-2).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.604458/full#supplementary-material

REFERENCES

Ali, M. K., Li, X. F., Tang, Q., Liu, X. Y., Chen, F., Xiao, J. F., et al. (2017). Regulation of inducible potassium transporter KdpFABC by the KdpD/KdpE two-component system in *Mycobacterium smegmatis*. *Front. Microbiol.* 8:570. doi: 10.3389/fmicb.2017.00570

Brownlee, G. G. (1971). Sequence of 6S RNA of *E. coli*. Nat. *New Biol.* 229, 147–149. doi: 10.1038/newbio229147a0

Caulier, S. C., Nannan, C., Gillis, A., Licciard, F., Bragard, C., and Mahillon, J. (2019). Overview of the antimicrobial compounds produced by members of the *Bacillus subtilis* group. *Front. Microbiol.* 10:302. doi: 10.3389/fmicb.2019.00302

Cavanagh, A. T., and Wassarman, K. M. (2013). 6S-1 RNA function leads to a delay in sporulation in *Bacillus subtilis*. *J. Bacteriol.* 195, 2079–2086. doi: 10.1128/JB.00050-13

Cavanagh, A. T., and Wassarman, K. M. (2014). 6S RNA, a global regulator of transcription in *Escherichia coli*, *Bacillus subtilis*, and beyond. *Annu. Rev. Microbiol.* 68, 45–60. doi: 10.1146/annurev-micro-092611-150135

Craig, J. E., Ford, M. J., Blaydon, D. C., and Sonenshein, A. L. (1997). A null mutation in the *Bacillus subtilis* aconitase gene causes a block in Spo0A-phosphate-dependentgene expression. *J. Bacteriol.* 179, 7351–7359. doi: 10.1128/jb.179.23.7351-7359.1997

Fayad, N., Awad, M. K., and Mahillon, J. (2019). Diversity of *Bacillus cereus* sensu *lato* mobilome. *BMC Genomics* 20:436. doi: 10.1186/s12864-019-5764-4

Flentie, K., Garner, A. L., and Stallings, C. L. (2016). *Mycobacterium tuberculosis* transcription machinery: ready to respond to host attacks. *J. Bacteriol.* 198, 1360–1373. doi: 10.1128/JB.00935-15

Fu, Y., Yu, Z. Q., Liu, S., Chen, B., Zhu, L., Li, Z., et al. (2018). C-di-GMP regulates various phenotypes and insecticidal activity of gram-positive *Bacillus thuringiensis*. *Front. Microbiol.* 9:45. doi: 10.3389/fmicb.2018.00045

He, J., Shao, X. H., Zheng, H. J., Li, M. S., Wang, J. P., Zhang, Q. Y., et al. (2010). Complete genome sequence of *Bacillus thuringiensis* mutant strain BMB171. *J. Bacteriol.* 192, 4074–4075. doi: 10.1128/JB.00562-10

Heilmann, B., Hakkila, K., Georg, J., Tyystjärvi, T., Hess, W. R., Axmann, I. M., et al. (2017). 6S RNA plays a role in recovery from nitrogen depletion in *Synechocystis* sp. PCC 6803. *BMC Microbiol.* 17:229. doi: 10.1186/s12866-017-1137-9

Hindley, J. (1967). Fractionation of ^{32}P-labelled ribonucleic acids on polyacrylamide gels and their characterization by fingerprinting. *J. Mol. Biol.* 30, 125–136. doi: 10.1016/0022-2836(67)90248-3

Hoch, P. G., Burenina, O. Y., Weber, M. H. W., Elkina, D. A., Nesterchuk, M. V., Sergiev, P. V., et al. (2015). Phenotypic characterization and complementation analysis of *Bacillus subtilis* 6S RNA single and double deletion mutants. *Biochimie* 117, 87–99. doi: 10.1016/j.biochi.2014.12.019

Ikehara, K., Okamoto, M., and Sugae, K. (1982). Induction of *Bacillus subtilis* sporulation by decoyinine and the concomitant disappearance of ppGpp in vegetative cells. *J. Biochem.* 91, 1089–1092. doi: 10.1093/oxfordjournals.jbchem.a133759

Janes, B. K., and Stibitz, S. (2006). Routine markerless gene replacement in *Bacillus anthracis*. *Infect. Immun.* 74, 1949–1953. doi: 10.1128/IAI.74.3.1949-1953.2006

Jiang, X., Diwa, A., and Belasco, J. G. (2000b). Regions of RNase E important for 5'-end-dependent RNA cleavage and autoregulated synthesis. *J. Bacteriol.* 182, 2468–2475. doi: 10.1128/jb.182.9.2468-2475.2000

Jiang, M., Shao, W., Perego, M., and Hoch, J. A. (2000a). Multiple histidine kinases regulate entry into stationary phase and sporulation in *Bacillus subtilis*. *Mol. Microbiol.* 38, 535–542. doi: 10.1046/j.1365-2958.2000.02148.x

Kalia, D., Merey, G., Nakayama, S., Zheng, Y., Zhou, J., Luo, Y. L., et al. (2013). Nucleotide, c-di-GMP, c-di-AMP, cGMP, cAMP, (p)ppGpp signaling in bacteria and implications in pathogenesis. *Chem. Soc. Rev.* 42, 305–341. doi: 10.1039/c2cs35206k

Kim, K. S., and Lee, Y. (2004). Regulation of 6S RNA biogenesis by switching utilization of both sigma factors and endoribonucleases. *Nucleic Acids Res.* 32, 6057–6068. doi: 10.1093/nar/gkh939

Li, L., Yang, C., Liu, Z., Li, F., and Yu, Z. (2000). Screening of acrystalliferous mutants from *Bacillus thuringiensis* and their transformation properties. *Wei Sheng Wu Xue Bao* 40, 85–90.

Liu, Y., Du, J., Lai, Q. L., Zeng, R. Y., Ye, D. Z., Xu, J., et al. (2017). Proposal of nine novel species of the *Bacillus cereus* group. *Int. J. Syst. Evol. Microbiol.* 67, 2499–2508. doi: 10.1099/ijsem.0.001821

Lopez, D., Vlamakis, H., and Kolter, R. (2009). Generation of multiple cell types in *Bacillus subtilis*. *FEMS Microbiol. Rev.* 33, 152–163. doi: 10.1111/j.1574-6976.2008.00148.x

Mackie, G. A. (1998). Ribonuclease E is a 5'-end-dependent endonuclease. *Nature* 395, 720–723. doi: 10.1038/27246

Madhugiri, R., Pessi, G., Voss, B., Hahn, J., Sharma, C. M., Reinhardt, R., et al. (2012). Small RNAs of the *Bradyrhizobium*/*Rhodopseudomonas* lineage and their analysis. *RNA Biol.* 9, 47–58. doi: 10.4161/rna.9.1.18008

Mikulík, K., Bobek, J., Zídková, J., and Felsberg, J. (2014). 6S RNA modulates growth and antibiotic production in *Streptomyces coelicolor*. *Appl. Microbiol. Biotechnol.* 98, 7185–7197. doi: 10.1007/s00253-014-5806-4

Miller, R. A., Beno, S. M., Kent, D. J., Carroll, L. M., Martin, N. H., Boor, K. J., et al. (2016). *Bacillus wiedmannii* sp. nov., a psychrotolerant and cytotoxic *Bacillus cereus* group species isolated from dairy foods and dairy environments. *Int. J. Syst. Evol. Microbiol.* 66, 4744–4753. doi: 10.1099/ijsem.0.001421

Mitani, T., Heinze, J. E., and Freese, E. (1977). Induction of sporulation in *Bacillus subtilis* by decoyinine or hadacidin. *Biochem. Biophys. Res. Commun.* 77, 1118–1125. doi: 10.1016/s0006-291x(77)80094-6

Predich, M., Nair, G., and Smith, I. (1992). *Bacillus subtilis* early sporulation genes *kinA*, *spo0F*, and *spo0A* are transcribed by the RNA polymerase containing sigma H. *J. Bacteriol.* 174, 2771–2778. doi: 10.1128/JB.174.9.2771-2778.1992

Qi, M. X., Mei, F., Wang, H., Sun, M., Wang, G. J., Yu, Z. N., et al. (2015). Function of global regulator CodY in *Bacillus thuringiensis* BMB171 by comparative proteomic analysis. *J. Microbiol. Biotechnol.* 25, 152–161. doi: 10.4014/jmb.1406.06036

Ren, J., Sang, Y., Qin, R., Cui, Z. L., and Yao, Y. F. (2017). 6S RNA is involved in acid resistance and invasion of epithelial cells in *Salmonella enterica* serovar Typhimurium. *Future Microbiol.* 12, 1045–1057. doi: 10.2217/fmb-2017-0055

Romeo, T., Vakulskas, C. A., and Babitzke, P. (2013). Posttranscriptional regulation on a global scale: form and function of Csr/Rsm systems. *Environ. Microbiol.* 15, 313–324. doi: 10.1111/j.1462-2920.2012.02794.x

Steuten, B., Schneider, S., and Wagner, R. (2014). 6S RNA: recent answers-future questions. *Mol. Microbiol.* 91, 641–648. doi: 10.1111/mmi.12484

Tang, Q., Yin, K., Qian, H. L., Zhao, Y. W., Wang, W., Chou, S. H., et al. (2016). Cyclic di-GMP contributes to adaption and virulence of *Bacillus thuringiensis* through a riboswitch-regulated collagen adhesion protein. *Sci. Rep.* 6:28807. doi: 10.1038/srep28807

Tock, M. R., Walsh, A. P., Carroll, G., and McDowall, K. J. (2000). The CafA protein required for the 5'-maturation of 16S rRNA is a 5'-end-dependent ribonuclease that has context-dependent broad sequence specificity. *J. Biol. Chem.* 275, 8726–8732. doi: 10.1074/jbc.275.12.8726

Trotochaud, A. E., and Wassarman, K. M. (2004). 6S RNA function enhances long-term cell survival. *J. Bacteriol.* 186, 4978–4985. doi: 10.1128/JB.186.15.4978-4985.2004

Wang, J. P., Ai, X. L., Mei, H., Fu, Y., Chen, B., Yu, Z. N., et al. (2013a). High-throughput identification of promoters and screening of highly active promoter-5'-UTR DNA region with different characteristics from *Bacillus thuringiensis*. *PLoS One* 8:e62960. doi: 10.1371/journal.pone.0062960

Wang, X., Cai, X., Ma, H. D., Yin, W., Zhu, L., Li, X. F., et al. (2019). A c-di-AMP riboswitch controlling *kdpFABC* operon transcription regulates the potassium transporter system in *Bacillus thuringiensis*. *Commun. Biol.* 2:151. doi: 10.1038/s42003-019-0414-6

Wang, X., Li, Z., Li, X., Qian, H. L., Cai, X., Li, X. F., et al. (2016). Poly-β-hydroxybutyrate metabolism is unrelated to the sporulation and parasporal crystal protein formation in *Bacillus thuringiensis*. *Front. Microbiol.* 7:836. doi: 10.3389/fmicb.2016.00836

Wang, J. P., Mei, H., Qian, H. L., Tang, Q., Liu, X. C., Yu, Z. N., et al. (2013b). Expression profile and regulation of spore and parasporal crystal formation-associated genes in *Bacillus thuringiensis*. *J. Proteome Res.* 12, 5487–5501. doi: 10.1021/pr4003728

Wang, J., Mei, H., Zheng, C., Qian, H. L., Cui, C., Fu, Y., et al. (2013c). The metabolic regulation of sporulation and par-asporal crystal formation in *Bacillus thuringiensis* revealed by transcriptomics and proteomics. *Mol. Cell. Proteomics* 12, 1363–1376. doi: 10.1074/mcp.M112.023986

Wassarman, K. M. (2018). 6S RNA, a global regulator of transcription. *Microbiol. Spectr.* 6:RWR-0019-2018. doi: 10.1128/microbiolspec.RWR-0019-2018

Wassarman, K. M., and Storz, G. (2000). 6S RNA regulates *E. coli* RNA polymerase activity. *Cell* 101, 613–623. doi: 10.1016/s0092-8674(00)80873-9

Wehner, S., Damm, K., Hartmann, R. K., and Marz, M. (2014). Dissemination of 6S RNA among bacteria. *RNA Biol.* 11, 1467–1478. doi: 10.4161/rna.29894

Zheng, C., Ma, Y., Wang, X., Xie, Y. Q., Ali, M. K., and He, J. (2015). Functional analysis of the sporulation-specific diadenylate cyclase CdaS in *Bacillus thuringiensis*. *Front. Microbiol.* 6:908. doi: 10.3389/fmicb.2015.00908

Zheng, C., Yu, Z. Q., Du, C. Y., Gong, Y. J., Yin, W., Li, X. F., et al. (2020). 2-Methylcitrate cycle: a well-regulated controller of *Bacillus* sporulation. *Environ. Microbiol.* 22, 1125–1140. doi: 10.1111/1462-2920.14901

Zhou, H., Zheng, C., Su, J. M., Chen, B., Fu, Y., Xie, Y. Q., et al. (2017). Characterization of a natural triple-tandem c-di-GMP riboswitch and application of the riboswitch-based dual-fluorescence reporter. *Sci. Rep.* 6:20871. doi: 10.1038/srep20871

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

REVIEW
published: 08 January 2021
doi: 10.3389/fmicb.2020.619618

NusG, an Ancient Yet Rapidly Evolving Transcription Factor

*Bing Wang and Irina Artsimovitch**

Department of Microbiology and the Center for RNA Biology, The Ohio State University, Columbus, OH, United States

Timely and accurate RNA synthesis depends on accessory proteins that instruct RNA polymerase (RNAP) where and when to start and stop transcription. Among thousands of transcription factors, NusG/Spt5 stand out as the only universally conserved family of regulators. These proteins interact with RNAP to promote uninterrupted RNA synthesis and with diverse cellular partners to couple transcription to RNA processing, modification or translation, or to trigger premature termination of aberrant transcription. NusG homologs are present in all cells that utilize bacterial-type RNAP, from endosymbionts to plants, underscoring their ancient and essential function. Yet, in stark contrast to other core RNAP components, NusG family is actively evolving: horizontal gene transfer and sub-functionalization drive emergence of NusG paralogs, such as bacterial LoaP, RfaH, and UpxY. These specialized regulators activate a few (or just one) operons required for expression of antibiotics, capsules, secretion systems, toxins, and other niche-specific macromolecules. Despite their common origin and binding site on the RNAP, NusG homologs differ in their target selection, interacting partners and effects on RNA synthesis. Even among housekeeping NusGs from diverse bacteria, some factors promote pause-free transcription while others slow the RNAP down. Here, we discuss structure, function, and evolution of NusG proteins, focusing on unique mechanisms that determine their effects on gene expression and enable bacterial adaptation to diverse ecological niches.

Keywords: antitermination, evolution, NusG, RfaH, transcriptional pausing, termination, virulence

Edited by:
Omar Orellana,
University of Chile, Chile

Reviewed by:
Paul Babitzke,
Pennsylvania State University (PSU), United States
Paolo Landini,
University of Milan, Italy

***Correspondence:**
Irina Artsimovitch
artsimovitch.1@osu.edu

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *20 October 2020*
Accepted: *07 December 2020*
Published: *08 January 2021*

Citation:
Wang B and Artsimovitch I (2021) NusG, an Ancient Yet Rapidly Evolving Transcription Factor. Front. Microbiol. 11:619618. doi: 10.3389/fmicb.2020.619618

INTRODUCTION

In every living cell, multi-subunit RNA polymerases (RNAPs) carry out the first step of gene expression, transcription of a DNA template into an RNA copy. Reflecting their common evolutionary origin in the last universal common ancestor (LUCA) and the basic mechanism of RNA synthesis, RNAPs share an overall architecture and structural elements that play key roles in the assembly of transcription complexes, substrate selection and catalysis, interactions with nucleic acids, etc. (Lane and Darst, 2010a,b). However, extant RNAPs differ greatly in subunit composition and sequence: core RNAPs are composed of 5–7 subunits in bacteria vs. 12+ subunits in archaea and eukaryotes, and even RNAPs from mesophilic bacteria *Escherichia coli* and *Bacillus subtilis* are only 50% identical. Differences in cellular transcriptional machinery are thought to reflect unique regulatory constraints imposed by diverse habitats. In support of this notion, even basal general transcription factors that assist RNAP during each step of the transcription cycle are not conserved between kingdoms. The sole exception to this trend is a transcription elongation factor NusG (Werner, 2012).

Bacterial Nus (N-utilization substance) proteins have been identified genetically based on their requirement for the coliphage λ development (Casjens and Hendrix, 2015). In *E. coli* and *Salmonella*, potentially harmful xenogenes are silenced by premature transcription termination by a hexameric RNA helicase Rho (Peters et al., 2012; Bossi et al., 2019). To escape silencing, bacteriophages have evolved antitermination mechanisms targeting Rho or RNAP (Santangelo and Artsimovitch, 2011). The immediate early gene *N* of phage λ is required for the expression of delayed-early genes. N nucleates the assembly of a large transcription antitermination complex (TAC) composed of RNAP and NusABEG proteins (Mason and Greenblatt, 1991; Krupp et al., 2019) and a similar TAC assembles during transcription of the *E. coli* ribosomal RNA operons (Squires et al., 1993; Huang et al., 2020). NusA and NusG are general transcription elongation factors, which are associated with RNAP transcribing all genes, at least in *E. coli* (Mooney et al., 2009a). NusE, *a.k.a.* the ribosomal protein S10, requires a binding partner NusB to remain soluble while not a part of the ribosome; NusB is selectively enriched on rRNA operons (Mooney et al., 2009a), consistent with its principal role in rRNA synthesis. Among the shared components of the TACs, NusG is the only factor that facilitates transcription elongation *in vivo* and *in vitro* (Burova et al., 1995; Burns et al., 1998; Zellars and Squires, 1999); by contrast, NusA increases RNAP pausing and intrinsic termination, whereas NusB/E have no effect (Belogurov and Artsimovitch, 2015).

All NusG-like proteins (NusG in bacteria; Spt5 in archaea and yeast, DSIF in mammals) bind to an evolutionary conserved site on the largest RNAP subunit (Klein et al., 2011; Martinez-Rucobo et al., 2011; Ehara et al., 2017; Kang et al., 2018; Vos et al., 2018). The NusG binding site is located on the tip of the RNAP clamp, a conserved flexible module that closes over the DNA binding channel. The clamp closes during the formation of a transcriptionally competent initiation complex, remains closed throughout elongation, and opens during termination (Belogurov and Artsimovitch, 2019); more subtle movements of the clamp have been proposed to accompany RNAP pausing, which serves as a prelude to termination (Kang et al., 2019). By keeping the clamp locked, NusG proteins are thought to promote continuous, pause-free RNA synthesis, an essential function given that the premature release of the RNA transcript is irreversible. The presence of a clamping factor in LUCA thus underscores the fundamental importance of transcription processivity, particularly on difficult templates (Werner, 2012).

The antipausing and, by inference, antitermination activity of NusG prompted its annotation as a transcription antiterminator. Likewise, many subsequently discovered bacterial NusG homologs have been shown to possess antitermination activity (Artsimovitch and Knauer, 2019). Nevertheless, this view has been challenged since the time of *E. coli* NusG discovery by the data in support of its role as a termination-promoting factor. NusG is essential in wild-type *E. coli* (Downing et al., 1990) and its depletion leads to defects in Rho-dependent termination (Sullivan and Gottesman, 1992). NusG aids Rho in silencing transcription of damaged and harmful RNAs genome-wide (Peters et al., 2012) and promotes efficient termination by Rho *in vitro* (Burns and Richardson, 1995). Indeed, the *nusG* gene can be deleted, albeit at a significant fitness cost, in an *E. coli* strain lacking the toxic *rac* prophage, which is silenced by Rho (Cardinale et al., 2008). Point mutations in *nusG* that lead to defects in transcription termination (Saxena and Gowrishankar, 2011) or interactions with the ribosome (Saxena et al., 2018) do not have significant fitness phenotypes.

Functional studies of NusG-like proteins from different bacteria support a picture in which these factors can mediate diverse effects on RNA synthesis (**Figure 1**). Through contacts to RNAP, nucleic acids, and auxiliary proteins, NusG homologs can suppress or promote transcriptional pausing and termination and bridge RNAP to other cellular machineries. Most unusually for a family of alternative transcription regulators, although binding to the same site on the transcribing RNAP, NusG-like proteins frequently have exactly opposite effects on the expression of some genes, most notably those encoding virulence determinants. Furthermore, even the housekeeping NusG proteins have seemingly opposite effects on RNA synthesis; for example, unlike its *E. coli* counterpart, *B. subtilis* NusG promotes RNAP pausing *in vitro* and *in vivo* (Yakhnin et al., 2016, 2020a). Below, we describe recent advances in our understanding of molecular mechanisms, evolution, and regulatory diversity of bacterial NusG-like proteins.

STRUCTURE AND TARGET CONSERVATION

NusG-like proteins have a similar structural core consisting of a NusG N-terminal domain (NGN) and a C-terminal domain with a 27-residue long Kyrpides-Ouzounis-Woese (KOW) motif common among RNA-binding proteins (Kyrpides et al., 1996; Ponting, 2002; **Figure 2**). Bacterial NusG alone can perform its function, while Spt5 has an obligatory partner—a small zinc finger protein Spt4 (called RpoE in archaea). Eukaryotic Spt5 contains several KOW domains, the first of which carries a large insertion, an N-terminal acidic region, and an unstructured C-terminal repeat (CTR) domain (**Figure 2A**); in metazoan DSIF, additional KOWs are present at the very C terminus of

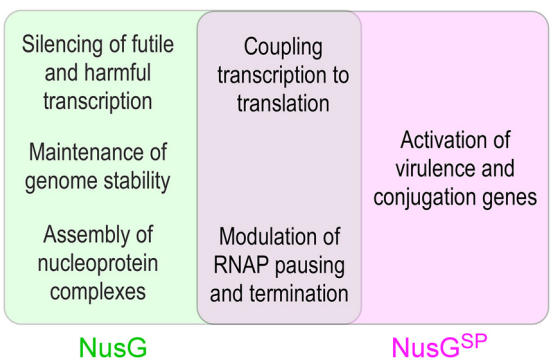


FIGURE 1 | Unique and overlapping cellular functions of housekeeping *E. coli* NusG and its specialized paralogs.

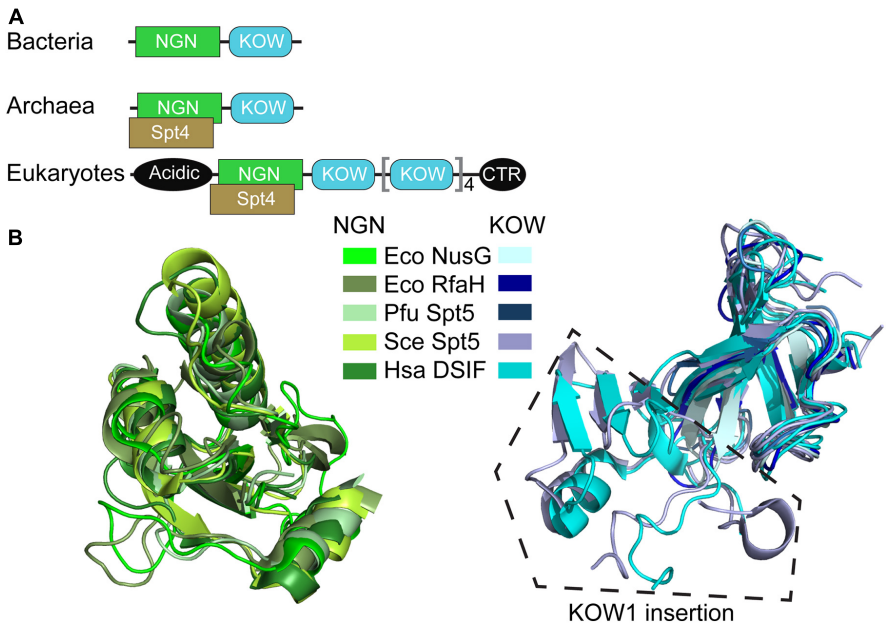


FIGURE 2 | Structural conservation of NusG-like proteins. **(A)** Domain organization. **(B)** Superposition of NGN and KOW domains. PDB IDs: *E. coli* (Eco) NusG-NGN: 2K06; Eco NusG-KOW: 2KVQ; Eco RfaH-NGN: 2OUG; Eco RfaH-KOW: 2LCL; *Pyrococcus furiosus* (Pfu) Spt5-NGN/KOW: 3P8B; *Saccharomyces cerevisiae* (Sce) Spt5-NGN: 2EXU; Sce Spt5-KOW1: 4YTK; Sce Spt5-KOW2/3: 4YTL; *Homo sapiens* (Hsa) DSIF-NGN: 3H7H; Hsa DSIF-KOW1: 5OIK; Hsa DSIF-KOW2: 2E6Z; Hsa DSIF-KOW3: 2DO3; Hsa DSIF-KOW4: 5OHO; Hsa DSIF-KOW5: 2E70. Sce Spt5-KOW1/2/3 have similar structures and are shown in the same color, as are Hsa DSIF-KOW1/2/3/4/5 domains.

the protein (Decker, 2020). Apart from the KOW1 insertion, the NGN and KOW domains from all life have very similar topologies (**Figure 2B**).

All NGN domains make very similar contacts to two conserved RNAP elements (Klein et al., 2011; Martinez-Rucobo et al., 2011; Ehara et al., 2017; Kang et al., 2018), the clamp helices (CH) in the largest RNAP subunit (β' in Bacteria) and the gate loop in the second largest subunit (β in Bacteria).

In addition, some NGNs make sequence-specific contacts to the non-template DNA strand in the transcription bubble of the transcription elongation complex (TEC; see below). The NGN binding site on the TEC is structurally analogous to binding sites of transcription initiation factors in promoter complexes; e.g., bacterial σ factors recognize non-template DNA sequences and an adjacent region on the β' CH during promoter-dependent initiation (Zhang et al., 2012). Consequently, NusG/Spt5 proteins compete with the cognate initiation factors for binding to RNAP, reducing pausing during transcription elongation and potentially facilitating promoter escape (Sevostyanova et al., 2008; Grohmann et al., 2011). Along with the housekeeping NusG present in every free-living cell, many species also contain NusG paralogs (Wang B. et al., 2020) that regulate expression of selected genes in a sequence- or condition-specific fashion.

While the "clamping" contacts between the NGN and TEC are sufficient for NusG/Spt5 effects on RNA synthesis (Mooney et al., 2009b; Hirtreiter et al., 2010), the KOW domains determine their regulatory properties. In *E. coli* NusG, interactions between the KOW domain and Rho facilitate termination (Lawson et al., 2018), whereas the KOW-ribosome interactions couple transcription to translation (Saxena et al., 2018). In eukaryotic Spt5, the presence of multiple KOWs and the CTR, which acts as a hub for recruitment of several RNA processing enzymes and other cellular factors (Decker, 2020), expands the range of regulatory interactions.

SILENCING ABERRANT TRANSCRIPTION

Accurate and timely execution of the gene expression program is essential for cell survival. By itself, RNAP is a passive interpreter of genetic information. Auxiliary proteins instruct RNAP to synthesize RNAs that are required for proper cellular function and prevent it from wasting resources on making useless or potentially harmful RNAs, such as antisense transcripts or mRNAs encoding toxic proteins. In *E. coli*, the housekeeping NusG travels with RNAP transcribing almost all genes (Mooney et al., 2009a), save a few controlled by its paralog RfaH (Belogurov et al., 2009), actively contributing to the transcriptome surveillance. *First*, NusG cooperates with Rho to silence transcription of aberrant RNAs; this is an essential function of *E. coli* NusG (Mitra et al., 2017). *Second*, NusG increases RNAP processivity by modifying properties of the TEC, a shared function of NusG proteins from all life. *Third*, NusG is an integral part of multi-component nucleoprotein complexes that promote facile synthesis and proper assembly of the ribosomal RNAs, and thus the ribosomes. *Finally*, NusG helps to protect

translatable mRNAs from premature release by Rho by bridging the RNAP and the ribosome.

Rho-Dependent Termination

Rho is an ATP-dependent, RecA-type hexameric helicase that terminates transcription of a wide variety of genes in bacteria. Initially viewed as a sequence-specific terminator that requires a C-rich Rho utilization (*rut*) element for loading onto the nascent RNA and subsequent TEC dissociation, Rho has recently emerged as a global multi-functional regulator (Mitra et al., 2017). In addition to its canonical role, inducing termination at the end of some genes (Peters et al., 2012), Rho silences transcriptional noise and expression of horizontally acquired genes, reduces translational stress, and prevents replication-transcription collisions. Genome-wide studies demonstrate that *E. coli* Rho travels with the elongating RNAP, together with NusG and NusA (Mooney et al., 2009a), from the onset of elongation, and acts on numerous cellular targets that lack easily recognizable *rut* sequences (Peters et al., 2012).

To silence AT-rich xenogenes and trigger the release of antisense transcripts or low-quality mRNAs independently of their sequence, Rho relies on help from NusG, which has been implicated in Rho termination at suboptimal, C-less sites (Peters et al., 2012). In a binary system lacking RNAP, NusG activates Rho by promoting isomerization from an open-ring, RNA-loading state, to a closed-ring, translocation-competent state, the transition otherwise triggered by a perfect *rut* element in the RNA (Lawson et al., 2018). The NusG KOW interacts with the C-terminal translocase domain of Rho (**Figure 3**), inducing conformational changes that favor the ring closure even on RNAs devoid of C residues (Lawson et al., 2018). NusG-Rho contacts are mediated by the same KOW region that binds to the ribosomal protein S10 (Burmann et al., 2010), explaining why the translating pioneering ribosome protects the mRNA from a spurious attack by Rho. By contrast, the corresponding Rho-binding residues are missing in RfaH (Lawson et al., 2018), explaining why RfaH does not bind to Rho.

However, the ring closure activity of NusG may not be the main mechanism by which NusG stimulates Rho-dependent termination. Consistent with biochemical data (Schmidt and Chamberlin, 1984; Epshtein et al., 2010) and genome-wide mapping (Mooney et al., 2009a) that support persistent Rho-RNAP interactions, a recent cryo-EM analysis of the *E. coli* TEC under attack by Rho reveals seven complexes thought to represent sequential steps in the termination pathway (Said et al., 2020). During the initial binding to the TEC, Rho makes numerous contacts to the RNAP subunits, NusA and NusG NGN (**Figure 4**), but captures the nascent RNA transcript only later in the pathway. Once engaged, Rho induces dramatic conformational changes in RNAP and Nus factors, which ultimately trap a moribund TEC in which the clamp is wide open and the RNA 3′ end is dislodged from the RNAP active site (Said et al., 2020), a model initially proposed by Nudler and colleagues (Epshtein et al., 2010). In this structurally defined pathway, NusG NGN assists Rho loading onto the RNA and then dissociates to allow for Rho-mediated RNAP clamp opening, whereas NusG KOW is invisible. Remarkably, the Rho ring remains opens even in the moribund TEC, implying that the NusG-promoted Rho helicase activity is required to unwind the RNA:DNA hybrid only after RNAP inactivation; this model is supported by a report that the *E. coli rho* gene becomes dispensable in the presence of a heterologous RNA:DNA helicase (Leela et al., 2013). The allosteric model of termination explains how Rho selectively binds to RNAs that are still being made and reinforces the notion that, even in bacteria, transcriptional regulators act in the context of multi-protein complexes, rather than on RNAP alone.

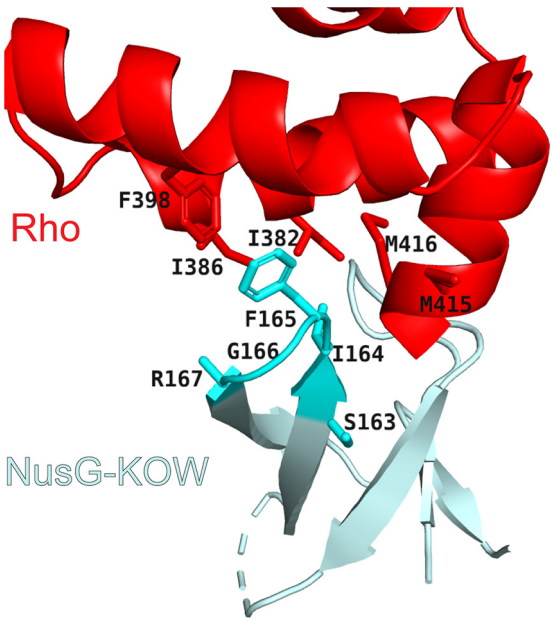


FIGURE 3 | Rho/NusG-KOW interface. Rho residues that contact NusG are shown as red sticks. NusG KOW residues implicated in Rho and S10 binding are shown as cyan sticks; PDB ID: 6DUQ.

Indeed, recent evidence suggests that Rho and NusG cooperate with the histone-like nucleoid-structuring (H-NS) protein, a prototypical xenogeneic silencer, to limit unwanted gene expression. In *E. coli*, Rho and H-NS co-localize on the chromosome (Chandraprakash and Seshasayee, 2014) and mutations in *rho* and *hns* lead to synergistic growth defects (Peters et al., 2012). In *Salmonella,* depletion of NusG leads to massive upregulation of H-NS silenced loci, which include pathogenicity islands and are devoid of *rut* sites; consistently, mutations that compromise Rho-*rut* contacts have no effect on NusG-mediated silencing (Bossi et al., 2019). While the molecular mechanism of this cooperation remains to be determined, it likely reflects RNAP stalling when running into nucleoprotein filaments assembled by H-NS and other nucleoid-associated proteins on the template DNA (Boudreau et al., 2018).

Inhibition of RNAP Pausing

During transcription of cellular DNA, RNAP frequently encounters unfavorable sequences or obstacles, such as DNA-bound proteins or DNA lesions, that slow the enzyme down

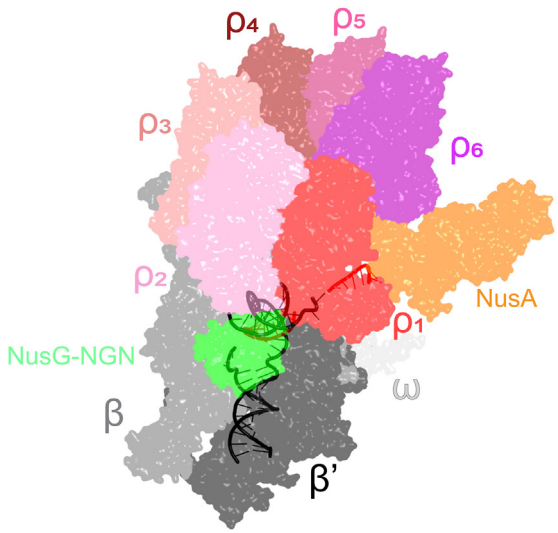


FIGURE 4 | A cryo-EM structure of the Rho engagement complex, in which Rho hexamer makes initial contacts with the transcribing RNAP bound to NusA and NusG. The DNA is shown in black; the RNA—in red. PDB ID: 6Z9P.

or induce arrest. Retrograde movement of the RNAP along the RNA and DNA chains, or backtracking, is a common mechanism of pausing and arrest (Nudler, 2012). Backtracked complexes are rendered inactive because the nascent RNA is extruded through the active site, blocking nucleotide addition (**Figure 5**). The arrested complexes are long-lived, blocking progression of other RNAPs and replisomes, and must be released or reactivated upon transcript cleavage. Cleavage of the backtracked RNA, which is mediated by the RNAP active site and is strongly enhanced by Gre cleavage factors (Sosunova et al., 2003), repositions the 3' end of the RNA in the active site. By preventing backtracking, an activity well-documented in the case of NusG and RfaH (Svetlov et al., 2007; Herbert et al., 2010), NusG-like proteins facilitate processive transcription and promote genome stability. Recent functional and structural data suggest a molecular mechanism of enhanced RNAP processivity, in which the NGN domain loops out the non-template DNA, bringing the upstream and downstream DNA duplexes closer together (Turtola and Belogurov, 2016; Kang et al., 2018; Nedialkov et al., 2018), and establishes contacts to the upstream DNA duplex (Krupp et al., 2019; Said et al., 2020). Together, these interactions alter the upstream DNA trajectory (**Figure 5**) and stabilize the upstream edge of the transcription bubble, which must melt to allow backtracking, explaining how NusG and RfaH inhibit backtracking (Svetlov et al., 2007; Herbert et al., 2010). In addition, the NGN domain, at least in the case of RfaH (Kang et al., 2018), disfavors subtle conformational changes (termed swiveling) that accompany the formation of hairpin-stabilized paused TEC (Kang et al., 2019) and constrains the path of the non-template DNA, preventing it from assuming non-productive conformations (Nedialkov et al., 2018); a similar mechanism has been proposed for yeast Spt5 (Crickard et al., 2016). Together, the NGN-promoted changes in the TEC ensure pause-free RNA synthesis, preventing arrest and termination.

NusG-Assisted Antitermination

To enact RNA surveillance, Rho travels with the elongating RNAP and probes the nascent RNA "translatability." RNAs that contain premature stop codons or are poorly translated, e.g., under conditions of proteotoxic stress, are released by Rho (Richardson, 1991). Yet a very large fraction of cellular RNA is never translated, most notably the most abundant and absolutely essential rRNA which comprises ~50% of the newly synthesized RNA during the exponential growth phase (Dennis et al., 2004). Thus, making rRNA rapidly while protecting it from Rho is key to the survival of cells. Similarly, phage replication is critically dependent on uninterrupted transcription of the phage genome, but Rho is known to broadly silence xenogenes, including phages (Mitra et al., 2017).

Protection of the phage λ early genes and *E. coli* rRNA operons (*rrn*) from Rho is conferred by multicomponent TACs. Recently solved cryo-EM structures of these TACs (**Figure 6**) revealed common and unique details of their action (Krupp et al., 2019; Huang et al., 2020). Both complexes assemble on *boxA* and *boxB* elements in the nascent RNA and share a set of NusABEG factors. Each complex also includes unique factors, N in the λN-TAC and an inositol monophosphatase SuhB dimer + the ribosomal protein S4 in the *rrn*-TAC.

The λN-TAC is resistant to pausing and termination elicited by hairpin signals and Rho. An intrinsically unstructured λN is the principal player which uses a range of mechanisms to modify the TEC (Krupp et al., 2019). λN snakes inside the RNAP, making contacts to multiple RNAP domains and repositioning others, and rearranges Nus factor interactions. λN stabilizes the elongation-competent state of RNAP, inhibiting the nascent RNA hairpin formation and its stabilization by NusA, supports the anti-backtracking and anti-swiveling action of the NusG NGN domain. In the λN-TAC, neither NusG domain can make contacts to Rho observed in the binary Rho-NusG complex (Lawson et al., 2018) and Rho-TEC (Said et al., 2020) structures. Consequently, in the λN-TAC, NusG anti-pausing activity is augmented while its termination-promoting activity is abolished.

Although the *rrn*-TAC has a different protein composition, analogous structural changes inhibit backtracking and NusA-stabilized hairpin pausing and sequester NusG from Rho, with a much larger, well-folded SuhB dimer playing a central role in restructuring of the TAC components instead of λN (Singh et al., 2016). Notably, in addition to promoting pause- and termination-free RNA synthesis, the *rrn*-TAC acts as a molecular chaperone that actively assists the folding and maturation of the nascent RNA (Huang et al., 2020). Similarly to the ribosome-associated chaperones, SuhB, S4 and Nus factors assemble into a ring around the RNA exit channel, extending the channel outward to accommodate a longer segment of the exiting RNA. The RNA is thus sequestered away from the upstream DNA, blocking formation of deleterious R-loops, and is held within a positively charged protein cage to promote folding of local secondary structures and annealing of distant segments, which

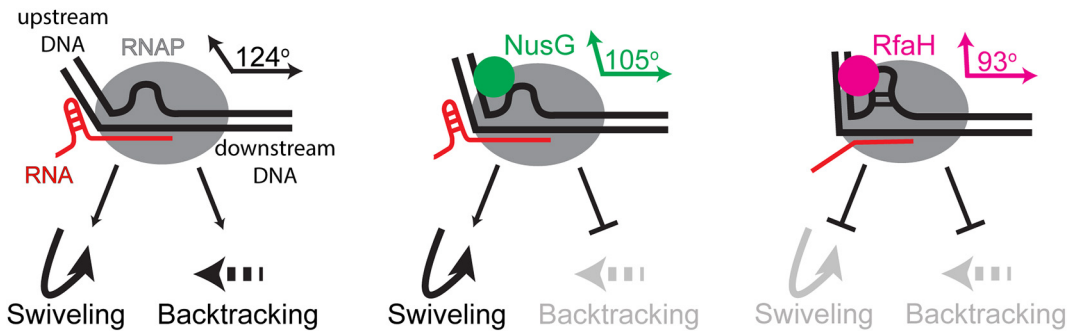


FIGURE 5 | Antipausing activities of *E. coli* NusG and RfaH. Upon encountering a pause-inducing sequence, RNAP can either backtrack or undergo conformational changes termed swiveling; the latter are stabilized by formation of a pause hairpin in the nascent RNA. The NGN domains of both proteins bind near the upstream edge of the transcription bubble, promoting forward and thus inhibiting backward translocation. Transient (NusG) or stable (RfaH) interactions with the non-template DNA strand bring the upstream and downstream DNA duplexes closer together (indicated by angles between these duplexes), an effect that is more pronounced with RfaH. RfaH also binds to the β' and β subunits with higher affinity, restricting the clamp movements to inhibit swiveling and hairpin-stabilized pausing. NusG lacks this activity.

is required for processing of rRNA precursors into mature forms (Young and Steitz, 1978).

NusG plays a supporting role in both TACs: e.g., λN alone has a short-range antitermination activity and requires the TAC assembly to act over long distances (Rees et al., 1996). By contrast, RfaH is a principal, self-sufficient antiterminator: RfaH acts over very long distances yet its activity is not affected by cellular factors, at least *in vitro* (Artsimovitch and Landick, 2002). Other NusG^{SP} may similarly act alone.

Transcription-Translation Coupling

In prokaryotic cells, the lack of a nuclear membrane provides an opportunity for direct physical interaction of the transcribing RNAP and the translating ribosome. The translation-coupled synthesis of the nascent mRNA is known as transcription-translation coupling. The coupling was directly observed by electron microscopy in 1970 in *E. coli* cells (Miller et al., 1970) and subsequently in archaeon *Thermococcus kodakarensis* (French et al., 2007). RNAP and ribosomes form a one-to-one complex with about 1 μM dissociation constant, which is already well within a physiologically relevant range, even in the absence of the nascent mRNA and accessory factors (Fan et al., 2017), resulting in factor-free coupling. Alternatively, the two complexes can be linked by bridging factors, e.g., via the NusG:S10 captured by NMR (Burmann et al., 2010). Substitutions at the *E. coli* NusG:S10 binding interface weakened NusG:S10 association *in vivo* and completely abolished it *in vitro* (Saxena et al., 2018).

The TEC-ribosome complexes, stabilized by general transcription factors, have been observed *in vitro* using cryo-EM (Wang C. et al., 2020; Webster et al., 2020) and analyzed inside cells using a combination of cross-linking mass spectrometry and cryo–electron tomography (O'Reilly et al., 2020). Evidence suggests that coupling may occur initially via direct RNAP:ribosome contacts and then is aided by accessory factors (Washburn et al., 2020). In the NusG/NusA coupled complex, the RNAP β' subunit contacts the 30S subunit protein S3, NusA simultaneously binds to α/β subunits and S2/S5, and finally NusG binds to β/β' and S10 (**Figure 7**). If the ribosome approaches the RNAP further, the collided state, in which the ribosome translocation and the factor-mediated coupling are no longer possible, forms (Wang C. et al., 2020; Webster et al., 2020). Preventing such unproductive collisions may be another function of NusA and NusG.

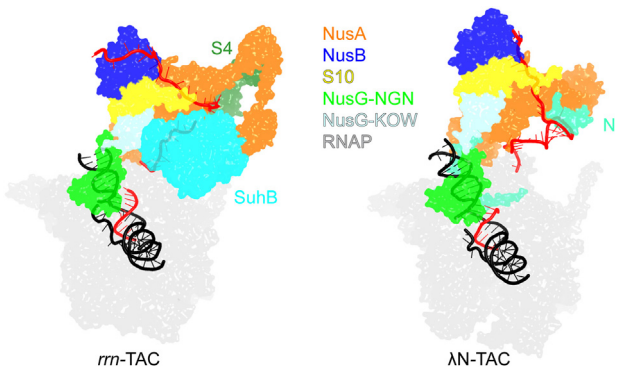


FIGURE 6 | Transcription antitermination complexes (TAC). Left, *rrn*-TAC; right, phage λN-TAC. RNA is in red, DNA is in black. The unique proteins that play key roles in antitermination are shown on each complex; the shared components are indicated in the middle. PDB ID: *rrn*-TAC, 6TQO; λN-TAC, 6GOV.

Since RNAP might often transcribe without a linked ribosome (Chen and Fredrick, 2018), the coupling events must carry important regulatory information (McGary and Nudler, 2013). The closely coupled ribosome prevents the formation of R-loops and RNAP backtracking, thereby promoting genome stability (Gowrishankar and Harinarayanan, 2004; Proshkin et al., 2010; Stevenson-Jones et al., 2020) and inhibits factor-independent termination by blocking the formation of nascent RNA hairpins (Roland et al., 1988). The coupled ribosome also prevents mRNA degradation, by blocking the access of RNaseE (Iost and Dreyfus, 1995), or premature Rho termination, by sequestering NusG and shielding the nascent RNA (Washburn et al., 2020).

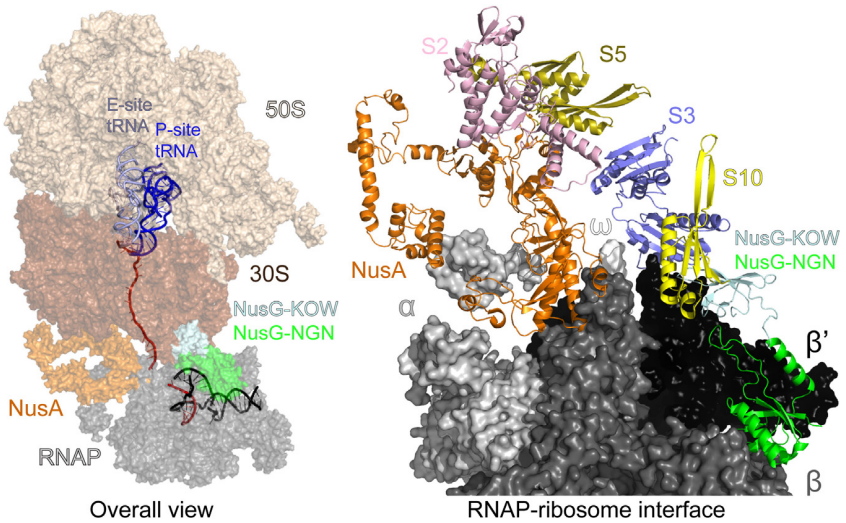


FIGURE 7 | Transcription-translation coupling. Left—an overall view: mRNA is in red, DNA is in black. The interface between the ribosomal 30S subunit and RNAP is stabilized by NusA and NusG. Right—a view of the coupling interface; mRNA, DNA, the entire 50S and most of the 30S subunit have been removed for clarity. PDB ID: 6 × 7F.

When the coupling is broken, e.g., by the ribosome pausing or stalling, Rho releases the nascent RNA, a phenomenon known as polarity (Richardson, 1991). Transcription attenuation is another regulatory mechanism dependent on coupling between the RNAP and the trailing ribosome, wherein the formation of an RNA hairpin induces RNAP pausing and the trailing ribosome pushes the RNAP out of the pause (Turnbough, 2019). By stabilizing the RNAP-ribosome tandem or aiding Rho, NusG controls the fate of the nascent RNA, promoting its translation or release.

B. subtilis (and Its NusG) Is Not at All Like *E. coli*

The universal conservation of the NusG structure and its binding site on the RNAP, as well as perceived common principles of gene expression control in bacteria, justified using the *E. coli* NusG as a paradigm. However, early and recent data suggest that, beyond occupying the same site on RNAP, even housekeeping NusGs, which are encoded within the conserved genomic locus, *secE-nusG-rplK-rplA* in evolutionary distant bacterial phyla (Wang B. et al., 2020), have relatively few common features. Comparison of NusG proteins from *E. coli* and *B. subtilis*, the best studied Gram-negative and Gram-positive model bacteria that grow very similarly in the lab, illustrates these differences.

In wild-type *E. coli*, *nusG* and *rho* genes are essential; their deletions can be obtained only in specially engineered strains (Leela et al., 2013) and confer significant growth defects. In contrast, neither gene is essential in *B. subtilis* (Ingham et al., 1999), in which Rho has limited effects on gene regulation (Nicolas et al., 2012), early stop codons do not induce polarity (Johnson et al., 2020), and most transcription termination is induced by hairpin signals (Mondal et al., 2016; Johnson et al., 2020). In contrast to *E. coli*, where NusG aids Rho in termination of *rut*-less RNAs (Lawson and Berger, 2019), Rho-dependent termination in *B. subtilis* is strongly linked to *cis*-encoded C-rich RNA elements (Johnson et al., 2020). Together, these results suggest that NusG is not involved in gene expression control by Rho in *B. subtilis* (and perhaps other related bacteria) and raise a possibility that an alternative mechanism of transcription noise silencing operates in these species.

Another key function of *E. coli* NusG is bridging the RNAP and the ribosome (**Figure 7**) to mediate transcription-translation coupling, which is thought to occur in all single-compartment cells (see above). In addition to preventing Rho-dependent termination, which may be irrelevant in *B. subtilis*, the coupled ribosome inhibits RNAP backtracking (Proshkin et al., 2010; Stevenson-Jones et al., 2020) and could disfavor the formation of deleterious R-loops (Gowrishankar et al., 2013). The pioneer round of translation may also prime the RNA for subsequent rounds of translation. Strikingly, a recent report demonstrates that transcription and translation are uncoupled in *B. subtilis* (Johnson et al., 2020), where RNAP moves along the template about twice as fast as the ribosome does. While in *E. coli* the coupled ribosome inhibits both intrinsic and Rho-dependent termination, termination in *B. subtilis* is unaffected by translation. The loss of coupling has a profound effect on operon structure: more than 70% of *B. subtilis* intrinsic terminators are positioned just downstream of the stop codon (Johnson et al., 2020), where they would be rendered inefficient by the trailing ribosome in *E. coli* (Roland et al., 1988). These findings are consistent with *in vitro* comparative analysis of *B. subtilis* and *E. coli* RNAP, which shows that *B. subtilis* enzyme transcribes faster and pauses less (Artsimovitch et al., 2000). In contrast, their ribosomes move at similar rates and are unable to catch up with the run-away *B. subtilis* RNAP (Johnson et al., 2020); even if *B. subtilis* NusG binds to the RNAP and the ribosome, it cannot bridge this gap.

In *E. coli*, RNAP pauses frequently and NusG facilitates RNA synthesis (Herbert et al., 2010). By contrast, *B. subtilis* RNAP rarely pauses and NusG stimulates pausing *in vitro* and *in vivo* (Yakhnin et al., 2020a). Unlike *E. coli* NusG, which is positioned next to the non-template DNA strand in the TEC but is not known to recognize any specific DNA elements (Kang et al., 2018), *B. subtilis* NusG specifically binds to T-rich DNA sequences and delays RNA chain elongation (Yakhnin et al., 2016). NusG-dependent RNAP pausing is required for regulation of several operons in *B. subtilis* (Yakhnin et al., 2020b); for example, NusG-dependent pausing in the *trp* and *rib* leader regions provides time for recruitment of an RNA-binding protein TRAP and for riboswitching by flavin mononucleotide, respectively. Sequence-specific pausing through non-template DNA contacts has been first shown for RfaH (Artsimovitch and Landick, 2002), which recognizes 12-nt *ops* elements in the *E. coli* genome (Belogurov et al., 2009); RfaH-induced RNAP delay is thought to facilitate the ribosome recruitment to the nascent RNA (see below) in a handful of leader regions. The *ops* sequence is a perfect match to the consensus pause sequence that induces pausing in *E. coli* (Larson et al., 2014; Vvedenskaya et al., 2014) but has additional recognition determinants for RfaH (Zuber et al., 2018).

By contrast, in *B. subtilis,* NusG recognizes a simpler consensus TTNTTT motif and stimulates pausing genome wide, favoring forward translocation of RNAP (Yakhnin et al., 2020a). Sequences that induce intrinsic, NusG-independent pausing of *B. subtilis* enzyme are also very different from the consensus pause elements documented in *E. coli*, and backtracking is not observed (Yakhnin et al., 2020a). Although the mechanism and regulation of pausing appear to be distinct, slowing RNAP is expected to be essential in both *B. subtilis* and *E. coli*. Pausing determines the overall rate of RNA chain synthesis, is an obligatory step in termination, and facilitates recruitment of regulatory factors (Kang et al., 2019). In both *E. coli* and *B. subtilis*, pausing has been implicated in attenuation control and co-transcriptional folding of riboswitches and catalytic RNAs (Landick et al., 1985; Pan et al., 1999; Perdrizet et al., 2012; Yakhnin et al., 2019), and contributes to coupling of transcription and translation in *E. coli* (McGary and Nudler, 2013). Pausing-defective *E. coli* RNAP variants do not support cell growth but can be rescued by small-molecule ligands that slow the RNAP down (Artsimovitch et al., 2003). In contrast to *E. coli* RNAP, which readily pauses at consensus sequences without the aid of accessory factors (Artsimovitch and Landick, 2000; Larson et al., 2014; Vvedenskaya et al., 2014), *B. subtilis* RNAP relies on NusG to slow it down (Yakhnin et al., 2016, 2020a). In this light, NusG can be viewed as a pause-promoting accessory subunit, a regulatory mechanism that could be widespread in bacteria (Yakhnin et al., 2020b). Indeed, *Thermus thermophilus* NusG reduces the RNA synthesis rate (Sevostyanova and Artsimovitch, 2010) and mycobacterial NusG promotes intrinsic termination (Czyz et al., 2014).

Is there any common function of NusG proteins? The conservation of the *boxA* and *boxB* RNA elements, all Nus factors, ribosomal proteins, and SuhB suggests that similar *rrn*-TACs may form in *B. subtilis*, a hypothesis supported by a report that *rrn* antitermination can be achieved in a heterologous *E. coli*/*B. subtilis* system (Arnvig et al., 2008). Observations that *B. subtilis* cells lacking NusG do not show defects in rRNA transcription argue that NusG is not required for rRNA synthesis (Yakhnin et al., 2020a). However, given that the principal role of the *E. coli rrn*-TACs appears to be in chaperoning of the nascent RNA (Huang et al., 2020), an analogous complex, with or without NusG, may be required to ensure the correct rRNA folding and processing in *B. subtilis*.

A TUSSLE FOR RNAP

In addition to housekeeping NusG/Spt5 proteins present in all free-living cells, many genomes encode one or more NusG paralogs (Wang B. et al., 2020). While the primary sequences of these proteins are very diverse, the high conservation of residues that comprise the high-affinity RNAP binding site suggests that all of them bind to the TEC similarly. Indeed, *E. coli* NusG and RfaH, which are only 17% identical, make very similar contacts to that RNAP β' subunit (Kang et al., 2018). However, in contrast to housekeeping NusG, which binds to RNAP and modulates transcription genome wide (Mooney et al., 2009a; Yakhnin et al., 2020a), these paralogs control expression of just a few target genes. Akin to alternative transcription initiation factors, these specialized NusGs ($NusG^{SP}$) comprise a set of alternative transcription elongation factors that compete for the transcribing RNAP, an analogy further strengthened by their recruitment to the same site on RNAP (Sevostyanova et al., 2008).

However, this analogy does not extend to functions and mechanisms of gene-specific recruitment. Every σ factor activates transcription of its cognate promoters by recruiting RNAP and facilitating DNA melting; just the promoter sequences differ. In a stark contrast, $NusG^{SP}$ factors activate expression of genes that the housekeeping NusG silences (**Figure 8**). These genes can be a few in number, but critical for bacterial evolution and pathogenesis because they encode conjugation and virulence determinants (see below).

Furthermore, while σ factors bind to specific DNA sequences in static promoter complexes, NusG homologs are recruited to a moving RNAP. The available data suggest that these proteins use different recruitment mechanisms, only in some cases relying on specific protein-DNA interactions. Housekeeping NusGs are abundant proteins that can bind the TEC by chance, irrespective of the transcribed sequence; indeed, specific interactions would slow RNAP down, a regulatory feature used in *B. subtilis* (Yakhnin et al., 2020a) but not in *E. coli*, in which NusG is sequence blind. By contrast, the best characterized $NusG^{SP}$, *E. coli* RfaH, uses a very complex mechanism to ensure efficient and selective recruitment to its targets (Zuber et al., 2018). RfaH is recruited to the TEC at operon polarity suppressor (*ops*; **Figure 9A**) sites (16 in *E. coli* MG1655 genome) which are present in leader regions of several operons silenced by NusG and Rho (Artsimovitch and Knauer, 2019). The *ops* element is a composite regulatory signal: it induces RNAP pausing and backtracking (Artsimovitch and Landick, 2000) and is directly recognized by RfaH (Artsimovitch and Landick, 2002). Pausing at *ops* is

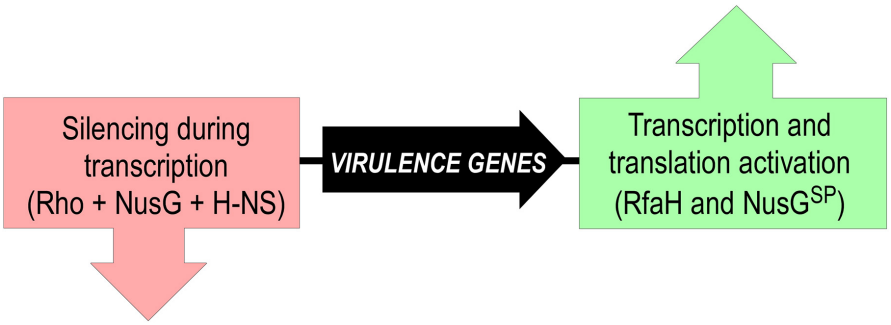


FIGURE 8 | Silencing and counter-silencing of virulence genes by NusG-like proteins.

essential for RfaH recruitment (Zuber et al., 2018): it (*i*) provides additional time for RfaH, which is present in few copies/cell, to find its target; and (*ii*) presents the *ops* bases in a small hairpin, with a conserved T residue flipped out for specific recognition by RfaH (Kang et al., 2018; Zuber et al., 2018). This is a one-time opportunity because, once the RNAP moves past *ops*, the recruitment window is closed; thus, RfaH must bind to RNAP at *ops* and stay bound until the end of RNA synthesis. To fend off 100-fold more abundant NusG (Schmidt et al., 2016), RfaH binds RNAP much tighter (Kang et al., 2018), essentially becoming an RNAP subunit for one round of RNA synthesis. RfaH maintains the ability to trigger pausing at a downstream (engineered) *ops* site while traveling with RNAP but reduces pausing at any other sequence (Belogurov et al., 2009).

If RfaH binds to RNAP very tightly during elongation, why does it need the *ops* signal in the first place? Unlike NusG, in which the RNAP-binding site on the NGN domain is exposed, this site is blocked by the KOW domain in free RfaH (**Figure 9**). Also unlike NusG, in which the KOW domain is in a β-barrel state (β-KOW; **Figure 2**), in this "autoinhibited" RfaH the KOW domain is folded as an α-helical hairpin (α-KOW; **Figure 9B**). To bind RNAP, RfaH must be "activated" by domain dissociation, which happens only in the presence of a complete *ops*-paused TEC (Zuber et al., 2019). The details of this process remain elusive, but the current model suggests that the NGN domain recognizes the *ops* hairpin *via* its exposed DNA-binding residues, forming a transient encounter complex and triggering the KOW dissociation (Artsimovitch and Knauer, 2019). It is possible that autoinhibition may be a common feature of NusG homologs. While in *E. coli* NusG the NGN and KOW domains move freely (Burmann et al., 2011), in NusG from a hyperthermophilic bacterium *Thermotoga maritima*, the two domains interact, masking the binding sites for RNAP, NusE, and Rho (Drögemüller et al., 2013). Domain dissociation enables *T. maritima* NusG-KOW binding to Rho and NusE, and these contacts may be stabilized by the NGN-RNAP contacts (Drögemüller et al., 2017).

RfaH recruitment relies on the multi-functional DNA element and elaborate structural rearrangements of the protein domains. Binding to a specific DNA element enables RfaH to control several operons scattered on the chromosome. But how is a wannabe NusG^{SP}, which has just surfaced following gene duplication, targeted to a specific locus in the presence of overwhelming numbers of NusG molecules? An "ancestral" mechanism, in which NusG^{SP} binds to the transcribing RNAP *in cis* has been proposed to explain this conundrum (Belogurov et al., 2009). This model is supported by bioinformatics analyses which reveal that the residues that mediate DNA contacts in RfaH arose late in evolution and that many NusG^{SP} are encoded within long xenogeneic operons, in contrast to the standalone *rfaH* gene (Wang B. et al., 2020). However, observations that some of these *cis*-encoded regulators act *in trans* (Chatzidaki-Livanis et al., 2010) suggest that NusG^{SP} recruitment strategies are multifaceted.

STRUCTURAL TRANSFORMATION OF RfaH

RfaH activation is not limited to the domain dissociation needed to expose the RNAP-binding site: the released α-KOW undergoes a dramatic transformation into a NusG-like β-KOW (**Figure 9B**) and binds to S10 similarly to NusG KOW (Burmann et al., 2012). The residues that make contacts with S10 are not available in the α-KOW domain, thus the free RfaH is autoinhibited with respect to both RNAP and ribosome binding, allowing RfaH to achieve high target specificity (Shi et al., 2017). The activated state persists until the TEC dissociates at a terminator and RfaH is released; the KOW then refolds into the α-helical hairpin and re-establishes contacts with the NGN, restoring autoinhibition (**Figure 9A**).

Interconversion between the alternative RfaH-KOW states is principally controlled by interdomain contacts: the KOW (re)folds into a β-barrel when expressed alone, separated from the NGN domain upon proteolytic cleavage of the linker, or as a result of interface-destabilizing substitutions (Burmann et al., 2012; Tomar et al., 2013; Shi et al., 2017). Deuteron incorporation reveals that the tip of the C-terminal α-hairpin is stably folded in the autoinhibited state, whereas the rest of the KOW is highly flexible, and its flexibility only decreases in the β-folded state (Galaz-Davison et al., 2020). The mechanism underlying this dramatic fold switch has been also pursued by computational approaches (Gc et al., 2014, 2015;

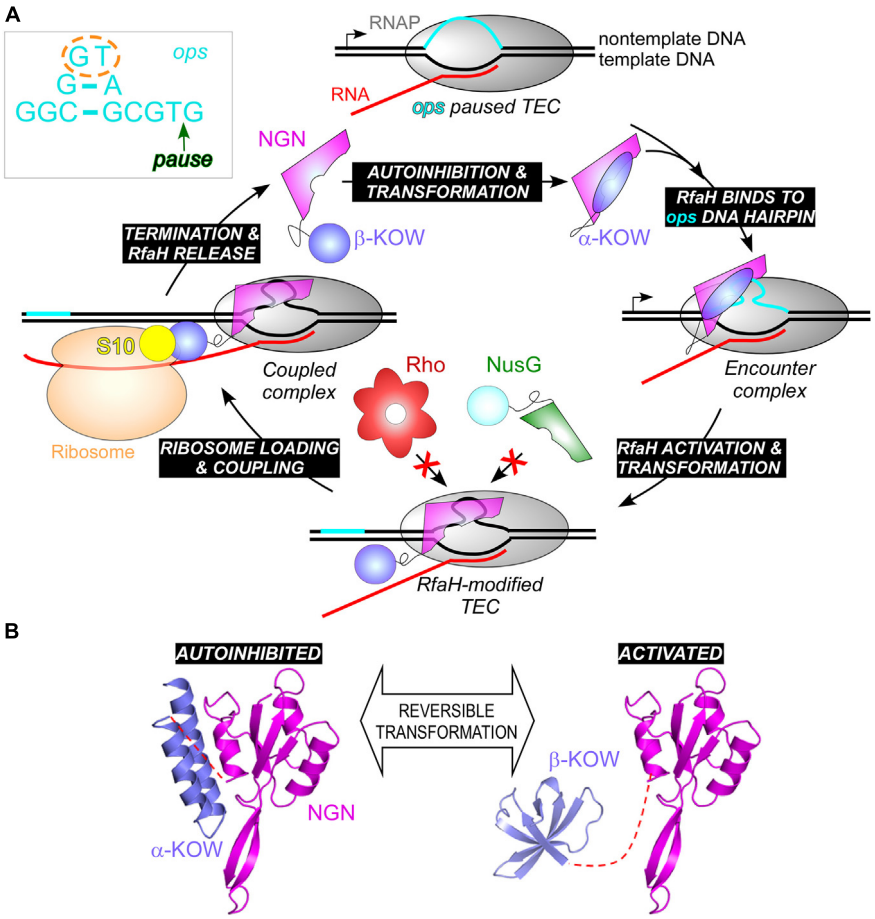


FIGURE 9 | (A) A full cycle of RfaH; see text for details. The inset shows the *ops* DNA element, which forms a short hairpin on the TEC surface; *ops* bases that make most interactions with RfaH in the complex are circled; the pause position is indicated by an arrow. **(B)** RfaH domain dissociation and refolding. PDB IDs: autoinhibited RfaH, 5OND; activated RfaH, 6C6S.

Balasco et al., 2015; Ramírez-Sarmiento et al., 2015; Xiong and Liu, 2015; Xun et al., 2016). Although the β-barrel is a preferred state of the isolated RfaH-KOW, its free energy is only slightly lower than that of the α-helical conformation. The separation of the two alternative states is dependent on large energy barriers resulting from the main chain hydrogen bonds of the α-helical hairpin. An all-atom Monte Carlo simulations study suggests a possibility that the encounter complex between the autoinhibited RfaH and the *ops*-TEC is characterized by net attractive interactions with the NGN and net repulsive interactions with the KOW. The resulting opposing forces on the two domains, in combination with the peculiar mechanical rigidity profile of the autoinhibited RfaH, might help trigger domain separation (Seifi et al., 2020). The α → β rearrangement essentially depends on an unstructured state: upon dissolution of the α-helical hairpin, the KOW assumes a disordered state and then follows a step-wise assembly into the final five-stranded β-barrel (Bernhardt and Hansmann, 2018; Joseph et al., 2019).

Among NusG homologs, *E. coli* RfaH is the only known transformer protein. However, it is possible that other KOW domains are capable of transformation. In particular, an amazingly broad repertoire of known cellular targets of eukaryotic NusG homologs (Decker, 2020) could be due to metamorphic behavior of their KOWs.

RfaH AS A TRANSLATION FACTOR

RfaH-controlled genes encode toxins, adhesins, LPS and capsule biosynthesis enzymes, type IV secretion apparatus, etc. located in long horizontally acquired operons (**Figure 10**), which are silenced by Rho. RfaH abolishes Rho-dependent termination (Sevostyanova et al., 2011) and the ability to bind Rho appears to be lost early in RfaH evolution (Wang B. et al., 2020). RfaH elicits dramatic, 50 + fold activation of gene expression *in vivo,* an effect that was initially assumed to be mediated by its direct antitermination effects on RNAP (Artsimovitch and Landick, 2002). Surprisingly, RNAP modification by RfaH makes only a minor contribution in the cell (Sevostyanova et al., 2011). Instead, RfaH inhibits Rho-dependent termination by outcompeting NusG and activating translation.

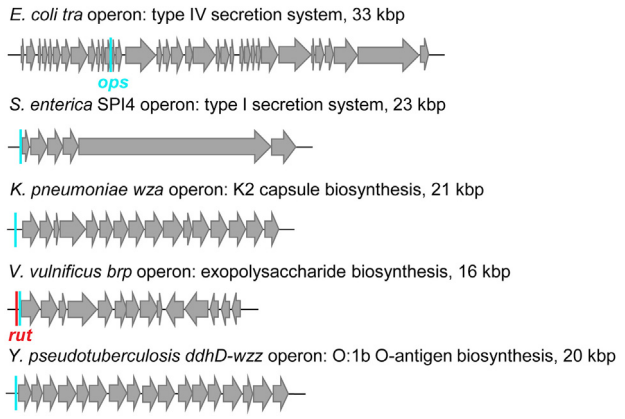


FIGURE 10 | Examples of RfaH-controlled operons. Positions of *ops* sites (cyan bars) and a *rut* site (red bar) are indicated.

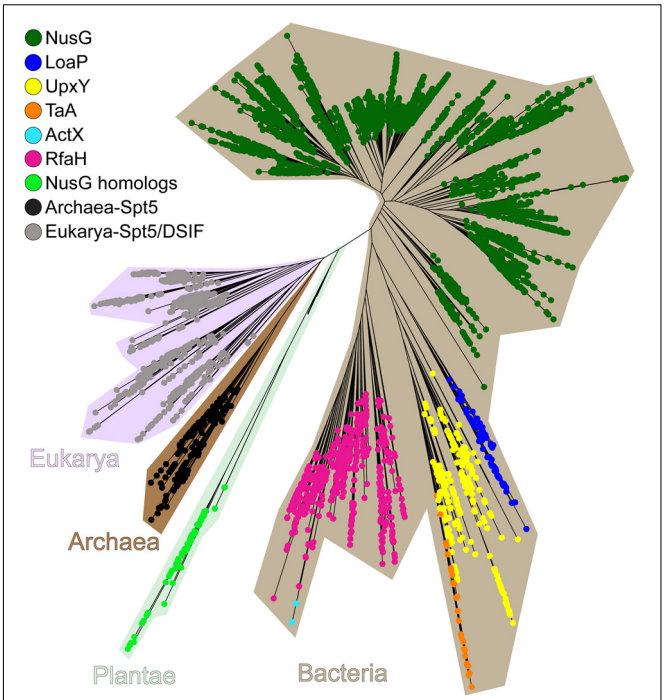


FIGURE 11 | The maximum-likelihood phylogenetic tree of NusG-like proteins. Plantae is an artificial group used solely for brevity.

RfaH-controlled genes lack Shine-Dalgarno elements, which recruit the ribosome through RNA base-pairing with the 16S rRNA (Rodnina, 2018) and have many rare codons, limiting their translation and making them easy targets for Rho. Observations that the transformed β-KOW directly binds S10 (Burmann et al., 2012) prompted a hypothesis that RfaH recruits the ribosome via β-KOW/S10 contacts and then couples transcription to translation during elongation. In support of this model, expression of SD-less reporters is completely dependent on RfaH, and substitutions of residues that interact with S10 abolish expression (Burmann et al., 2012). In addition to the ribosome recruitment, by bridging the RNAP and the ribosome during elongation, RfaH may prevent uncoupling at rare codons; the ribosome stalling exposes mRNA to Rho (Elgamal et al., 2016). RfaH may be particularly important during synthesis of excessively long proteins such as *Salmonella* pathogenicity island IV giant 600 kDa adhesin (**Figure 9B**), which requires RfaH for expression (Main-Hester et al., 2008). Remarkably, the *ops*-RfaH module supports efficient expression of an SD-less reporter *in vivo*, ∼20% relative to that driven by a perfect SD element (Burmann et al., 2012).

Although RfaH and NusG make similar contacts to S10 (Burmann et al., 2012), their effects on translation are expected to be different. NusG binds to the RNAP transiently (Kang et al., 2018) and late in the operon, well after the first ORF (Mooney et al., 2009a). In contrast, RfaH binds to RNAP upstream of the first ORF and remains stably associated with the EC until termination (Belogurov et al., 2009). It is possible that RfaH recruits the ribosome to the *ops*-paused RNAP and promotes ribosome scanning for a downstream initiation codon. Future studies will reveal the details of translation activation by RfaH, but the available data suggest that this universally conserved transcription antiterminator may be acting primarily as an RNAP-tethered translation initiation/elongation factor and may employ the first protein-mediated ribosome recruitment mechanism outside of viruses.

DIVERSITY OF THE NusG FAMILY

Specialized NusG paralogs (**Figure 11**) are evolving in very different ecological niches but may have similar functions—to promote expression of long or silenced operons. Functional data implicate several NusG^{SP} in transcription antitermination of very long gene clusters, whereas for others this function is inferred from their genomic associations. *Bacillus amyloliquefaciens* LoaP inhibits termination in two operons producing antibiotics difficidin and macrolactin (Goodson et al., 2017). Differently from RfaH, which is rather inefficient against intrinsic terminators (Artsimovitch and Landick, 2002; Carter et al., 2004), LoaP promotes readthrough of the hairpin termination signals (Goodson et al., 2017). Polyketide antibiotic TA made by *Myxococcus xanthus* inhibits bacterial cell wall synthesis and is produced by a 40 kb operon which is activated by NusG^{SP} called TaA (Paitan et al., 1999) by an unknown mechanism. Human gut bacterium *Bacteroides fragilis* synthesizes eight capsular polysaccharides from separate operons, which are activated by UpxY family of NusG^{SP}. UpxY proteins prevent premature transcriptional termination within the 5′ leaders upstream from the *upxY* gene (Chatzidaki-Livanis et al., 2009).

While functional data are available for just a few NusG^{SP}, recent bioinformatics analysis suggests that these proteins fall into eight different clusters, which differ in their primary sequence signatures as well as regulatory contexts. Some NusG^{SP}, such as RfaH, form one group and are encoded by single cistrons, whereas others (e.g., *loaP*, *taA*, and *upxY*) are adjacent to their target operons (Wang B. et al., 2020). ActX, which

is closely related to RfaH (**Figure 11**), is encoded within pilus biosynthesis operons on antibiotic-resistant plasmids in *E. coli* and *Klebsiella pneumoniae* (Núñez et al., 1997), but its regulatory function remains unknown. Analysis of genomic contexts can be instrumental in predicting functional associations (Moreno-Hagelsieb and Santoyo, 2015). Gene neighbors of NusG^{SP} (except for RfaH-like stand-alone genes) are enriched in genes involved in cell envelope biogenesis, with glycosyltransferases, nucleoside-diphosphate-sugar epimerases, and exopolysaccharide biosynthesis enzymes being the most common (Wang B. et al., 2020). However, notable differences exist among distinct clusters; for example, some NusG^{SP} are adjacent to Tat protein secretion system, others are encoded near undecaprenyl pyrophosphate synthase and H-NS genes. A group of regulators from *Shewanella* are encoded within putative exopolysaccharide operons, an arrangement resembling *B. fragilis* operons controlled by UpxY proteins (Chatzidaki-Livanis et al., 2010). Future studies will be required to determine functional significance of these associations.

Extensive duplications, sub-functionalization, and horizontal transfer underpin the evolution of NusG paralogs. One NusG copy has gradually evolved into RfaH, starting from an "early" loss of binding to Rho terminator while tightening contacts to RNAP and culminating with the "late" acquisition of residues that interact with the *ops* DNA element and confer autoinhibition (Wang B. et al., 2020). While in most NusG homologs these changes do not alter the core domain structure, some factors acquired additional domains thought to promote adaptation to their unique niches. For example, in *T. maritima* NusG, an extra domain DII supports NusG recruitment to the TEC and stabilizes the NusG:RNAP complex, a necessary adaptation to high temperatures in the *T. maritima* natural habitat (Drögemüller et al., 2017).

In addition to Spt5, NusG homologs are also encoded in the genomes of all major land plant and algal lineages except for some green algal species (Wang B. et al., 2020). These bacterial regulators have recognizable chloroplast-localization signals and are presumably retained to assist the bacterial-type RNAPs that mediate chloroplast transcription. A NusG homolog *of Arabidopsis thaliana* has been identified as a component of the active transcriptional machinery in chloroplasts (Pfalz et al., 2006), and a Rho ortholog has been shown to terminate transcription by plastid-encoded RNAP (Yang et al., 2020).

NusG PARALOGS AND VIRULENCE

Extensive functional studies have established RfaH as the paradigm for the regulation of transcription elongation, translation initiation, and protein folding. However, RfaH is also a key virulence factor. RfaH activates the expression of capsule, cell wall, toxins, adhesins, and pilus biosynthesis operons (**Figure 9B**), which are important for virulence and conjugal transfer in several Gram-negative pathogens including *E. coli*, *K. pneumoniae*, *Vibrio vulnificus*, *Salmonella enterica*, *Yersinia pseudotuberculosis*, and *Yersinia pestis* (Kong et al., 2011; Bachman et al., 2015; Garrett et al., 2016; Hoffman et al., 2017). RfaH effects on gene expression are very large (50+ fold); consequently, the loss of *rfaH* leads to dramatic defects in virulence, e.g., 10^4 decrease in *K. pneumoniae* survival in the lung (Bachman et al., 2015).

The first protein secretion process discovered in bacteria was the hemolysin A (HlyA) type 1 secretion system (T1SS), which is found in uropathogenic *E. coli* strains (Thomas et al., 2014). HlyA is a 107 kDa protein that induces hemolysis by creating pores in the erythrocyte membrane (Skals et al., 2009). RfaH, a.k.a. HlyT, has been identified genetically as an activator of the *hly* operon (Thomas et al., 2014). Inactivation of *rfaH* dramatically decreases virulence of uropathogenic *E. coli* strain in a murine model of urinary tract infection (Nagy et al., 2002). The capability to colonize the intestinal tract by efficiently competing with the commensal microbiota has been considered as a multifactorial virulence property. RfaH also plays a role in the infectious process during colonization of the intestinal tract: *rfaH* mutants are susceptible to bile salts and show reduced gut colonization capacity (Nagy et al., 2005).

Antibiotic-resistant *K. pneumoniae* is an urgent public health threat and a leading cause of pneumonia in hospitalized patients (David et al., 2019). Functional genomic profiling of four diverse serum-resistant *K. pneumoniae* strains reveals that the deletion of *rfaH* dramatically reduces resistance to serum complement system in all strains (Short et al., 2020). *Vibrio vulnificus* is another opportunistic human pathogen responsible for the majority of seafood-associated deaths worldwide, and antibiotic resistance has developed (Heng et al., 2017). Loss of *rfaH* also makes *V. vulnificus* highly sensitive to human serum (Garrett et al., 2016). Expression of the *brp* exopolysaccharide operon mediates surface adherence of *V. vulnificus*, and the presence of *ops* and *rut* sites in the leader region suggests RfaH-dependent antitermination (Chodur and Rowe-Magnus, 2018). *S. enterica* serovar Typhimurium is a primary enteric pathogen infecting both humans and animals and a major cause of diarrheal diseases, with antibiotic resistance on the rise (Fàbrega and Vila, 2013; Knodler and Elfenbein, 2019). *Salmonella* harbors five pathogenicity islands (SPI) required for infection in vertebrate hosts. Among them, SPI4 plays a role in the initial interaction with the intestinal epithelium and possibly contributes to long-term persistence (Gerlach et al., 2007). *S. enterica* RfaH is required for the expression of SPI4, which encodes a T1SS and its adhesin substrate (Main-Hester et al., 2008), as well as the expression of secreted and surface-associated polysaccharides (Lindberg and Hellerqvist, 1980; Bailey et al., 1997). Mutants of *S. enterica* serovar Typhimurium lacking *rfaH* are efficient as vaccines against salmonellosis and induce strong serum immune responses (Nagy et al., 2006; Liu et al., 2016). Given their association with capsular and TSS operons (Wang B. et al., 2020), other NusG paralogs likely play important roles during pathogenesis.

Antibiotic resistance determinants are frequently encoded on conjugative plasmids and can be rapidly transferred between bacteria (Wang et al., 2017). RfaH activates the F plasmid conjugation operon (Beutin and Achtman, 1979) and RfaH homologs are encoded on some clinical resistant plasmids (Wang B. et al., 2020), suggesting that they may contribute to plasmid

transfer. A recent study showed that deletions of seven genes, including *rfaH*, prevented cefotaxime-induced up-regulation of *traF* and decreased the conjugative transfer of the resistance plasmid (Liu et al., 2019).

RfaH proteins from *Vibrio*, *Yersinia*, *Salmonella*, and *Klebsiella* bind to the *E. coli* TEC *in vitro* and complement the *E. coli rfaH* gene deletion (Carter et al., 2004). Small molecule inhibitors that block recruitment of *E. coli* and *K. pneumoniae* RfaH to RNAP (Svetlov et al., 2018) may have a potential to inhibit virulence and the spread of antibiotic resistance.

CONCLUDING REMARKS

NusG homologs comprise the only universally conserved family of transcription factors, which includes housekeeping regulators and their specialized paralogs (**Figure 11**). Despite highly similar core domain architectures and interactions with RNAP, NusG-like proteins exert amazingly diverse, and frequently opposite, effects on gene expression. Bacterial NusG homologs can inhibit or stimulate transcription termination, accelerate RNA synthesis by suppressing RNAP backtracking or slow transcription down by halting RNAP at specific sequences, bridge the RNAP to the ribosome during translation elongation or recruit the ribosome to mRNAs that lack canonical ribosome binding sites, and likely perform other functions that remain to be discovered.

This regulatory plasticity depends on dynamic interactions of the NGN and KOW domains with each other, RNAP, single and double-stranded nucleic acids, and many auxiliary cellular proteins. While bound to the TEC through contacts mediated by highly conserved residues within RNAP and NGN, NusG homologs employ divergent residues in their NGN and KOW domains to enact a range of responses demanded by specific cellular circumstances. Some NusG paralogs augment their regulatory prowess by undergoing an unprecedented and reversible refolding of an entire KOW domain, during which the protein turns inside out. The presence of NusG in all free-living organisms, sometimes in several copies, confirms its unique place in gene expression control, from LUCA to present life forms.

AUTHOR CONTRIBUTIONS

BW prepared all original figures and wrote the first draft. IA revised and expanded the manuscript. Both authors prepared figures and edited the draft while preparing a revised manuscript.

FUNDING

Our research was supported by the National Institutes of Health (GM067153 to IA).

REFERENCES

Arnvig, K. B., Zeng, S., Quan, S., Papageorge, A., Zhang, N., Villapakkam, A. C., et al. (2008). Evolutionary comparison of ribosomal operon antitermination function. *J. Bacteriol.* 190, 7251–7257. doi: 10.1128/jb.00760-08

Artsimovitch, I., Chu, C., Lynch, A. S., and Landick, R. (2003). A new class of bacterial RNA polymerase inhibitor affects nucleotide addition. *Science* 302, 650–654. doi: 10.1126/science.1087526

Artsimovitch, I., and Knauer, S. H. (2019). Ancient transcription factors in the news. *mBio* 10:e01547-18.

Artsimovitch, I., and Landick, R. (2000). Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. *Proc. Natl. Acad. Sci. U.S.A.* 97, 7090–7095. doi: 10.1073/pnas.97.13.7090

Artsimovitch, I., and Landick, R. (2002). The transcriptional regulator RfaH stimulates RNA chain synthesis after recruitment to elongation complexes by the exposed nontemplate DNA strand. *Cell* 109, 193–203. doi: 10.1016/s0092-8674(02)00724-9

Artsimovitch, I., Svetlov, V., Anthony, L., Burgess, R. R., and Landick, R. (2000). RNA polymerases from *Bacillus subtilis* and *Escherichia coli* differ in recognition of regulatory signals in vitro. *J. Bacteriol.* 182, 6027–6035. doi: 10.1128/jb.182.21.6027-6035.2000

Bachman, M. A., Breen, P., Deornellas, V., Mu, Q., Zhao, L., Wu, W., et al. (2015). Genome-wide identification of *Klebsiella pneumoniae* fitness genes during lung infection. *mBio* 6:e00775-15.

Bailey, M. J., Hughes, C., and Koronakis, V. (1997). RfaH and the ops element, components of a novel system controlling bacterial transcription elongation. *Mol. Microbiol.* 26, 845–851. doi: 10.1046/j.1365-2958.1997.6432014.x

Balasco, N., Barone, D., and Vitagliano, L. (2015). Structural conversion of the transformer protein RfaH: new insights derived from protein structure prediction and molecular dynamics simulations. *J. Biomol. Struct. Dyn.* 33, 2173–2179. doi: 10.1080/07391102.2014.994188

Belogurov, G. A., and Artsimovitch, I. (2015). Regulation of transcript elongation. *Annu. Rev. Microbiol.* 69, 49–69. doi: 10.1146/annurev-micro-091014-104047

Belogurov, G. A., and Artsimovitch, I. (2019). The mechanisms of substrate selection, catalysis, and translocation by the elongating RNA polymerase. *J. Mol. Biol.* 431, 3975–4006. doi: 10.1016/j.jmb.2019.05.042

Belogurov, G. A., Mooney, R. A., Svetlov, V., Landick, R., and Artsimovitch, I. (2009). Functional specialization of transcription elongation factors. *EMBO J.* 28, 112–122. doi: 10.1038/emboj.2008.268

Bernhardt, N. A., and Hansmann, U. H. E. (2018). Multifunnel landscape of the fold-switching protein RfaH-CTD. *J. Phys. Chem. B* 122, 1600–1607. doi: 10.1021/acs.jpcb.7b11352

Beutin, L., and Achtman, M. (1979). Two *Escherichia coli* chromosomal cistrons, sfrA and sfrB, which are needed for expression of F factor tra functions. *J. Bacteriol.* 139, 730–737. doi: 10.1128/JB.139.3.730-737.1979

Bossi, L., Ratel, M., Laurent, C., Kerboriou, P., Camilli, A., Eveno, E., et al. (2019). NusG prevents transcriptional invasion of H-NS-silenced genes. *PLoS Genet.* 15:e1008425. doi: 10.1371/journal.pgen.1008425

Boudreau, B. A., Hron, D. R., Qin, L., van der Valk, R. A., Kotlajich, M. V., Dame, R. T., et al. (2018). StpA and Hha stimulate pausing by RNA polymerase by promoting DNA-DNA bridging of H-NS filaments. *Nucleic Acids Res.* 46, 5525–5546. doi: 10.1093/nar/gky265

Burmann, B. M., Knauer, S. H., Sevostyanova, A., Schweimer, K., Mooney, R. A., Landick, R., et al. (2012). An α Helix to β Barrel domain switch transforms the transcription factor RfaH into a translation factor. *Cell* 150, 291–303. doi: 10.1016/j.cell.2012.05.042

Burmann, B. M., Scheckenhofer, U., Schweimer, K., and Rosch, P. (2011). Domain interactions of the transcription-translation coupling factor *Escherichia coli* NusG are intermolecular and transient. *Biochem. J.* 435, 783–789. doi: 10.1042/BJ20101679

Burmann, B. M., Schweimer, K., Luo, X., Wahl, M. C., Stitt, B. L., Gottesman, M. E., et al. (2010). A NusE:NusG complex links transcription and translation. *Science* 328, 501–504. doi: 10.1126/science.1184953

Burns, C. M., and Richardson, J. P. (1995). NusG is required to overcome a kinetic limitation to Rho function at an intragenic terminator. *Proc. Natl. Acad. Sci. U.S.A.* 92, 4738–4742. doi: 10.1073/pnas.92.11.4738

Burns, C. M., Richardson, L. V., and Richardson, J. P. (1998). Combinatorial effects of NusA and NusG on transcription elongation and Rho-dependent termination in *Escherichia coli*. *J. Mol. Biol.* 278, 307–316. doi: 10.1006/jmbi.1998.1691

Burova, E., Hung, S. C., Sagitov, V., Stitt, B. L., and Gottesman, M. E. (1995). *Escherichia coli* NusG protein stimulates transcription elongation rates in vivo and in vitro. *J. Bacteriol.* 177, 1388–1392. doi: 10.1128/jb.177.5.1388-1392.1995

Cardinale, C. J., Washburn, R. S., Tadigotla, V. R., Brown, L. M., Gottesman, M. E., and Nudler, E. (2008). Termination factor Rho and its cofactors NusA and NusG silence foreign DNA in *E. coli*. *Science* 320, 935–938. doi: 10.1126/science.1152763

Carter, H. D., Svetlov, V., and Artsimovitch, I. (2004). Highly divergent RfaH orthologs from pathogenic proteobacteria can substitute for *Escherichia coli* RfaH both in vivo and in vitro. *J. Bacteriol.* 186, 2829–2840. doi: 10.1128/jb.186.9.2829-2840.2004

Casjens, S. R., and Hendrix, R. W. (2015). Bacteriophage lambda: early pioneer and still relevant. *Virology* 47, 310–330. doi: 10.1016/j.virol.2015.02.010

Chandraprakash, D., and Seshasayee, A. S. (2014). Inhibition of factor-dependent transcription termination in *Escherichia coli* might relieve xenogene silencing by abrogating H-NS-DNA interactions in vivo. *J. Biosci.* 39, 53–61. doi: 10.1007/s12038-014-9413-4

Chatzidaki-Livanis, M., Coyne, M. J., and Comstock, L. E. (2009). A family of transcriptional antitermination factors necessary for synthesis of the capsular polysaccharides of *Bacteroides fragilis*. *J. Bacteriol.* 191, 7288–7295. doi: 10.1128/jb.00500-09

Chatzidaki-Livanis, M., Weinacht, K. G., and Comstock, L. E. (2010). Trans locus inhibitors limit concomitant polysaccharide synthesis in the human gut symbiont *Bacteroides fragilis*. *Proc. Natl. Acad. Sci. U.S.A.* 107, 11976–11980. doi: 10.1073/pnas.1005039107

Chen, M., and Fredrick, K. (2018). Measures of single- versus multiple-round translation argue against a mechanism to ensure coupling of transcription and translation. *Proc. Natl. Acad. Sci. U.S.A.* 115, 10774–10779. doi: 10.1073/pnas.1812940115

Chodur, D. M., and Rowe-Magnus, D. A. (2018). Complex control of a genomic island governing biofilm and rugose colony development in *Vibrio vulnificus*. *J. Bacteriol.* 200:e00190-18.

Crickard, J. B., Fu, J., and Reese, J. C. (2016). Biochemical analysis of Yeast suppressor of Ty 4/5 (Spt4/5) reveals the importance of nucleic acid interactions in the prevention of RNA polymerase II arrest. *J. Biol. Chem.* 291, 9853–9870. doi: 10.1074/jbc.M116.716001

Czyz, A., Mooney, R. A., Iaconi, A., and Landick, R. (2014). Mycobacterial RNA polymerase requires a U-tract at intrinsic terminators and is aided by NusG at suboptimal terminators. *mBio* 5:e00931-14.

David, S., Reuter, S., Harris, S. R., Glasner, C., Feltwell, T., Argimon, S., et al. (2019). Epidemic of carbapenem-resistant *Klebsiella pneumoniae* in Europe is driven by nosocomial spread. *Nat. Microbiol.* 4, 1919–1929. doi: 10.1038/s41564-019-0492-8

Decker, T.-M. (2020). Mechanisms of transcription elongation factor DSIF (Spt4–Spt5). *J. Mol. Biol.* doi: 10.1016/j.jmb.2020.09.016 [Epub ahead of print].

Dennis, P. P., Ehrenberg, M., and Bremer, H. (2004). Control of rRNA synthesis in *Escherichia coli*: a systems biology approach. *Microbiol. Mol. Biol. Rev.* 68, 639–668. doi: 10.1128/mmbr.68.4.639-668.2004

Downing, W. L., Sullivan, S. L., Gottesman, M. E., and Dennis, P. P. (1990). Sequence and transcriptional pattern of the essential *Escherichia coli* secE-nusG operon. *J. Bacteriol.* 172, 1621–1627. doi: 10.1128/jb.172.3.1621-1627.1990

Drögemüller, J., Schneider, C., Schweimer, K., Strauß, M., Wöhrl, B. M., Rösch, P., et al. (2017). *Thermotoga maritima* NusG: domain interaction mediates autoinhibition and thermostability. *Nucleic Acids Res.* 45, 446–460. doi: 10.1093/nar/gkw1111

Drögemüller, J., Stegmann, C. M., Mandal, A., Steiner, T., Burmann, B. M., Gottesman, M. E., et al. (2013). An autoinhibited state in the structure of *Thermotoga maritima* NusG. *Structure* 21, 365–375. doi: 10.1016/j.str.2012.12.015

Ehara, H., Yokoyama, T., Shigematsu, H., Yokoyama, S., Shirouzu, M., and Sekine, S. I. (2017). Structure of the complete elongation complex of RNA polymerase II with basal factors. *Science* 357, 921–924. doi: 10.1126/science.aan8552

Elgamal, S., Artsimovitch, I., and Ibba, M. (2016). Maintenance of transcription-translation coupling by elongation factor P. *mBio* 7:e01373-16.

Epshtein, V., Dutta, D., Wade, J., and Nudler, E. (2010). An allosteric mechanism of Rho-dependent transcription termination. *Nature* 463, 245–249. doi: 10.1038/nature08669

Fàbrega, A., and Vila, J. (2013). *Salmonella enterica* serovar Typhimurium skills to succeed in the host: virulence and regulation. *Clin. Microbiol. Rev.* 26, 308–341. doi: 10.1128/cmr.00066-12

Fan, H., Conn, A. B., Williams, P. B., Diggs, S., Hahm, J., Gamper, H. B., et al. (2017). Transcription–translation coupling: direct interactions of RNA polymerase with ribosomes and ribosomal subunits. *Nucleic Acids Res.* 45, 11043–11055. doi: 10.1093/nar/gkx719

French, S. L., Santangelo, T. J., Beyer, A. L., and Reeve, J. N. (2007). Transcription and translation are coupled in Archaea. *Mol. Biol. Evol.* 24, 893–895. doi: 10.1093/molbev/msm007

Galaz-Davison, P., Molina, J. A., Silletti, S., Komives, E. A., Knauer, S. H., Artsimovitch, I., et al. (2020). Differential local stability governs the metamorphic fold switch of bacterial virulence factor RfaH. *Biophys. J.* 118, 96–104. doi: 10.1016/j.bpj.2019.11.014

Garrett, S. B., Garrison-Schilling, K. L., Cooke, J. T., and Pettis, G. S. (2016). Capsular polysaccharide production and serum survival of *Vibrio vulnificus* are dependent on antitermination control by RfaH. *FEBS Lett.* 590, 4564–4572. doi: 10.1002/1873-3468.12490

Gc, J. B., Bhandari, Y. R., Gerstman, B. S., and Chapagain, P. P. (2014). Molecular dynamics investigations of the α-Helix to β-Barrel conformational transformation in the RfaH transcription factor. *J. Phys. Chem. B* 118, 5101–5108. doi: 10.1021/jp502193v

Gc, J. B., Gerstman, B. S., and Chapagain, P. P. (2015). The role of the interdomain interactions on RfaH dynamics and conformational transformation. *J. Phys. Chem. B* 119, 12750–12759. doi: 10.1021/acs.jpcb.5b05681

Gerlach, R. G., Jäckel, D., Stecher, B., Wagner, C., Lupas, A., Hardt, W. D., et al. (2007). *Salmonella* pathogenicity Island 4 encodes a giant non-fimbrial adhesin and the cognate type 1 secretion system. *Cell. Microbiol.* 9, 1834–1850. doi: 10.1111/j.1462-5822.2007.00919.x

Goodson, J. R., Klupt, S., Zhang, C., Straight, P., and Winkler, W. C. (2017). LoaP is a broadly conserved antiterminator protein that regulates antibiotic gene clusters in *Bacillus amyloliquefaciens*. *Nat. Microbiol.* 2:17003. doi: 10.1038/nmicrobiol.2017.3

Gowrishankar, J., and Harinarayanan, R. (2004). Why is transcription coupled to translation in bacteria? *Mol. Microbiol.* 54, 598–603. doi: 10.1111/j.1365-2958.2004.04289.x

Gowrishankar, J., Leela, J. K., and Anupama, K. (2013). R-loops in bacterial transcription: their causes and consequences. *Transcription* 4, 153–157. doi: 10.4161/trns.25101

Grohmann, D., Nagy, J., Chakraborty, A., Klose, D., Fielden, D., Ebright, R. H., et al. (2011). The initiation factor TFE and the elongation factor Spt4/5 compete for the RNAP clamp during transcription initiation and elongation. *Mol. Cell* 43, 263–274. doi: 10.1016/j.molcel.2011.05.030

Heng, S.-P., Letchumanan, V., Deng, C.-Y., Ab Mutalib, N.-S., Khan, T. M., Chuah, L.-H., et al. (2017). *Vibrio vulnificus*: an environmental and clinical burden. *Front. Microbiol.* 8:997. doi: 10.3389/fmicb.2017.00997

Herbert, K. M., Zhou, J., Mooney, R. A., Porta, A. L., Landick, R., and Block, S. M. (2010). *E. coli* NusG inhibits backtracking and accelerates pause-free transcription by promoting forward translocation of RNA polymerase. *J. Mol. Biol.* 399, 17–30. doi: 10.1016/j.jmb.2010.03.051

Hirtreiter, A., Damsma, G. E., Cheung, A. C., Klose, D., Grohmann, D., Vojnic, E., et al. (2010). Spt4/5 stimulates transcription elongation through the RNA polymerase clamp coiled-coil motif. *Nucleic Acids Res.* 38, 4040–4051. doi: 10.1093/nar/gkq135

Hoffman, J. M., Sullivan, S., Wu, E., Wilson, E., and Erickson, D. L. (2017). Differential impact of lipopolysaccharide defects caused by loss of RfaH in *Yersinia pseudotuberculosis* and *Yersinia pestis*. *Sci. Rep.* 7:10915.

Huang, Y. H., Hilal, T., Loll, B., Bürger, J., Mielke, T., Böttcher, C., et al. (2020). Structure-based mechanisms of a molecular RNA polymerase/chaperone machine required for ribosome biosynthesis. *Mol. Cell* 79, 1024–1036.e5. doi: 10.1016/j.molcel.2020.08.010

Ingham, C. J., Dennis, J., and Furneaux, P. A. (1999). Autogenous regulation of transcription termination factor Rho and the requirement for Nus factors in *Bacillus subtilis*. *Mol. Microbiol.* 31, 651–663. doi: 10.1046/j.1365-2958.1999.01205.x

Iost, I., and Dreyfus, M. (1995). The stability of *Escherichia coli* lacZ mRNA depends upon the simultaneity of its synthesis and translation. *EMBO J.* 14, 3252–3261. doi: 10.1002/j.1460-2075.1995.tb07328.x

Johnson, G. E., Lalanne, J. B., Peters, M. L., and Li, G. W. (2020). Functionally uncoupled transcription-translation in *Bacillus subtilis*. *Nature* 585, 124–128. doi: 10.1038/s41586-020-2638-5

Joseph, J. A., Chakraborty, D., and Wales, D. J. (2019). Energy landscape for fold-switching in regulatory protein RfaH. *J. Chem. Theory Comput.* 15, 731–742. doi: 10.1021/acs.jctc.8b00912

Kang, J. Y., Mishanina, T. V., Landick, R., and Darst, S. A. (2019). Mechanisms of transcriptional pausing in Bacteria. *J. Mol. Biol.* 431, 4007–4029. doi: 10.1016/j.jmb.2019.07.017

Kang, J. Y., Mooney, R. A., Nedialkov, Y., Saba, J., Mishanina, T. V., Artsimovitch, I., et al. (2018). Structural basis for transcript elongation control by NusG family universal regulators. *Cell* 173, 1650–1662.e14. doi: 10.1016/j.cell.2018.05.017

Klein, B. J., Bose, D., Baker, K. J., Yusoff, Z. M., Zhang, X., and Murakami, K. S. (2011). RNA polymerase and transcription elongation factor Spt4/5 complex structure. *Proc. Natl. Acad. Sci. U.S.A.* 108, 546–550. doi: 10.1073/pnas.1013828108

Knodler, L. A., and Elfenbein, J. R. (2019). Salmonella enterica. *Trends Microbiol.* 27, 964–965. doi: 10.1016/j.tim.2019.05.002

Kong, Q., Yang, J., Liu, Q., Alamuri, P., Roland, K. L., Curtiss, R. III, et al. (2011). Effect of deletion of genes involved in lipopolysaccharide core and O-antigen synthesis on virulence and immunogenicity of *Salmonella enterica* serovar Typhimurium. *Infect. Immun.* 79, 4227–4239. doi: 10.1128/iai.05398-11

Krupp, F., Said, N., Huang, Y.-H., Loll, B., Bürger, J., Mielke, T., et al. (2019). Structural basis for the action of an all-purpose transcription anti-termination factor. *Mol. Cell* 74, 143–157.e5. doi: 10.1016/j.molcel.2019.01.016

Kyrpides, N. C., Woese, C. R., and Ouzounis, C. A. (1996). KOW: a novel motif linking a bacterial transcription factor with ribosomal proteins. *Trends Biochem. Sci.* 21, 425–426. doi: 10.1016/s0968-0004(96)30036-4

Landick, R., Carey, J., and Yanofsky, C. (1985). Translation activates the paused transcription complex and restores transcription of the trp operon leader region. *Proc. Natl. Acad. Sci. U.S.A.* 82, 4663–4667. doi: 10.1073/pnas.82.14.4663

Lane, W. J., and Darst, S. A. (2010a). Molecular evolution of multisubunit RNA polymerases: sequence analysis. *J. Mol. Biol.* 395, 671–685. doi: 10.1016/j.jmb.2009.10.062

Lane, W. J., and Darst, S. A. (2010b). Molecular evolution of multisubunit RNA polymerases: structural analysis. *J. Mol. Biol.* 395, 686–704. doi: 10.1016/j.jmb.2009.10.063

Larson, M. H., Mooney, R. A., Peters, J. M., Windgassen, T., Nayak, D., Gross, C. A., et al. (2014). A pause sequence enriched at translation start sites drives transcription dynamics in vivo. *Science* 344, 1042–1047. doi: 10.1126/science.1251871

Lawson, M. R., and Berger, J. M. (2019). Tuning the sequence specificity of a transcription terminator. *Curr. Genet.* 65, 729–733. doi: 10.1007/s00294-019-00939-1

Lawson, M. R., Ma, W., Bellecourt, M. J., Artsimovitch, I., Martin, A., Landick, R., et al. (2018). Mechanism for the regulated control of bacterial transcription termination by a universal adaptor protein. *Mol. Cell* 71, 911–922.e4. doi: 10.1016/j.molcel.2018.07.014

Leela, J. K., Syeda, A. H., Anupama, K., and Gowrishankar, J. (2013). Rho-dependent transcription termination is essential to prevent excessive genome-wide R-loops in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 110, 258–263. doi: 10.1073/pnas.1213123110

Lindberg, A. A., and Hellerqvist, C. G. (1980). Rough mutants of *Salmonella typhimurium*: immunochemical and structural analysis of lipopolysaccharides from rfaH mutants. *J. Gen. Microbiol.* 116, 25–32. doi: 10.1099/00221287-116-1-25

Liu, G., Olsen, J. E., and Thomsen, L. E. (2019). Identification of genes essential for antibiotic-induced up-regulation of plasmid-transfer-genes in cephalosporin resistant *Escherichia coli*. *Front. Microbiol.* 10:2203. doi: 10.3389/fmicb.2019.02203

Liu, Q., Liu, Q., Yi, J., Liang, K., Liu, T., Roland, K. L., et al. (2016). Outer membrane vesicles derived from *Salmonella typhimurium* mutants with truncated LPS induce cross-protective immune responses against infection of *Salmonella enterica* serovars in the mouse model. *Int. J. Med. Microbiol.* 306, 697–706. doi: 10.1016/j.ijmm.2016.08.004

Main-Hester, K. L., Colpitts, K. M., Thomas, G. A., Fang, F. C., and Libby, S. J. (2008). Coordinate regulation of *Salmonella* pathogenicity island 1 (SPI1) and SPI4 in *Salmonella enterica* serovar Typhimurium. *Infect. Immun.* 76, 1024–1035. doi: 10.1128/iai.01224-07

Martinez-Rucobo, F. W., Sainsbury, S., Cheung, A. C., and Cramer, P. (2011). Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity. *EMBO J.* 30, 1302–1310. doi: 10.1038/emboj.2011.64

Mason, S. W., and Greenblatt, J. (1991). Assembly of transcription elongation complexes containing the N protein of phage lambda and the *Escherichia coli* elongation factors NusA, NusB, NusG, and S10. *Genes Dev.* 5, 1504–1512. doi: 10.1101/gad.5.8.1504

McGary, K., and Nudler, E. (2013). RNA polymerase and the ribosome: the close relationship. *Curr. Opin. Microbiol.* 16, 112–117. doi: 10.1016/j.mib.2013.01.010

Miller, O. L., Hamkalo, B. A., and Thomas, C. A. (1970). Visualization of bacterial genes in action. *Science* 169, 392–395. doi: 10.1126/science.169.3943.392

Mitra, P., Ghosh, G., Hafeezunnisa, M., and Sen, R. (2017). Rho protein: roles and mechanisms. *Annu. Rev. Microbiol.* 71, 687–709. doi: 10.1146/annurev-micro-030117-020432

Mondal, S., Yakhnin, A. V., Sebastian, A., Albert, I., and Babitzke, P. (2016). NusA-dependent transcription termination prevents misregulation of global gene expression. *Nat. Microbiol.* 1:15007. doi: 10.1038/nmicrobiol.2015.7

Mooney, R. A., Davis, S. E., Peters, J. M., Rowland, J. L., Ansari, A. Z., and Landick, R. (2009a). Regulator trafficking on bacterial transcription units in vivo. *Mol. Cell* 33, 97–108. doi: 10.1016/j.molcel.2008.12.021

Mooney, R. A., Schweimer, K., Rosch, P., Gottesman, M., and Landick, R. (2009b). Two structurally independent domains of *E. coli* NusG create regulatory plasticity via distinct interactions with RNA polymerase and regulators. *J. Mol. Biol.* 391, 341–358. doi: 10.1016/j.jmb.2009.05.078

Moreno-Hagelsieb, G., and Santoyo, G. (2015). Predicting functional interactions among genes in prokaryotes by genomic context. *Adv. Exp. Med. Biol.* 883, 97–106. doi: 10.1007/978-3-319-23603-2_5

Nagy, G., Danino, V., Dobrindt, U., Pallen, M., Chaudhuri, R., Emödy, L., et al. (2006). Down-regulation of key virulence factors makes the *Salmonella enterica* serovar Typhimurium rfaH mutant a promising live-attenuated vaccine candidate. *Infect. Immun.* 74, 5914–5925. doi: 10.1128/iai.00619-06

Nagy, G., Dobrindt, U., Grozdanov, L., Hacker, J., and Emödy, L. (2005). Transcriptional regulation through RfaH contributes to intestinal colonization by *Escherichia coli*. *FEMS Microbiol. Lett.* 244, 173–180. doi: 10.1016/j.femsle.2005.01.038

Nagy, G., Dobrindt, U., Schneider, G., Khan, A. S., Hacker, J., and Emödy, L. (2002). Loss of regulatory protein RfaH attenuates virulence of uropathogenic *Escherichia coli*. *Infect. Immun.* 70, 4406–4413. doi: 10.1128/iai.70.8.4406-4413.2002

Nedialkov, Y., Svetlov, D., Belogurov, G. A., and Artsimovitch, I. (2018). Locking the nontemplate DNA to control transcription. *Mol. Microbiol.* 109, 445–457. doi: 10.1111/mmi.13983

Nicolas, P., Mader, U., Dervyn, E., Rochat, T., Leduc, A., Pigeonneau, N., et al. (2012). Condition-dependent transcriptome reveals high-level regulatory architecture in *Bacillus subtilis*. *Science* 335, 1103–1106. doi: 10.1126/science.1206848

Nudler, E. (2012). RNA polymerase backtracking in gene regulation and genome instability. *Cell* 149, 1438–1445. doi: 10.1016/j.cell.2012.06.003

Núñez, B., Avila, P., and De La Cruz, F. (1997). Genes involved in conjugative DNA processing of plasmid R6K. *Mol. Microbiol.* 24, 1157–1168. doi: 10.1046/j.1365-2958.1997.4111778.x

O'Reilly, F. J., Xue, L., Graziadei, A., Sinn, L., Lenz, S., Tegunov, D., et al. (2020). In-cell architecture of an actively transcribing-translating expressome. *Science* 369, 554–557. doi: 10.1126/science.abb3758

Paitan, Y., Orr, E., Ron, E. Z., and Rosenberg, E. (1999). A NusG-like transcription anti-terminator is involved in the biosynthesis of the polyketide antibiotic TA of *Myxococcus xanthus*. *FEMS Microbiol. Lett.* 170, 221–227. doi: 10.1111/j.1574-6968.1999.tb13377.x

Pan, T., Artsimovitch, I., Fang, X. W., Landick, R., and Sosnick, T. R. (1999). Folding of a large ribozyme during transcription and the effect of the elongation

factor NusA. *Proc. Natl. Acad. Sci. U.S.A.* 96, 9545–9550. doi: 10.1073/pnas.96.17.9545

Perdrizet, G. A. II, Artsimovitch, I., Furman, R., Sosnick, T. R., and Pan, T. (2012). Transcriptional pausing coordinates folding of the aptamer domain and the expression platform of a riboswitch. *Proc. Natl. Acad. Sci. U.S.A.* 109, 3323–3328. doi: 10.1073/pnas.1113086109

Peters, J. M., Mooney, R. A., Grass, J. A., Jessen, E. D., Tran, F., and Landick, R. (2012). Rho and NusG suppress pervasive antisense transcription in *Escherichia coli*. *Genes Dev.* 26, 2621–2633. doi: 10.1101/gad.196741.112

Pfalz, J., Liere, K., Kandlbinder, A., Dietz, K.-J., and Oelmüller, R. (2006). pTAC2, -6, and -12 are components of the transcriptionally active plastid chromosome that are required for plastid gene expression. *Plant Cell* 18, 176–197. doi: 10.1105/tpc.105.036392

Ponting, C. P. (2002). Novel domains and orthologues of eukaryotic transcription elongation factors. *Nucleic Acids Res.* 30, 3643–3652. doi: 10.1093/nar/gkf498

Proshkin, S., Rahmouni, A. R., Mironov, A., and Nudler, E. (2010). Cooperation between translating ribosomes and RNA polymerase in transcription elongation. *Science* 328, 504–508. doi: 10.1126/science.1184939

Ramírez-Sarmiento, C. A., Noel, J. K., Valenzuela, S. L., and Artsimovitch, I. (2015). Interdomain contacts control native state switching of RfaH on a dual-funneled landscape. *PLoS Comput. Biol.* 11:e1004379. doi: 10.1371/journal.pcbi.1004379

Rees, W. A., Weitzel, S. E., Yager, T. D., Das, A., and von Hippel, P. H. (1996). Bacteriophage lambda N protein alone can induce transcription antitermination in vitro. *Proc. Natl. Acad. Sci. U.S.A.* 93, 342–346. doi: 10.1073/pnas.93.1.342

Richardson, J. P. (1991). Preventing the synthesis of unused transcripts by Rho factor. *Cell* 64, 1047–1049. doi: 10.1016/0092-8674(91)90257-y

Rodnina, M. V. (2018). Translation in prokaryotes. *Cold Spring Harb. Perspect. Biol.* 10:a032664. doi: 10.1101/cshperspect.a032664

Roland, K. L., Liu, C. G., and Turnbough, C. L. Jr. (1988). Role of the ribosome in suppressing transcriptional termination at the pyrBI attenuator of *Escherichia coli* K-12. *Proc. Natl. Acad. Sci. U.S.A.* 85, 7149–7153. doi: 10.1073/pnas.85.19.7149

Said, N., Hilal, T., Sunday, N. D., Khatri, A., Bürger, J., Mielke, T., et al. (2020). Steps toward translocation-independent RNA polymerase inactivation by terminator ATPase ρ. *Science* doi: 10.1126/science.abd1673 [Epub ahead of print].

Santangelo, T. J., and Artsimovitch, I. (2011). Termination and antitermination: RNA polymerase runs a stop sign. *Nat. Rev. Microbiol.* 9, 319–329. doi: 10.1038/nrmicro2560

Saxena, S., and Gowrishankar, J. (2011). Compromised factor-dependent transcription termination in a nusA mutant of *Escherichia coli*: spectrum of termination efficiencies generated by perturbations of Rho, NusG, NusA, and H-NS family proteins. *J. Bacteriol.* 193, 3842–3850. doi: 10.1128/jb.00221-11

Saxena, S., Myka, K. K., Washburn, R., Costantino, N., Court, D. L., and Gottesman, M. E. (2018). *Escherichia coli* transcription factor NusG binds to 70S ribosomes. *Mol. Microbiol.* 108, 495–504. doi: 10.1111/mmi.13953

Schmidt, A., Kochanowski, K., Vedelaar, S., Ahrné, E., Volkmer, B., Callipo, L., et al. (2016). The quantitative and condition-dependent *Escherichia coli* proteome. *Nat. Biotechnol.* 34, 104–110. doi: 10.1038/nbt.3418

Schmidt, M. C., and Chamberlin, M. J. (1984). Binding of rho factor to *Escherichia coli* RNA polymerase mediated by nusA protein. *J. Biol. Chem.* 259, 15000–15002.

Seifi, B., Aina, A., and Wallin, S. (2020). Structural fluctuations and mechanical stabilities of the metamorphic protein RfaH. *Proteins* doi: 10.1002/prot.26014 [Epub ahead of print].

Sevostyanova, A., and Artsimovitch, I. (2010). Functional analysis of *Thermus thermophilus* transcription factor NusG. *Nucleic Acids Res.* 38, 7432–7445. doi: 10.1093/nar/gkq623

Sevostyanova, A., Belogurov, G. A., Mooney, R. A., Landick, R., and Artsimovitch, I. (2011). The β subunit gate loop is required for RNA polymerase modification by RfaH and NusG. *Mol. Cell* 43, 253–262. doi: 10.1016/j.molcel.2011.05.026

Sevostyanova, A., Svetlov, V., Vassylyev, D. G., and Artsimovitch, I. (2008). The elongation factor RfaH and the initiation factor sigma bind to the same site on the transcription elongation complex. *Proc. Natl. Acad. Sci. U.S.A.* 105, 865–870. doi: 10.1073/pnas.0708432105

Shi, D., Svetlov, D., Abagyan, R., and Artsimovitch, I. (2017). Flipping states: a few key residues decide the winning conformation of the only universally conserved transcription factor. *Nucleic Acids Res.* 45, 8835–8843. doi: 10.1093/nar/gkx523

Short, F. L., Di Sario, G., Reichmann, N. T., Kleanthous, C., Parkhill, J., and Taylor, P. W. (2020). Genomic profiling reveals distinct routes to complement resistance in *Klebsiella pneumoniae*. *Infect. Immun.* 88: e00043-20.

Singh, N., Bubunenko, M., Smith, C., Abbott, D. M., Stringer, A. M., Shi, R., et al. (2016). SuhB associates with Nus factors to facilitate 30S ribosome biogenesis in *Escherichia coli*. *mBio* 7:e00114-16.

Skals, M., Jorgensen, N. R., Leipziger, J., and Praetorius, H. A. (2009). α-Hemolysin from *Escherichia coli* uses endogenous amplification through P2X receptor activation to induce hemolysis. *Proc. Natl. Acad. Sci. U.S.A.* 106, 4030–4035. doi: 10.1073/pnas.0807044106

Sosunova, E., Sosunov, V., Kozlov, M., Nikiforov, V., Goldfarb, A., and Mustaev, A. (2003). Donation of catalytic residues to RNA polymerase active center by transcription factor Gre. *Proc. Natl. Acad. Sci. U.S.A.* 100, 15469–15474. doi: 10.1073/pnas.2536698100

Squires, C. L., Greenblatt, J., Li, J., Condon, C., and Squires, C. L. (1993). Ribosomal RNA antitermination in vitro: requirement for Nus factors and one or more unidentified cellular components. *Proc. Natl. Acad. Sci. U.S.A.* 90, 970–974. doi: 10.1073/pnas.90.3.970

Stevenson-Jones, F., Woodgate, J., Castro-Roa, D., and Zenkin, N. (2020). Ribosome reactivates transcription by physically pushing RNA polymerase out of transcription arrest. *Proc. Natl. Acad. Sci. U.S.A.* 117, 8462–8467. doi: 10.1073/pnas.1919985117

Sullivan, S. L., and Gottesman, M. E. (1992). Requirement for *E. coli* NusG protein in factor-dependent transcription termination. *Cell* 68, 989–994. doi: 10.1016/0092-8674(92)90041-a

Svetlov, D., Shi, D., Twentyman, J., Nedialkov, Y., Rosen, D. A., Abagyan, R., et al. (2018). In silico discovery of small molecules that inhibit RfaH recruitment to RNA polymerase. *Mol. Microbiol.* 110, 128–142. doi: 10.1111/mmi.14093

Svetlov, V., Belogurov, G. A., Shabrova, E., Vassylyev, D. G., and Artsimovitch, I. (2007). Allosteric control of the RNA polymerase by the elongation factor RfaH. *Nucleic Acids Res.* 35, 5694–5705. doi: 10.1093/nar/gkm600

Thomas, S., Holland, I. B., and Schmitt, L. (2014). The type 1 secretion pathway — the hemolysin system and beyond. *Biochim. Biophys. Acta* 1843, 1629–1641. doi: 10.1016/j.bbamcr.2013.09.017

Tomar, S. K., Knauer, S. H., Nandymazumdar, M., Rosch, P., and Artsimovitch, I. (2013). Interdomain contacts control folding of transcription factor RfaH. *Nucleic Acids Res.* 41, 10077–10085. doi: 10.1093/nar/gkt779

Turnbough, C. L. Jr. (2019). Regulation of bacterial gene expression by transcription attenuation. *Microbiol. Mol. Biol. Rev.* 83:e00019-19.

Turtola, M., and Belogurov, G. A. (2016). NusG inhibits RNA polymerase backtracking by stabilizing the minimal transcription bubble. *eLife* 5:e18096. doi: 10.7554/eLife.18096

Vos, S. M., Farnung, L., Urlaub, H., and Cramer, P. (2018). Structure of paused transcription complex Pol II-DSIF-NELF. *Nature* 560, 601–606. doi: 10.1038/s41586-018-0442-2

Vvedenskaya, I. O., Vahedian-Movahed, H., Bird, J. G., Knoblauch, J. G., Goldman, S. R., Zhang, Y., et al. (2014). Interactions between RNA polymerase and the "core recognition element" counteract pausing. *Science* 344, 1285–1289. doi: 10.1126/science.1253458

Wang, B., Gumerov, V. M., Andrianova, E. P., Zhulin, I. B., and Artsimovitch, I. (2020). Origins and molecular evolution of the NusG paralog RfaH. *mBio* 11:e02717-20.

Wang, C., Molodtsov, V., Firlar, E., Kaelber, J. T., Blaha, G., Su, M., et al. (2020). Structural basis of transcription-translation coupling. *Science* 369, 1359–1365. doi: 10.1126/science.abb5317

Wang, Y., Tian, G. B., Zhang, R., Shen, Y., Tyrrell, J. M., Huang, X., et al. (2017). Prevalence, risk factors, outcomes, and molecular epidemiology of mcr-1-positive *Enterobacteriaceae* in patients and healthy adults from China: an epidemiological and clinical study. *Lancet Infect. Dis.* 17, 390–399. doi: 10.1016/s1473-3099(16)30527-8

Washburn, R. S., Zuber, P. K., Sun, M., Hashem, Y., Shen, B., Li, W., et al. (2020). *Escherichia coli* NusG links the lead ribosome with the transcription elongation complex. *iScience* 23:101352. doi: 10.1016/j.isci.2020.101352

Webster, M. W., Takacs, M., Zhu, C., Vidmar, V., Eduljee, A., Abdelkareem, M., et al. (2020). Structural basis of transcription-translation coupling and collision in bacteria. *Science* 369, 1355–1359. doi: 10.1126/science.abb5036

Werner, F. (2012). A nexus for gene expression-molecular mechanisms of Spt5 and NusG in the three domains of life. *J. Mol. Biol.* 417, 13–27. doi: 10.1016/j.jmb.2012.01.031

Xiong, L., and Liu, Z. (2015). Molecular dynamics study on folding and allostery in RfaH. *Proteins* 83, 1582–1592. doi: 10.1002/prot.24839

Xun, S., Jiang, F., and Wu, Y. D. (2016). Intrinsically disordered regions stabilize the helical form of the C-terminal domain of RfaH: a molecular dynamics study. *Bioorg. Med. Chem.* 24, 4970–4977. doi: 10.1016/j.bmc.2016.08.012

Yakhnin, A. V., FitzGerald, P. C., McIntosh, C., Yakhnin, H., Kireeva, M., Turek-Herman, J., et al. (2020a). NusG controls transcription pausing and RNA polymerase translocation throughout the *Bacillus subtilis* genome. *Proc. Natl. Acad. Sci. U.S.A.* 117, 21628–21636. doi: 10.1073/pnas.2006873117

Yakhnin, A. V., Kashlev, M., and Babitzke, P. (2020b). NusG-dependent RNA polymerase pausing is a frequent function of this universally conserved transcription elongation factor. *Crit. Rev. Biochem. Mol. Biol.* 55, 716–728. doi: 10.1080/10409238.2020.1828261

Yakhnin, A. V., Murakami, K. S., and Babitzke, P. (2016). NusG is a sequence-specific RNA polymerase pause factor that binds to the non-template DNA within the paused transcription bubble. *J. Biol. Chem.* 291, 5299–5308. doi: 10.1074/jbc.M115.704189

Yakhnin, H., Yakhnin, A. V., Mouery, B. L., Mandell, Z. F., Karbasiafshar, C., Kashlev, M., et al. (2019). NusG-dependent RNA polymerase pausing and tylosin-dependent ribosome stalling are required for tylosin resistance by inducing 23S rRNA methylation in *Bacillus subtilis*. *mBio* 10:e02665-19.

Yang, Z., Li, M., and Sun, Q. (2020). RHON1 co-transcriptionally resolves R-Loops for *Arabidopsis* chloroplast genome maintenance. *Cell Rep.* 30, 243–256.e5. doi: 10.1016/j.celrep.2019.12.007

Young, R. A., and Steitz, J. A. (1978). Complementary sequences 1700 nucleotides apart form a ribonuclease III cleavage site in *Escherichia coli* ribosomal precursor RNA. *Proc. Natl. Acad. Sci. U.S.A.* 75, 3593–3597. doi: 10.1073/pnas.75.8.3593

Zellars, M., and Squires, C. L. (1999). Antiterminator-dependent modulation of transcription elongation rates by NusB and NusG. *Mol. Microbiol.* 32, 1296–1304. doi: 10.1046/j.1365-2958.1999.01442.x

Zhang, Y., Feng, Y., Chatterjee, S., Tuske, S., Ho, M. X., Arnold, E., et al. (2012). Structural basis of transcription initiation. *Science* 338, 1076–1080. doi: 10.1126/science.1227786

Zuber, P. K., Artsimovitch, I., NandyMazumdar, M., Liu, Z., Nedialkov, Y., Schweimer, K., et al. (2018). The universally-conserved transcription factor RfaH is recruited to a hairpin structure of the non-template DNA strand. *eLife* 7:e36349. doi: 10.7554/eLife.36349

Zuber, P. K., Schweimer, K., Rosch, P., Artsimovitch, I., and Knauer, S. H. (2019). Reversible fold-switching controls the functional cycle of the antitermination factor RfaH. *Nat. Commun.* 10:702.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 15 January 2021
doi: 10.3389/fmicb.2020.607693

Modification of Transfer RNA Levels Affects Cyclin Aggregation and the Correct Duplication of Yeast Cells

Loreto Arias[1], Fabián Martínez[1], Daniela González[1], Rodrigo Flores-Ríos[1], Assaf Katz[1], Mario Tello[2], Sandra Moreira[1] and Omar Orellana[1]**

[1] Programa de Biología Celular y Molecular, Instituto de Ciencias Biomédicas, Facultad de Medicina, Universidad de Chile, Santiago, Chile, [2] Departamento de Biología, Facultad de Química y Biología, Universidad de Santiago de Chile, Santiago, Chile

OPEN ACCESS

Edited by:
Hari S. Misra, Bhabha Atomic Research Centre (BARC), India

Reviewed by:
Herve Seligmann, Karlsruhe Institut of Technology, Germany
Martin Kollmar, Max Planck Institute for Biophysical Chemistry, Germany

***Correspondence:**
Sandra Moreira
sandra.moreira.r@gmail.com
Omar Orellana
oorellan@med.uchile.cl

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *17 September 2020*
Accepted: *21 December 2020*
Published: *15 January 2021*

Citation:
Arias L, Martínez F, González D, Flores-Ríos R, Katz A, Tello M, Moreira S and Orellana O (2021) Modification of Transfer RNA Levels Affects Cyclin Aggregation and the Correct Duplication of Yeast Cells. Front. Microbiol. 11:607693. doi: 10.3389/fmicb.2020.607693

Codon usage bias (the preferential use of certain synonymous codons (optimal) over others is found at the organism level (intergenomic) within specific genomes (intragenomic) and even in certain genes. Whether it is the result of genetic drift due to GC/AT content and/or natural selection is a topic of intense debate. Preferential codons are mostly found in genes encoding highly-expressed proteins, while lowly-expressed proteins usually contain a high proportion of rare (lowly-represented) codons. While optimal codons are decoded by highly expressed tRNAs, rare codons are usually decoded by lowly-represented tRNAs. Whether rare codons play a role in controlling the expression of lowly- or temporarily-expressed proteins is an open question. In this work we approached this question using two strategies, either by replacing rare glycine codons with optimal counterparts in the gene that encodes the cell cycle protein Cdc13, or by overexpression the $tRNA^{Gly}$ that decodes rare codons from the fission yeast, *Schizosaccharomyces pombe*. While the replacement of synonymous codons severely affected cell growth, increasing tRNA levels affected the aggregation status of Cdc13 and cell division. These lead us to think that rare codons in lowly-expressed cyclin proteins are crucial for cell division, and that the overexpression of tRNA that decodes rare codons affects the expression of proteins containing these rare codons. These codons may be the result of the natural selection of codons in genes that encode lowly-expressed proteins.

Keywords: transfer RNA, cell cycle, cyclin, codon usage, protein aggregation

INTRODUCTION

Degeneracy or redundancy of the genetic code implies that more than one codon (2, 4 or 6 codons) exist for 18 of the 20 genetically encoded amino acids. Codons encoding the same amino acid are called synonymous. Despite considerable information demonstrating that the choice of one synonymous codon over another is not random (Bulmer, 1991; reviewed in Quax et al., 2015), the term "silent" codons is still in use, since codon exchange does not alter protein sequences. Each organism has a defined codon usage bias (CUB). CUB greatly varies among species and within the same genome.

Whether CUB is the result of mutational drift forced by the nucleotide composition of DNA and/or natural selection for translation efficiency or accuracy has been a matter of debate (Bulmer, 1991; Brandis and Hughes, 2016; Dilucca et al., 2018). The most frequent codons (optimal codons) in rapidly-growing unicellular organisms are usually decoded by highly-expressed tRNAs, and the less frequent codons (rare or non-optimal codons) are decoded by less-expressed tRNAs (Ikemura, 1985; Kanaya et al., 1999; Chaney and Clark, 2015; Quax et al., 2015). In these organisms, CUB may optimize or deoptimize translation (translation efficiency) for a group of related mRNAs, coordinating their expression. However, these interpretations are controversial since some researchers have established that codon usage is an important factor in protein expression in trypanosomatids (Jeacock et al., 2018), where protein concentration can be estimated from protein coding sequences, while in other studies it has been determined that translation elongation speed is independent of codon usage bias (Ingolia et al., 2011). Moreover, codon bias has been identified as a major factor in determining both mRNA (Presnyak et al., 2015) and protein levels (Zhou et al., 2016). However, the extent to which mRNA translation efficiency links codon bias to protein levels remains unclear. This issue has gained importance in recent years as it has been linked to many factors that affect gene expression, such as mRNA stability, protein levels, folding, and localization (Chaney and Clark, 2015; reviewed in Bali and Bebok, 2015; Buhr et al., 2016; Hanson and Coller, 2018). Among these factors, the effect of codon choice on codon reading speed is the least understood. Since protein folding is co-translational for many proteins, codon changes that affect codon reading speed may also alter the folding of encoded proteins, leading to abnormal protein functioning. Experimental evidence and genome-wide analyses suggest that regions between protein domains are enriched in non-optimal codons, while structured domains are mostly encoded by optimal codons (Zhou et al., 2015). Despite these data, it is difficult to predict the synonymous mutations that lead to abnormal protein expression, folding, or function and the response of cells to these defects.

Lowly expressed proteins have a tendency to use rare codons. Even though there is a correlation between gene expression and rare codon use in various species (Hiraoka et al., 2009; Ray et al., 2014), it is still not clear if the tendency to use non-optimal codons plays a role in protein expression levels or in protein folding (Supek, 2016).

Other aspects of genetic information may be altered by the replacement of synonymous codons. It is relevant to consider these potential effects since cell function can be altered. It has been proposed that codons are selected to prevent off-frame translation after ribosomal slippage (ambush hypothesis) (Seligmann and Pollock, 2004; Seligmann, 2010, 2012, 2019; Køížek and Køížek, 2012), although this idea has been disputed (Morgens et al., 2013; Chatenay et al., 2017). Additionally, circular codes, that is, codes within the genetic code, have been proposed as playing a role in ensuring translation accuracy (Arquès and Michel, 1996). Proposed as primordial codes, they are conserved in evolution, with implications in the interaction of mRNAs with ribosomes and tRNAs (Michel and Thompson, 2020). Codon bias may be implicated in the conservation of circular codes in organisms and replacement of synonymous codons may alter such codes.

Circadian-clock proteins of *Neurospora crassa* and a cyanobacterium are examples of how the replacement of non-optimal codons with optimal ones affects the function of encoded proteins (Xu et al., 2013; Zhou et al., 2013) and alters circadian rhythms. A close correlation between the translation of non-optimal codons and the level of decoding, tRNA has been revealed in the gene encoding the cystic fibrosis transmembrane conductance regulator (CFTR). The replacement of an optimal with a non-optimal codon affects the function of CFTR. This effect is compensated by the upregulation of tRNA decoding non-optimal codon (Kirchner et al., 2017). Based on the above mentioned, the classical role of tRNAs as adaptor molecules for the incorporation of amino acids in nascent proteins has been expanded by findings that demonstrate changes in global tRNA abundance in response to different cellular processes (Torrent et al., 2018; Yang et al., 2020).

Eukaryotic cell-cycle proteins (cyclins) appear to be enriched in non-optimal codons, with some differences that depend on the cell-cycle phase in which they are expressed. The levels of total tRNA and of some aminoacyl-tRNA synthetases are also cell-cycle dependent in yeast (Frenkel-Morgenstern et al., 2012). Cdc13 is one of the most widely studied cyclins from the fission yeast, *S. pombe*. It forms a complex and activates Cdc2 (Cdk1), an important kinase in G2/M transition (Humaidan et al., 2018). Cdc13 is preferentially expressed during this transition. The complex travels to the nucleus and phosphorylates a number of substrates that are crucial for the progress of the cell cycle. Although the protein structure of Cdc13 has not yet been determined, a secondary structure is predicted, and some relevant regions have been identified. Cdc13 has a conserved hydrophobic patch (MRGILTDW) that is not required for cells to undergo the S phase *in vivo* but is required to target Cdc13 to the equivalent to the spindle pole body in yeast (SPB) and for mitosis (Basu et al., 2020). Mutation in this hydrophobic patch alters Cdc13 localization, preventing centrosomal localization at the onset of mitosis. Finally, the complex is degraded at the mitosis stage. Since *cdc13* contains several non-optimal codons, we hypothesize that the presence of non-optimal codons in *cdc13* is required to modulate the proper level of the Cdc13 protein. The variable levels of tRNAs during the cell cycle (Frenkel-Morgenstern et al., 2012) may control the expression of *cdc13*. To test this hypothesis, we followed two approaches: (1) to introduce synonymous mutations to replace non-optimal with optimal codons in *cdc13*, and (2), to modify the concentration of tRNAs that decode non-optimal codons. Our results show that these alterations significantly affect both Cdc13 distribution in soluble and aggregated fractions and cell duplication.

MATERIALS AND METHODS

Yeast Strains and Media

Schizosaccharomyces pombe 972h-Sleu1-32 (LP36) strain was used in homologous recombination experiments.

Supplementary Table 1 lists the primers used in this study. *S. pombe* was grown on YES medium (5 g/l of yeast extract, 30 g/l of glucose), YPD (10 g/l of yeast extract, 20 g/l of peptone, 20 g/l of glucose), and Edinburgh minimal medium 2 (EMM2, United States Biological).

Plasmid Construction for Homologous *cdc13* Recombination

The endogenous *cdc13* gene was replaced by a homologous recombination with the wild-type or mutated *cdc13*, both containing a 7xHis-tag at the 3′ end. For the homologous recombination with the wild-type *cdc13* gene, we first cloned the flanking regions of the Cdc13 coding sequence in the pFA6a-KanMX6 vector (Addgene), as described below. Flanking regions were amplified from *S. pombe* genomic DNA (gDNA), using Herculase II Fusion DNA polymerase (Agilent Genomics) in accordance with the manufacturer's instructions, using cdc13-histag forward and reverse primer sets as described in **Supplementary Table 1**. The digestion products were ligated to the pFA6a-KanMX6 vector (first the 3′ flanking region, followed by the 5′ flanking region), previously digested with the corresponding restriction enzymes (*Sac*I and *Eco*RI to 3′ flanking region; *Nde*I and *Bam*HI to 5′ flanking region). The *E. coli* JM109 strain was transformed with the corresponding construction by chemical transformation (Sambrook and Russell, 2001; Froger and Hall, 2007) and five clones were analyzed by DNA sequencing to corroborate the correct incorporation of 7xHis-tag. The final vector was named pFA6a-KanMX6-5′-3′ and confirmed by DNA sequencing.

The synonymous substitutions were designed based on *S. pombe* codon usage (Forsburg, 1994; Hiraoka et al., 2009) replacing low-usage codons GGG/GGA (five codons) with their optimal counterpart GGT. **Supplementary Table 1** shows the primers that were used. *Cdc13* mutants were constructed according to the following methodology: first, the mutant *cdc13* sequence was amplified in five separate PCRs, using the following primers listed in **Supplementary Table 1**: cdc13-histag-F and CDC13_GGT1_R (reverse primer containing the first mutated codon) were used to amplify the first segment; CDC13_GGT1_F and CDC13_GGT2_R (reverse primer containing the second mutated codon) were used to amplify the second segment; CDC13_GGT2_F and CDC13_GGT3_R (reverse primer containing the third mutated codon) were used to amplify the third segment; CDC13_GGT3_F and CDC13_GGT45_R (reverse primer containing the last two mutated codons) were used to amplify the fourth segment; and CDC13_GGT45_F forward primer containing the last two mutated codons) and cdc13-histag-R were used to amplify the last segment. All PCR products were purified using the Real Genomics HiYieldTM Gel/PCR DNA fragments Extraction system commercial kit (Real Genomics) in accordance with the manufacturer's instructions, and then the purified PCR products were used as templates for four amplifications by PCR: segments I and II were used as templates to join both segments, amplifying with cdc13-histag-F and CDC13_GGT2_R primers. In another experiment, segments III and IV were used as templates to join both segments, amplifying with CDC13_GGT2_F and CDC13_GGT45_R primers (**Supplementary Table 1**). Both PCR products were purified the same way and were used as templates to join the four segments by amplified PCR with cdc13-histag-F and CDC13_GGT45R primers (**Supplementary Table 1**). This product was joined with fragment V as described previously, amplified with cdc13 histag-F and cdc13-histag-R primer sets (**Supplementary Table 1**), resulting in the complete *cdc13* with the mutations. The final product was purified and digested with *Nde*I and *Bam*HI restriction enzymes, and then ligated to the pFA6a-KanMX6-5′-3′ vector. The product was transformed into the *E. coli* JM109 strain by chemical transformation (Sambrook and Russell, 2001; Froger and Hall, 2007) and 5 clones were analyzed by DNA sequencing to corroborate the correct incorporation of the mutations.

Homologous Recombination

Purified plasmids containing the flanking regions of the wild-type plus mutant coding sequences of Cdc13 and the His-tag were used to amplify the sequence used for homologous recombination (Oldenburg et al., 1997), using Herculase II Fusion DNA polymerase (Agilent Genomics) in accordance with the manufacturer's instructions, and using pREP41F′ and pREP41R′ primers (**Supplementary Table 1**). 200 ng of the corresponding PCR product and 300 ng of the pET15b vector (as a carrier) were electroporated in 100 μl of electrocompetent *S. pombe* cells were prepared with the described protocol (Forsburg and Rhind, 2006). Yeasts were washed with 1 M sorbitol and recovered in YPD medium at 30°C for 3 h. The cells were then pelleted, resuspended in 200 μl of YE medium, and plated in YPD agar prepared with 200 μg/ml of G418 (geneticin) antibiotic (Sigma-Aldrich). Cells were grown for 4–5 days at 30°C and single colonies were picked and grown on YPD medium supplemented with 200 μg/ml of G418 antibiotic. The incorporation of synonymous mutations at the right position was corroborated by DNA sequencing.

Overexpression of tRNAs

Fragments containing tRNA genes ($tRNA^{Gly}_{UCC}$, $tRNA^{Gly}_{GCC}$, and $tRNA^{Arg}_{UCC}$) were amplified by PCR from *S. pombe* genomic DNA using Herculase II Fusion DNA polymerase (Agilent Genomics) in accordance with the manufacturer's instructions and using the primer set described in **Supplementary Table 1**. The PCR products contained 400 base pairs for $tRNA^{Gly}_{UCC}$ and 420 base pairs for $tRNA^{Gly}_{GCC}$ and $tRNA^{Arg}_{UCC}$, including the regulatory elements necessary for tRNA transcription. The products were purified using the commercial kit Real Genomics HiYieldTM Gel/PCR DNA Fragment Extraction (Real Genomics), in accordance with the manufacturer's instructions. The purified products were digested with *Bam*HI and *Nde*I and cloned into pREP41, a high copy number vector. Vectors with tRNA genes were transformed into yeast by electroporation, as described previously (Forsburg and Rhind, 2006).

Northern Blot Analyses

Northern blot analysis was performed using biotinylated probes synthesized by IDT Technologies (**Supplementary Table 2**).

Samples were transferred to positively charged nylon membranes for 2 h and 15 min at 20 volts in 0.5X TBE buffer (45 mM Tris-borate, 1 mM EDTA). RNA was then fixed by UV radiation (120,000 μjoules) and membranes were blocked for 30 min at 41°C in pre-hybridization solution (6X SSC (150 mM NaCl, 15 mM Sodium Citrate), 5X Denhardt's, 0.1% SDS, 100 μg/ml salmon sperm DNA). After blocking, probes were added directly to the hybridization solution (6X SSC, 0.1% SDS, 100 μg/ml salmon sperm DNA) and incubated overnight. Membranes were incubated for 3 min at room temperature with solution A (2X SSC, 0.1% SDS) and then incubated twice for 15 min at 41°C in solution B (0.1X SCC, 0.1% SDS). Next, membranes were blocked for 30 min at room temperature with blocking solution (1% casein in maleic buffer; 0.1 M maleic acid, 0.15 M NaCl pH 7.5). Then, 0.1 μg/ml of streptavidin- horseradish peroxidase was added to the blocking solution. Membranes were incubated for 30 min at room temperature and then cleaned twice for 15 min with maleic acid buffer 0.3% (v/v) tween-20. Finally, membranes were cleaned for 3 min in predetection buffer (0.1 M Tris–HCl, 0.1 M NaCl, pH 9.5) and developed using a chemiluminescent kit [SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific)].

Determination of tRNA Aminoacylation Levels *in vivo*

Total RNA (obtained from the different strains that overexpress $tRNA^{Gly}$ or $tRNA^{Arg}$) was purified under acidic conditions and the 3′ extreme nucleotide was then eliminated by sodium periodate oxidation followed by β-elimination (Salazar et al., 2001; Choi et al., 2003). For this purpose, yeast was grown in EMM2 at 30°C, until OD600 of 0.9–1.0, and then pelleted at 10,000 × *g* for 6 min at room temperature. The pellet was resuspended in 500 μl of 0.3 M sodium acetate at pH 5.2, 1 mM of EDTA, followed by the addition of 500 μl of acid phenol. The mix was incubated for 10 min on ice and then centrifuged for 6 min at 10,000 × *g*. The supernatant was recovered, and RNA was precipitated adding three volumes of 100% ethanol at −80°C overnight. Samples were centrifuged for 30 min at 14,000 × *g* at 4°C. Pellets were cleaned with 0.5 ml of 75% ethanol, 10 mM sodium acetate pH 5.2 and then resuspended in 50 μl of water. Each sample was divided into two tubes (A and B), each with 25 μl aliquots. 1.42 μl of 3 M sodium acetate at pH 5.2 was added to the A tubes, and then stored at −80°C. tRNA in the B tubes was deacylated by adding 6.25 μl of 1 M Tris acetate at pH 9.0 and incubated for 60 min at 37°C. Samples in the B tubes were precipitated by adding 3.13 μl of 3 M sodium acetate at pH 5.2 and 62.5 μl of ethanol at 100%, and stored for 30 min at −80°C. Samples were centrifuged for 30 min at 13,000 × *g* at 4°C, pellets were washed with 70% ethanol and centrifuged at 13,000 × *g* for 5 min at 4°C. The pellet was dried and resuspended in 26.4 μl of 160 mM sodium acetate pH 5.2. Samples from the A tubes were thawed and 4.8 μl of freshly prepared 250 mM sodium periodate was added to each sample. Tubes were covered in aluminum paper and incubated for 90 min on ice, and then 12.97 μl of 20% glucose was added. After an additional 90 min of incubation on ice, 4.3 μl of 3 M sodium acetate at pH 5.2 and 87 μl of ethanol were added. Samples were stored for at least 30 min at −80°C and centrifuged for 30 min at13,000 × *g* at 4°C. Pellets were resuspended in 250 μl of 0.5 M lysine at pH 8.0 and incubated for 60 min at 45°C. Then, 25 μl of 3 M sodium acetate at pH 5.2 and 500 μl of ethanol were added and samples were stored for at least 30 min at −80°C. Tubes were centrifuged again for 30 min at 13,000 × *g* at 4°C and then washed with 70% ethanol. Finally, pellets were dried and resuspended in 15 μl of RNAse-free water. Samples were analyzed by gel electrophoresis (8 M urea, 10% polyacrylamide), followed by Northern blot analysis.

mRNA Isolation and Quantification

To quantify *cdc13* mRNA, wild-type and over-expressed tRNA strains were grown to the late exponential phase (A600, 1.0) in 15 ml of EMM2 medium under standard conditions. Yeasts were then pelleted at 2,250 × *g* for 5 min at room temperature and washed once with sterilized water. Cell walls were disrupted with 0.12 μg of zymolyase 20T (United States Biological) in 1 M of sorbitol for 30 min at 37°C. Cells were pelleted and resuspended in 200 μl of TRIzol (Thermo Scientific) and vortexed three times for 1 min each (intercalated with 1 min on ice). Then, 40 μl of chloroform was added to the mix and immediately centrifuged at 12,000 × *g* for 20 min at 4°C. The supernatant was precipitated with 0.7 volumes of isopropanol at −80°C overnight. RNA was pelleted at 15,000 × *g* for 30 min and washed once with 80% ethanol. RNA was resuspended in 30 μl of sterilized miliQ water and quantified in a nanodrop spectrometer (BioTek), and then visualized in a 1% agarose gel to check RNA integrity. 1 μg of RNA was treated with DNAse I (Roche), according to the instructions provided in the manual. 500 ng of RNA was used for cDNA synthesis followed by Real-Time PCR with Brilliant II QRT-PCR, AffinityScript Two-Step Master Mix (Agilent), according to the manufacturer's instructions. The primers used in the Real-Time PCR are listed in **Supplementary Table 1** (cdc13rtF and cdc13rtR primers). Actin was used as a control in the quantifications.

Cell-Cycle Synchronization by Hydroxyurea

Hydroxyurea (HU) was used to synchronize cells in the early-mid S phase, as described previously (Luche and Forsburg, 2009). Cells were grown in 5 ml of EMM2 at 30°C overnight, with constant agitation. Subsequently, the saturated culture was brought to an OD600 of 0.1 and incubated for 3 h or until an OD600 ∼0.2–0.6 HU was added at a final concentration of 15 mM and incubated for 4 h at 30°C with constant agitation. After the time, cells were collected by centrifuging at 4,000 × *g* for 5 min at room temperature. The supernatant was discarded, and the resulting sediment was washed twice with half the volume used of EMM2, previously incubated at 30°C, to remove residual HU. The sediment was resuspended in 1 ml of sterile EMM2 and incubated at 30°C with constant agitation, during all the experiment time (24 h) and a 10 μl aliquot was taken every 3 h. The phenotype was analyzed observing the cells by optical microscopy. A 40X objective (total magnification: 400) was used.

The photos were obtained using a camera attached to a Canon SLR model Eos Rebel T3 lens microscopy.

Total and Aggregated Protein Isolation and Quantification

Wild-type and mutant strains were grown at the exponential phase (OD600, 1.0) in 20 ml of EMM2 medium at 30°C with constant agitation. Yeasts were harvested by centrifugation at 2,250 × *g* for 5 min at room temperature, and the pellet was used to analyze protein aggregation as described previously (Rand and Grant, 2006). Briefly, cells were pelleted and resuspended in lysis buffer [50 mM potassium phosphate buffer, pH 7, 1 mM EDTA, 5% glycerol, 1 mM of phenylmethylsulfonyl fluoride, and Complete Mini protease inhibitor cocktail (Roche)]. Cell disruption was carried out by three vortex cycles (1 min of vortex and 1 min on ice) with 220 mg of acid-washed glass beads (Sigma-Aldrich G8772). Membrane proteins were removed by washing twice with 320 μl lysis buffer and 80 μl of 10% NP-40 (Sigma-Aldrich), and the final aggregated protein extract was resuspended in 20 μl of 1X Laemmli sample buffer. Total and aggregated protein extracts were analyzed by Western blot using an anti His-tag antibody in a 1:1,000 dilution (His Tag Antibody MAB050R-100, R&D Systems). Western blot against tubulin was performed as an internal control (T5168 monoclonal anti-a-tubulin clone B-5-1-2, Sigma-Aldrich). Bands representing WT, the pREP41_tRNA strain, and tubulin (for total protein), or WT and pREP41_tRNA strain (for aggregated proteins) were quantified using ImageJ software.

Statistical Significance

The results are presented as mean ± standard deviations of the number of independent experiments indicated (biological replicates) (n) or as a representative example of experiments performed at least three times independently. Data were analyzed statistically using GraphPad Prism 6.0 software. The results were analyzed using the unpaired *t*-student test to determine significant differences among the experimental conditions. A p value < 0.05 was considered the limit of significance.

RESULTS

Replacement of Rare Glycine Codons for Synonymous Preferred Codons in *cdc13* Severely Affects *S. pombe* Growth

We used the gene encoding Cdc13, which is rich in non-optimal glycine codons to evaluate the effect of replacing non-optimal for optimal codons. GGA and GGG glycine codons are highly represented in *cdc13* (5 of the 13 glycine codons, 38.4% compared to less than 1% in highly-expressed genes) (**Supplementary Table 3**). Thus, we choose to replace these rare codons with the optimal GGT codon (86% of the glycine codons in highly-expressed genes) (**Figure 1** and **Supplementary Table 3**), as described above. Surprisingly, the replacement of GGA/GGG codons with GGT in *cdc13*, yielded two types of colonies, small (most of them) and large (the remaining few) (**Supplementary Figure 1**, upper panel). The small colonies did not grow on either solid orin liquid media. The large colonies were cultured in liquid medium, where they grew slowly. Cell shape was observed by light microscopy at 16 h of culturing in rich medium. The mutant cells presented a phenotype similar to that of wild-type strains, although some cells were elongated, similar to what has been observed in yeast where the cell cycle is blocked after the G2 phase (Oltra et al., 2004; **Supplementary Figure 1**, lower panel). We believe that the replacement of rare glycine codons by optimal synonymous codons produced an altered Cdc13 protein that is non-functional to the cells, consequently cells stopped proliferating after a few duplications. Inspection of the *cdc13* sequence around the replaced codons did not reveal any alteration of off-frame stop codons (ambush hypothesis) (Seligmann and Pollock, 2004) that might account for the growth defect. However, large colonies proliferate even though *cdc13* contains all the rare glycine codon replacements (confirmed by DNA sequencing). The effect of replacing *cdc13* with mutated genes should be investigated, as should the differences in growth between the large and small colonies.

tRNA$^{Gly}_{UCC}$ Overexpression Does Not Affect the Level of Other tRNA^{Gly}

As noted above, tRNA levels in *S. cerevisiae* vary during the cell cycle, with the maximum at G2-M transition (Frenkel-Morgenstern et al., 2012). Since the replacement of rare glycine codons by optimal synonymous codons had a dramatic effect on *S. pombe* proliferation, we thought that the translation of rare codons may be controlled by the concentration of the decoding tRNA. The rare glycine codons GGA and GGG in *S. pombe* are decoded by the tRNA$^{Gly}_{UCC}$, encoded by three almost identical genes (low copy number compared to other tRNA^{Gly} genes). The tRNA$^{Gly}_{CCC}$ encoded by a single copy of the gene also decodes GGG codons (**Supplementary Table 3**).

We determined whether the expression of tRNA$^{Gly}_{UCC}$ is regulated in wild-type cells (as in *S. cerevisiae*) (Pang et al., 2014) by treating the cells with hydrogen peroxide and measuring the level of this tRNA. *S. pombe* cells were exposed to 5 or 10 mM of hydrogen peroxide and catalase activity was measured to confirm the response of the cells to oxidative stress. Increased catalase activity was observed under both conditions (data not shown). At 10 mM of hydrogen peroxide, the level of tRNA$^{Gly}_{UCC}$ increased fivefold, indicating that the level of tRNA$^{Gly}_{UCC}$ is regulated in response to this external challenge (**Supplementary Figure 2A**).

To test whether an increase in tRNA$^{Gly}_{UCC}$ concentration affects the level of Cdc13, we overexpressed this tRNA in *S. pombe*. For this purpose, we cloned the corresponding tRNA gene (which contains an internal RNA polymerase III promoter) in the high copy number vector pREP41, with LEU2 as selection marker. *S. pombe* cells were transformed with this construct (**Supplementary Figure 2B**), with the empty vector as a control. Northern blot analysis using a specific probe for tRNA$^{Gly}_{UCC}$ revealed a 7-to-8-fold increase in the level of this tRNA compared to the control (**Figure 2A**). The overexpression of tRNA$^{Gly}_{UCC}$ did not affect the levels of other tRNA^{Gly} (tRNA$^{Gly}_{GCC}$, tRNA$^{Gly}_{CCC}$) (**Figure 2A**). We also tested the aminoacylation status of

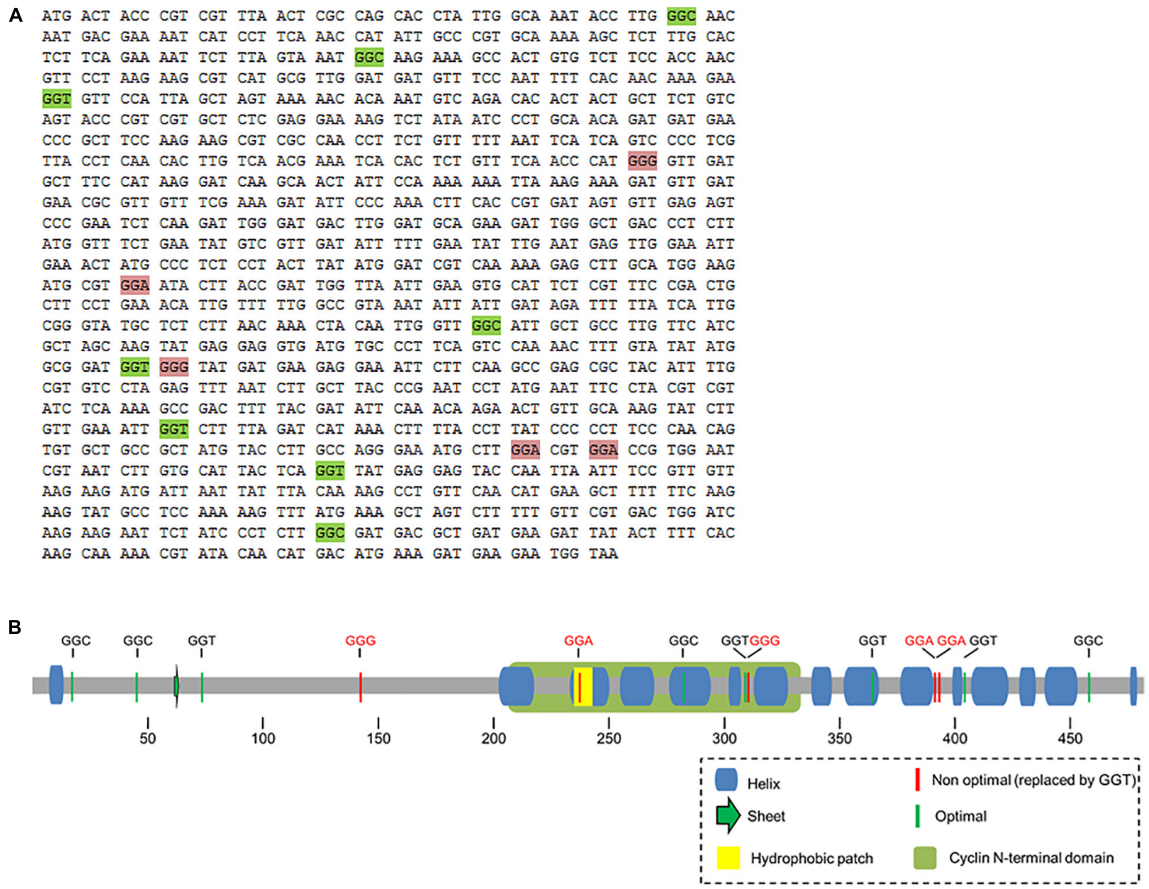


FIGURE 1 | (A) Sequence of *cdc13* and localization of Gly codons. The coding sequence of *cdc13* is shown and unreplaced Gly codons (green boxes) and replaced codons (pink boxes) are indicated. **(B)** Schematic representation of secondary structure of Cdc13 from *S. pombe.* Location of glycine codons in the secondary structure of Cdc13 and the corresponding optimal (green) and non-optimal (red) codons in mRNA. The yellow square represents the hydrophobic motif MRGILTDW and the green rectangle represents the cyclin domain.

these tRNA^{Gly} with overexpression of tRNA$^{Gly}{}_{UCC}$. The ratio of aminoacylated/deacylated tRNA in cells transformed with the empty vector and cells that overexpressed tRNA$^{Gly}{}_{UCC}$ was determined. We observed a ∼20% increase in the amino acylated fraction of tRNA$^{Gly}{}_{UCC}$ compared to the control (**Figures 2B,C**) (although it was not a statistically significant) The fact that the level of this tRNA increased 7-to-8-fold in cells transformed with the vector containing the tRNA$^{Gly}{}_{UCC}$ gene (**Figure 2A**) implies that the amount of Gly-tRNA^{Gly} increases 10-to-11-fold compared to the control. However, the overexpression of tRNA$^{Gly}{}_{UCC}$ did not significantly change the amino acylation levels of other tested tRNA^{Gly} isoacceptors (**Figures 2B,C**).

Overexpression of tRNA$^{Gly}{}_{UCC}$ Affects the Distribution of Cdc13 Into Soluble and Aggregated Fractions

We used cells transformed with pREP4-tRNA$^{Gly}{}_{UCC}$ and control cells to evaluate the effect of tRNA$^{Gly}{}_{UCC}$ overexpression on Cdc13 expression. Neither total Cdc13 protein levels (**Figures 3A–C**) nor its mRNA levels (**Figure 3D**) were altered by tRNA$^{Gly}{}_{UCC}$ overexpression. However, when the soluble and aggregated fractions were separated, a dramatic increase in Cdc13 was observed in the aggregated fraction (**Figure 3A**, quantified in **Figure 3B**). Together, these results suggest that increasing tRNA$^{Gly}{}_{UCC}$ levels increases the Cdc13 in the aggregated fraction without affecting its level. Further investigation is needed to determine if this is the result of Cdc13 misfolding because of a change in the mRNA translation rate. An alternative explanation for this defect is that other proteins (including chaperones) are altered by tRNA$^{Gly}{}_{UCC}$ overexpression and aggregate with Cdc13. Another explanation is that tRNA$^{Gly}{}_{UCC}$ overexpression causes increased mismatches with non-cognate codons that are not compensated (Seligmann, 2011), giving rise to translation errors that induce protein aggregations. These and other alternative explanation need to be investigated.

Overexpression of tRNA$^{Gly}{}_{UCC}$ Affects the Progression of the *S. pombe* Cell Cycle

To test whether the overexpression of tRNA^{Gly} affects the cell cycle progression of *S. pombe*, yeast cultures were synchronized with HU. The shapes of cells that overexpress tRNA$^{Gly}{}_{UCC}$

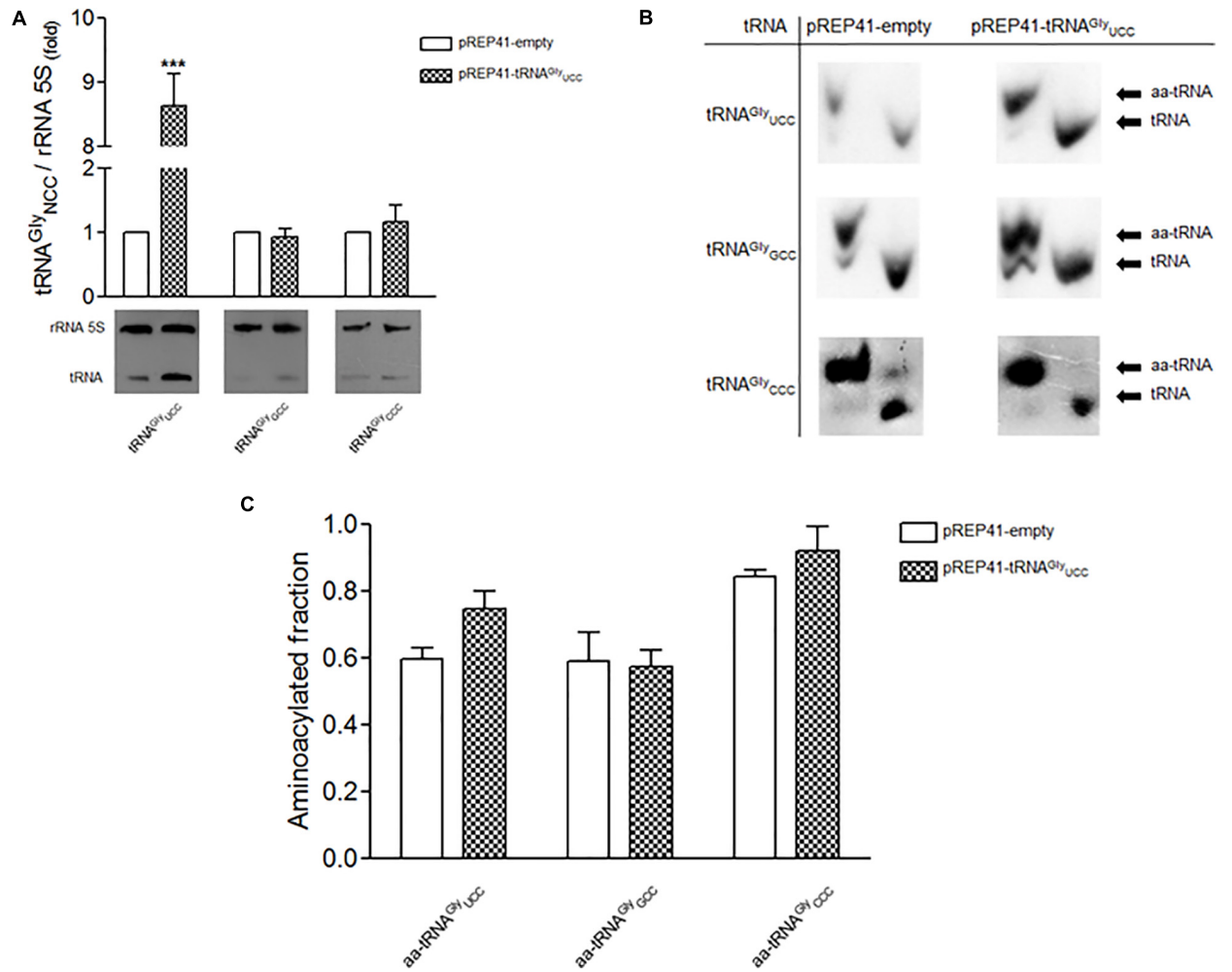


FIGURE 2 | Overexpression of tRNA$^{Gly}_{UCC}$ (low gene dosage) and the effect on the level and aminoacylation of tRNA^{Gly} isoacceptors. **(A)** Lower panel. Northern blot of tRNA$^{Gly}_{UCC}$, tRNA$^{Gly}_{GCC}$ and tRNA$^{Gly}_{CCC}$ after overexpression of tRNA$^{Gly}_{UCC}$. Upper panel, relative levels of tRNA^{Gly} isoacceptors in cells that overexpress tRNA$^{Gly}_{UCC}$. 5S rRNA was used as a reference (folds of change (***$p < 0{,}001$, *t*-student). **(B)** Aminoacylation of tRNA^{Gly} isoacceptors. Representative Northern blot analysis of total tRNA purified under acidic conditions from *S. pombe* transformed with pREP41-empty (control) or pREP41-tRNA$^{Gly}_{UCC}$, detected with specific probes against tRNA$^{Gly}_{UCC}$, tRNA$^{Gly}_{GCC}$, and tRNA$^{Gly}_{CCC}$. Aminoacylated tRNAs (aa-tRNA) and non-aminoacylated tRNAs (tRNA) are indicated. **(C)** Quantification of Northern blots. Relative levels of aminoacylated fraction of tRNA$^{Gly}_{UCC}$, tRNA$^{Gly}_{GCC}$, and tRNA$^{Gly}_{CCC}$ for each condition, were quantified by densitometric analysis.

and control cells were monitored under a microscope at 12 h (three cell cycles). Under these conditions, we observed that nearly 100% of cells that overexpress tRNA$^{Gly}_{UCC}$ had elongated shapes (**Figure 4A**). The average size of cells that overexpress tRNA$^{Gly}_{UCC}$ was at least twice that of the control cells, but some were four times as large (**Figure 4B**). More than one septum was observed microscopically in the elongated cells. Unsynchronized cells also evidenced elongated shape, but to a lesser extent.

In order to evaluate whether the elongated phenotype is a specific result of the overexpression of tRNA$^{Gly}_{UCC}$ and not the overexpression of any other tRNA^{Gly}, we overexpressed the high copy number tRNA$^{Gly}_{GCC}$ gene in *S. pombe* that decodes both optimal glycine codons. We observed a twofold increase in tRNA$^{Gly}_{GCC}$ (**Figure 5A**). As tRNA$^{Gly}_{GCC}$ is encoded by eight copies (**Supplementary Table 3**), we expected relatively high levels of this tRNA in the cells. However, *S. pombe* showed no elongated phenotypes (**Figure 5B**). Similar experiments involved overexpressing tRNA$^{Arg}_{UCU}$. This tRNA decodes the arginine AGA (preferred) and AGG (rare) codons present in *S. pombe cdc13*. The tRNA gene was cloned as indicated for tRNA$^{Gly}_{UCC}$. A 2.5-fold increase in the cellular level of tRNA was observed (data not shown). However, the overexpression of tRNA$^{Arg}_{UCU}$ did not alter cell shape, as tRNA$^{Gly}_{UCC}$ overexpression did (data not shown).

DISCUSSION

Translation efficiency and accuracy are largely achieved by binding to the ribosome of the proper aminoacyl-tRNA that

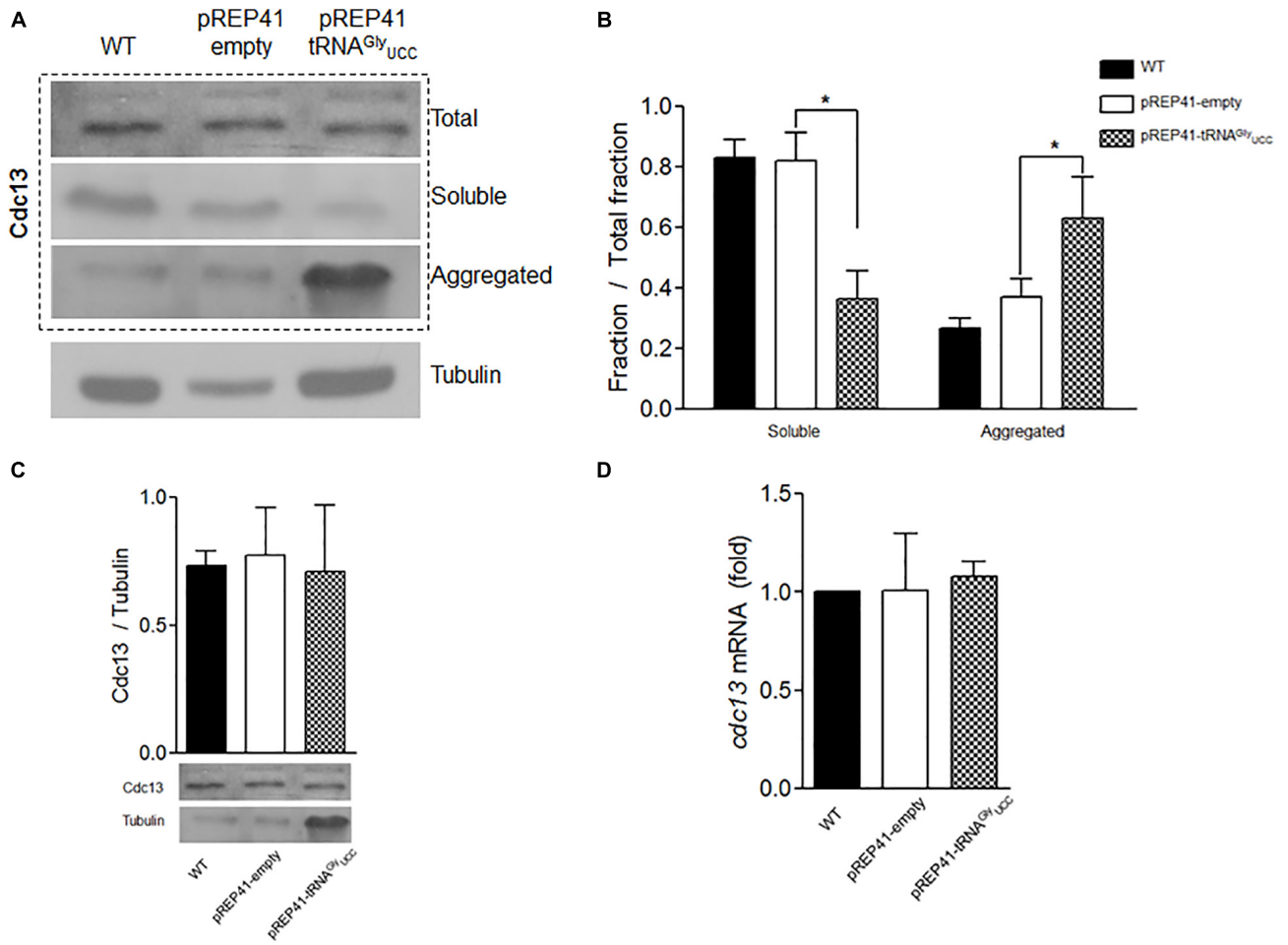


FIGURE 3 | Overexpression of tRNA$^{Gly}_{UCC}$ in *S. pombe* affects the aggregation status of Cdc13. **(A)** Western blot of Cdc13 in total extract or soluble and aggregated fractions. **(B)** Quantification of soluble and aggregated Cdc13 compared to total extract. **(C)** Total Cdc13 in cells that overexpress tRNA$^{Gly}_{UCC}$ compared to controls. **(D)** Relative levels of *cdc13* mRNA compared to those of controls (pREP41-empty). (*$p < 0.05$, *t*-student).

competes with a plethora of non-cognate or near-cognate aminoacyl-tRNAs (all at different concentrations) to enter the A site, Ikemura (1985), Kanaya et al. (1999), Frenkel-Morgenstern et al. (2012), Kirchner et al. (2017), Torrent et al. (2018), Yang et al. (2020). This gives rise to the notion that optimal codons are translated by highly represented transfer RNAs. The translation speed of certain codons has been explained as the result of the time required by the ribosome to find the proper aminoacyl-tRNA to translate the codon in the A site. Rare codons are usually translated by lowly-represented tRNAs. The ribosome must deal with stochastic binding of these tRNAs in competition with the entire pool of highly represented tRNAs, which slows down the translation of rare codons. There are a few examples where the level of transfer RNAs alters or regulates the translation of genes crucial for cellular processes based on the presence of non-optimal codons, in particular mRNAs. A mutation that replaces an optimal codon by a rare synonymous codon in human CFTR gene lead to a misfolded and malfunctioning proteins. An increase in the level of tRNA decoding such a codon restores the function of CFTR (Kirchner et al., 2017). In yeast, stress-responsive genes are enriched in codons that use rare tRNAs. The tRNAs of cells exposed to different stresses are reprogrammed to respond to stress by enhancing stress mRNA translation (Torrent et al., 2018). Certain tRNAs are preferentially expressed in human cancer cells under the control of an RNA polymerase III transcription factor. The knockdown of these tRNAs reduces the proliferation of cancer cells, which indicates their crucial role in the reprogramming cell proliferation (Yang et al., 2020).

The cell cycle is a complex process that requires the temporal expression of a number of proteins that regulate the functioning of the different phases of the process in coordination. The fission yeast *S. pombe* has been a model to study the cell cycle of eukaryotes (Yanagida, 2002). Many of the cell cycle proteins from *S. pombe* are temporarily expressed based on the transcription and translation of corresponding genes, as well as the degradation of gene products in a well-coordinated process (Hayles et al., 2013). Several of these proteins are enriched in rare codons,

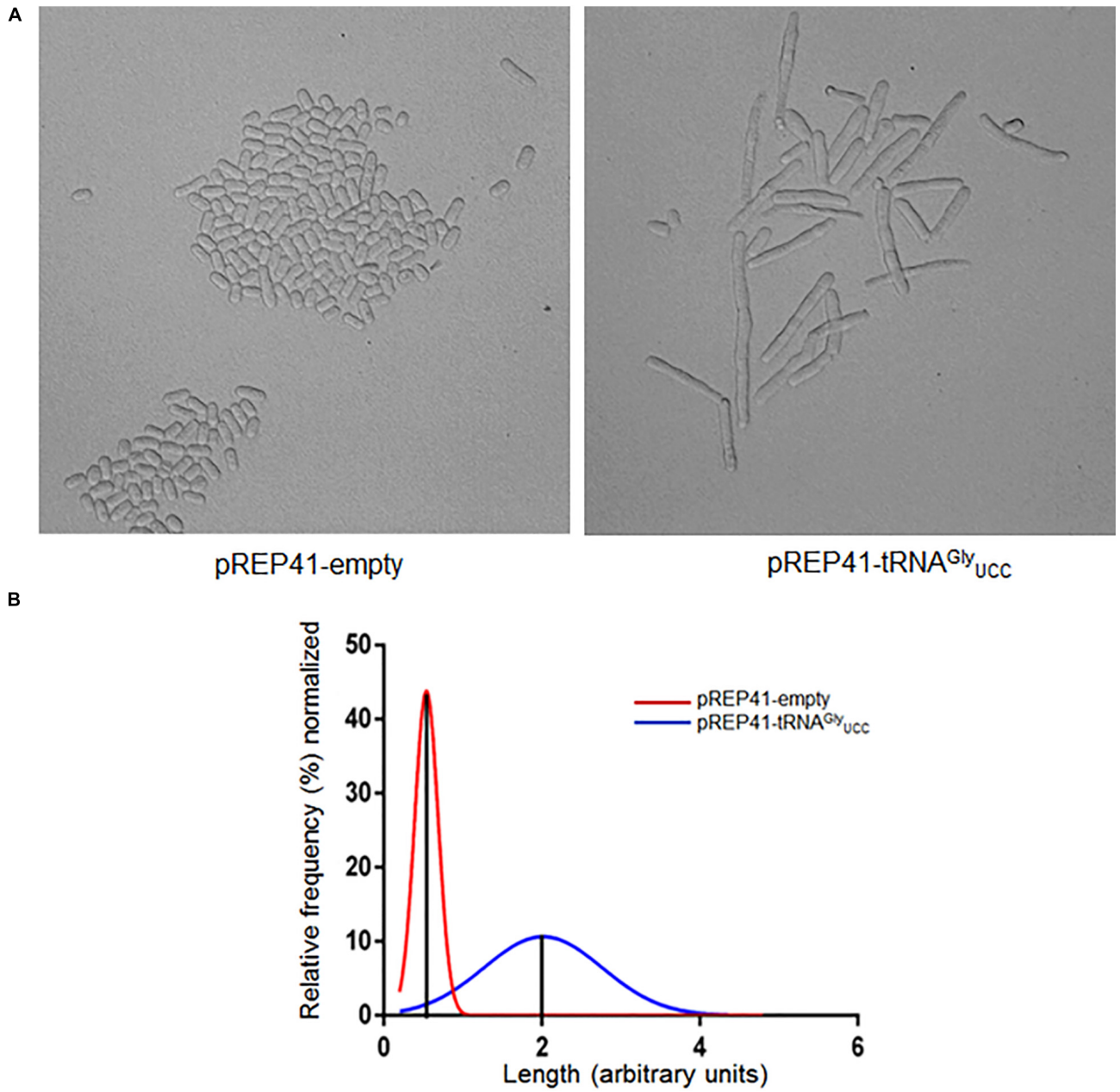


FIGURE 4 | Overexpression of tRNA$^{Gly}_{UCC}$ causes morphological changes, inducing cell elongation. **(A)** Phenotype observed under a microscope of cells synchronized with 15 mM HU for 4 hrs. The images were obtained under a microscope at 400x at 12 h post-incubation with HU (cell cycle stage: G2). **(B)** Histograms of the length of cells that overexpress tRNA$^{Gly}_{UCC}$ (blue line) normalized by the average length of control cells (red line).

giving rise to the notion that translation efficiency is in part a way to control the level of some proteins. The levels of transfer RNAs oscillate markedly, with increases in the G2 phase, concomitant with an increase in the activity of several amino acyl tRNA synthetases, including GlyRS (Frenkel-Morgenstern et al., 2012). The two molecular events match with the expression patterns of certain cyclins. These observations reinforce the idea that translation of these proteins is regulated cyclically by tRNA levels.

Cdc13 is a crucial protein that controls the G2-M transition of the *S. pombe* cell cycle. The presence of five non-optimal codons of the thirteen codons for glycine in *cdc13* suggests that the level of this protein is controlled at the translational level. The data obtained in this work reveals that alterations of rare codons decoding glycine in the gene that encodes Cdc13 have a profound impact on cell proliferation. First, the replacement of all five codons by optimal counterparts, where 2 of the 5 rare glycine

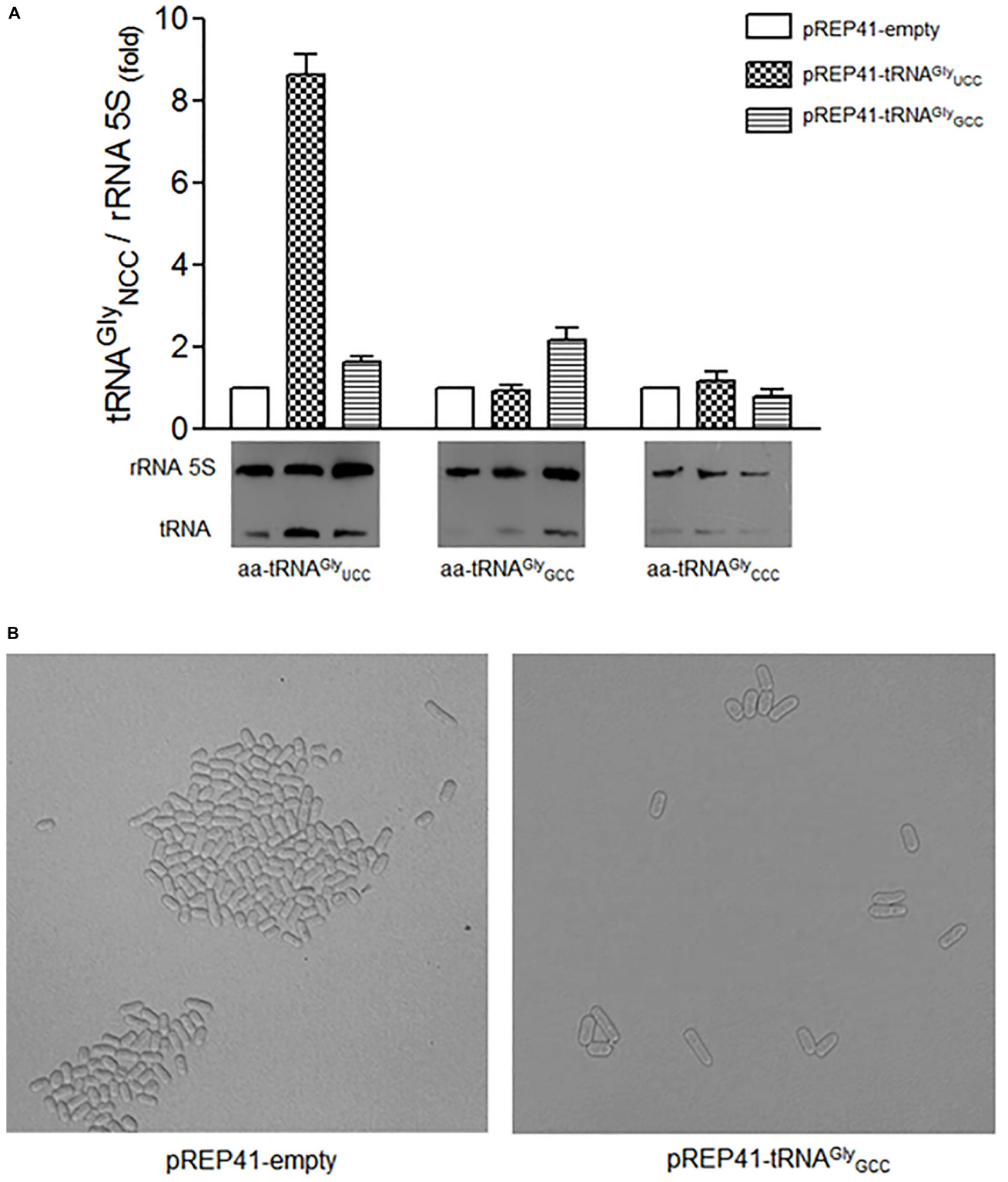


FIGURE 5 | Overexpression of tRNA$^{Gly}_{GCC}$ did not change the level of other tRNA^{Gly} isoaceptors and did not cause morphological changes in cell elongation. **(A)** Relative levels of tRNA$^{Gly}_{UCC}$, tRNA$^{Gly}_{GCC}$, tRNA$^{Gly}_{CCC}$ isoacceptors in cells transformed with the indicated constructs. **(B)** Phenotype observed under a microscope of cells overexpressing tRNA$^{Gly}_{GCC}$ synchronized with 15 mM HU for 4 hrs. The images were obtained under a microscope at 400x at 12 hrs post-incubation with HU (cell cycle stage: G2).

codons are in the cyclin N-terminal domain (**Figure 1**), and the other two are located in tandem near the C-terminus, results in the almost complete impairment of the cells to proliferate in solid or liquid media (**Supplementary Figure 1**). It is believed that synonymous codons tend to be translated more slowly than optimal counterparts (Shah and Gilchrist, 2011; Boël et al., 2016) suggesting an alteration of co-translational protein folding. Other reports suggest that protein synthesis is under selective evolutionary pressure by co-translational folding (Chaney et al., 2017; Jacobs and Shakhnovich, 2017). One hypothesis is that subtle modifications in the elongation rate affect the folding mechanism (Braselmann et al., 2013), although cells have molecular chaperones that help in folding proteins adequately, including folding nascent polypeptides (Kramer et al., 2019). According to our results it is possible that the replacement of rare glycine codons by optimal counterparts results in misfolded Cdc13, as has been suggested (Zhou et al., 2015), This would reduce the levels of functional Cdc13, although there may be other interpretations of the results (see below).

Studies have shown that a complex between Cdc2 (Cdk1) and Cdc13 is required for the cells in *S. pombe* to enter mitosis. It has been reported that the deletion of the *cdc13* gene gives rise to small cells that do not enter mitosis, although some of them can continuously replicate, giving rise to elongated cells with giant nuclei (Hayles et al., 1994). Patterson et al. (2019) described another example of elongated cells and demonstrated that Cdc13 expression below wild-type levels results in larger cells. They found a correlation between Cdc13 expression levels and cell size at division. Using mutant cells with a thermosensitive mutation in *cdc13,* they found that at the restrictive temperature, the complex was largely in the insoluble fraction, which prevents the cell from entering mitosis (Hayles et al., 1994). Our results show that overexpression of $tRNA^{Gly}_{UCC}$ is accompanied by a substantial increase in Cdc13 in the aggregated protein fraction (up to 50% of the all Cdc13 proteins, **Figure 3**), along with the formation of elongated cells (**Figure 4**). The effects observed in our work may be the result of several different events, such as aggregation of other proteins that drag Cdc13, impairment of chaperones to properly fold Cdc13, or effects on other cyclins. However, the observed effects are consistent with the role of Cdc13 in forming the complex described above (Hayles et al., 1994; Humaidan et al., 2018) and the cell replication problem as a consequence of replacing rare Gly codons in *cdc13.* Further experiments are required to confirm whether there is a direct effect on Cdc13 folding and aggregation that alters the cell cycle. Nevertheless, these two effects seem to be specific to $tRNA^{Gly}_{UCC}$, as they are not observed when the tRNAs that decode optimal glycine or arginine codons are overexpressed.

The overexpression of Gly-$tRNA^{Gly}_{UCC}$ probably exerts an effect on not only *cdc13* mRNA, but also many other mRNAs containing the decoded codons. Thus, the observed cell division phenotype maybe the consequence of the altered expression of other genes involved in the cell cycle. Analysis of the effect of global proteome alterations on $tRNA^{Gly}_{UUC}$ overexpression, or the overexpression of any other tRNAs will certainly give insights into the role of rare codons in the cell cycle, as well as the selective pressure that allows the natural selection of rare codons in cell cycle proteins.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

LA conducted most of the experiments as part of her Magister thesis. FM, DG, and RF-R contributed with part of the experimental data. MT conducted all bioinformatic analysis. SM supervised the thesis of LA and contributed with her expertise on *S. pombe* manipulation. AK contributed with ongoing discussion and reviewing the manuscript. OO provided the funds, conceptual questions, ongoing discussion and training of students. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by Fondecyt, Chile by grants 1150834 and 1190552 to OO, 1191074 to AK, and 3150366 to SM. LA was recipient of the Conicyt, Chile fellowship for graduate studies.

ACKNOWLEDGMENTS

We thank Fondecyt, Chile for financial support and the Universidad de Chile and Universidad de Santiago for logistical support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.607693/full#supplementary-material

REFERENCES

Arquès, D. G., and Michel, C. J. (1996). A complementary circular code in the protein coding genes. *J. Theor. Biol.* 182, 45–58. doi: 10.1006/jtbi.1996.0142

Bali, V., and Bebok, Z. (2015). Decoding mechanisms by which silent codon changes influence protein biogenesis and function. *Int. J. Biochem. Cell Biol.* 64, 58–74. doi: 10.1016/j.biocel.2015.03.011

Basu, S., Roberts, E. L., Jones, A. W., Swaffer, M. P., Snijders, A. P., and Nurse, P. (2020). The hydrophobic patch directs cyclin B to centrosomes to promote

global CDK phosphorylation at mitosis. *Curr. Biol.* 30, 883.e4–892.e4. doi: 10.1016/j.cub.2019.12.053

Boël, G., Letso, R., Neely, H., Price, W. N., Wong, K. H., Su, M., et al. (2016). Codon influence on protein expression in *E. coli* correlates with mRNA levels. *Nature* 529, 358–363. doi: 10.1038/nature16509

Brandis, G., and Hughes, D. (2016). The selective advantage of synonymous codon usage bias in *Salmonella*. *PLoS Genet.* 12:e1005926. doi: 10.1371/journal.pgen.1005926

Braselmann, E., Chaney, J. L., and Clark, P. L. (2013). Folding the proteome. *Trends Biochem. Sci.* 38, 337–344. doi: 10.1016/j.tibs.2013.05.001

Buhr, F., Jha, S., Thommen, M., Mittelstaet, J., Kutz, F., Schwalbe, H., et al. (2016). Synonymous codons direct cotranslational folding toward different protein conformations. *Mol. Cell.* 61, 341–351. doi: 10.1016/j.molcel.2016.01.008

Bulmer, M. (1991). The selection-mutation-drift theory of synonymous codon usage. *Genetics* 129, 897–907.

Chaney, J. L., and Clark, P. L. (2015). Roles for synonymous codon usage in protein biogenesis. *Annu. Rev. Biophys.* 44, 143–166. doi: 10.1146/annurev-biophys-060414-034333

Chaney, J. L., Steele, A., Carmichael, R., Rodriguez, A., Specht, A. T., Ngo, K., et al. (2017). Widespread position-specific conservation of synonymous rare codons within coding sequences. *PLoS Comput. Biol.* 13:e1005531. doi: 10.1371/journal.pcbi.1005531

Chatenay, D., Cocco, S., Greenbaum, B., Monasson, R., and Netter, P. (2017). "Evolutionary constraints on coding sequences at the nucleotidic level: a statistical physics approach," in *Evolutionary Biology: Self/Nonself Evolution, Species and Complex Traits Evolution, Methods and Concepts*, ed. P. Pontarotti (Cham: Springer), 329–367. doi: 10.1007/978-3-319-61569-1_18

Choi, H., Gabriel, K., Schneider, J., Otten, S., and McClain, W. H. (2003). Recognition of acceptor-stem structure of tRNA(Asp) by *Escherichia coli* aspartyl-tRNA synthetase. *RNA* 9, 386–393. doi: 10.1261/rna.2139703

Dilucca, M., Cimini, G., and Giansanti, A. (2018). Essentiality, conservation, evolutionary pressure and codon bias in bacterial genomes. *Gene* 663, 178–188. doi: 10.1016/j.gene.2018.04.017

Forsburg, S. L. (1994). Codon usage table for *Schizosaccharomyces pombe*. *Yeast* 10, 1045–1047. doi: 10.1002/yea.320100806

Forsburg, S. L., and Rhind, N. (2006). Basic methods for fission yeast. *Yeast* 23, 173–183. doi: 10.1002/yea.1347

Frenkel-Morgenstern, M., Danon, T., Christian, T., Igarashi, T., Cohen, L., Hou, Y. M., et al. (2012). Genes adopt non-optimal codon usage to generate cell cycle-dependent oscillations in protein levels. *Mol. Syst. Biol.* 8:572. doi: 10.1038/msb.2012.3

Froger, A., and Hall, J. E. (2007). Transformation of plasmid DNA into E. coli using the heat shock method. *J. Vis. Exp.* 6:253. doi: 10.3791/253

Hanson, G., and Coller, J. (2018). Codon optimality, bias and usage in translation and mRNA decay. *Nat. Rev. Mol. Cell Biol.* 19, 20–30. doi: 10.1038/nrm.2017.91

Hayles, J., Fisher, D., Woollard, A., and Nurse, P. (1994). Temporal order of S phase and mitosis in fission yeast is determined by the state of the p34cdc2-mitotic B cyclin complex. *Cell* 78, 813–822. doi: 10.1016/s0092-8674(94)90542-8

Hayles, J., Wood, V., Jeffery, L., Hoe, K. L., Kim, D. U., Park, H. O., et al. (2013). A genome-wide resource of cell cycle and cell shape genes of fission yeast. *Open Biol.* 3:130053. doi: 10.1098/rsob.130053

Hiraoka, Y., Kawamata, K., Haraguchi, T., and Chikashige, Y. (2009). Codon usage bias is correlated with gene expression levels in the fission yeast Schizosaccharomyces pombe. *Genes Cells* 14, 499–509. doi: 10.1111/j.1365-2443.2009.01284.x

Humaidan, D., Breinig, F., and Helms, V. (2018). Adding phosphorylation events to the core oscillator driving the cell cycle of fission yeast. *PLoS One* 13:e0208515. doi: 10.1371/journal.pone.0208515

Ikemura, T. (1985). Codon usage and tRNA content in unicellular and multicellular organisms. *Mol. Biol. Evol.* 2, 13–34. doi: 10.1093/oxfordjournals.molbev.a040335

Ingolia, N. T., Lareau, L. F., and Weissman, J. S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* 147, 789–802. doi: 10.1016/j.cell.2011.10.002

Jacobs, W. M., and Shakhnovich, E. I. (2017). Evidence of evolutionary selection for cotranslational folding. *Proc. Natl. Acad. Sci. U.S.A.* 114, 11434–11439. doi: 10.1073/pnas.1705772114

Jeacock, L., Faria, J., and Horn, D. (2018). Codon usage bias controls mRNA and protein abundance in trypanosomatids. *eLife* 7:e32496. doi: 10.7554/eLife.32496

Kanaya, S., Yamada, Y., Kudo, Y., and Ikemura, T. (1999). Studies of codon usage and tRNA genes of 18 unicellular organisms and quantification of *Bacillus subtilis* tRNAs: gene expression level and species-specific diversity of codon usage based on multivariate analysis. *Gene* 238, 143–155. doi: 10.1016/s0378-1119(99)00225-5

Kirchner, S., Cai, Z., Rauscher, R., Kastelic, N., Anding, M., Czech, A., et al. (2017). Alteration of protein function by a silent polymorphism linked to tRNA abundance. *PLoS Biol.* 15:e2000779. doi: 10.1371/journal.pbio.2000779

Kramer, G., Shiber, A., and Bukau, B. (2019). Mechanisms of cotranslational maturation of newly synthesized proteins. *Annu. Rev. Biochem.* 88, 337–364. doi: 10.1146/annurev-biochem-013118-111717

Køížek, M., and Køížek, P. (2012). Why has nature invented three stop codons of DNA and only one start codon? *J. Theor. Biol.* 304, 183–187. doi: 10.1016/j.jtbi.2012.03.026

Luche, D. D., and Forsburg, S. L. (2009). Cell-cycle synchrony for analysis of S. pombe DNA replication. *Methods Mol. Biol.* 521, 437–448. doi: 10.1007/978-1-60327-815-7_24

Michel, C. J., and Thompson, J. D. (2020). Identification of a circular code periodicity in the bacterial ribosome: origin of codon periodicity in genes? *RNA Biol.* 17, 571–583. doi: 10.1080/15476286.2020.1719311

Morgens, D. W., Chang, C. H., and Cavalcanti, A. R. (2013). Ambushing the ambush hypothesis: predicting and evaluating off-frame codon frequencies in prokaryotic genomes. *BMC Genomics* 14:418. doi: 10.1186/1471-2164-14-418

Oldenburg, K. R., Vo, K. T., Michaelis, S., and Paddon, C. (1997). Recombination-mediated PCR-directed plasmid construction in vivo in yeast. *Nucleic Acids Res.* 25, 451–452. doi: 10.1093/nar/25.2.451

Oltra, E., Verde, F., Werner, R., and D'Urso, G. (2004). A novel RING-finger-like protein Ini1 is essential for cell cycle progression in fission yeast. *J. Cell Sci.* 117(Pt 6), 967–974. doi: 10.1242/jcs.00946

Pang, Y. L., Abo, R., Levine, S. S., and Dedon, P. C. (2014). Diverse cell stresses induce unique patterns of tRNA up- and down-regulation: tRNA-seq for quantifying changes in tRNA copy number. *Nucleic Acids Res.* 42:e170. doi: 10.1093/nar/gku945

Patterson, J. O., Rees, P., and Nurse, P. (2019). Noisy cell-size-correlated expression of cyclin B drives probabilistic cell-size homeostasis in fission yeast. *Curr. Biol.* 29, 1379.e4–1386.e4. doi: 10.1016/j.cub.2019.03.011

Presnyak, V., Alhusaini, N., Chen, Y. H., Martin, S., Morris, N., Kline, N., et al. (2015). Codon optimality is a major determinant of mRNA stability. *Cell* 160, 1111–1124. doi: 10.1016/j.cell.2015.02.029

Quax, T. E., Claassens, N. J., Söll, D., and van der Oost, J. (2015). Codon bias as a means to fine-tune gene expression. *Mol. Cell* 59, 149–161. doi: 10.1016/j.molcel.2015.05.035

Rand, J. D., and Grant, C. M. (2006). The thioredoxin system protects ribosomes against stress-induced aggregation. *Mol. Biol. Cell* 17, 387–401. doi: 10.1091/mbc.e05-06-0520

Ray, S. K., Baruah, V. J., Satapathy, S. S., and Banerjee, R. (2014). Cotranslational protein folding reveals the selective use of synonymous codons along the coding sequence of a low expression gene. *J. Genet.* 93, 613–617. doi: 10.1007/s12041-014-0429-1

Salazar, J. C., Zúñiga, R., Raczniak, G., Becker, H., Söll, D., and Orellana, O. (2001). A dual-specific Glu-tRNA(Gln) and Asp-tRNA(Asn) amidotransferase is involved in decoding glutamine and asparagine codons in *Acidithiobacillus ferrooxidans*. *FEBS Lett.* 500, 129–131. doi: 10.1016/s0014-5793(01)02600-x

Sambrook, J., and Russell, D. (2001). *Molecular Cloning A laboratory Manual*, 3rd Edn. New York, NY: Cold Spring Harbor Laboratory Press.

Seligmann, H. (2010). The ambush hypothesis at the whole-organism level: off frame, 'hidden' stops in vertebrate mitochondrial genes increase developmental stability. *Comput. Biol. Chem.* 34, 80–85. doi: 10.1016/j.compbiolchem.2010.03.001

Seligmann, H. (2011). Error compensation of tRNA misacylation by codon-anticodon mismatch prevents translational amino acid misinsertion. *Comput. Biol. Chem.* 35, 81–95. doi: 10.1016/j.compbiolchem.2011.03.001

Seligmann, H. (2012). Coding constraints modulate chemically spontaneous mutational replication gradients in mitochondrial genomes. *Curr. Genomics* 13, 37–54. doi: 10.2174/138920212799034802

Seligmann, H. (2019). Localized context-dependent effects of the "ambush" hypothesis: more off-frame stop codons downstream of shifty codons. *DNA Cell Biol.* 38, 786–795. doi: 10.1089/dna.2019.4725

Seligmann, H., and Pollock, D. D. (2004). The ambush hypothesis: hidden stop codons prevent off-frame gene reading. *DNA Cell Biol.* 23, 701–705. doi: 10.1089/dna.2004.23.701

Shah, P., and Gilchrist, M. A. (2011). Explaining complex codon usage patterns with selection for translational efficiency, mutation bias, and genetic drift. *Proc. Natl. Acad. Sci. U.S.A.* 108, 10231–10236. doi: 10.1073/pnas.1016719108

Supek, F. (2016). The code of silence: widespread associations between synonymous codon biases and gene function. *J. Mol. Evol.* 82, 65–73. doi: 10.1007/s00239-015-9714-8

Torrent, M., Chalancon, G., de Groot, N. S., Wuster, A., and Babu, M. M. (2018). Cells alter their tRNA abundance to selectively regulate protein synthesis during stress conditions. *Sci. Signal.* 11:eaat6409. doi: 10.1126/scisignal.aat6409

Xu, Y., Ma, P., Shah, P., Rokas, A., Liu, Y., and Johnson, C. H. (2013). Non-optimal codon usage is a mechanism to achieve circadian clock conditionality. *Nature* 495, 116–120. doi: 10.1038/nature11942

Yanagida, M. (2002). The model unicellular eukaryote, Schizosaccharomyces pombe. *Genome Biol.* 3:COMMENT2003. doi: 10.1186/gb-2002-3-3-comment2003

Yang, J., Smith, D. K., Ni, H., Wu, K., Huang, D., Pan, S., et al. (2020). SOX4-mediated repression of specific tRNAs inhibits proliferation of human glioblastoma cells. *Proc. Natl. Acad. Sci. U.S.A.* 117, 5782–5790. doi: 10.1073/pnas.1920200117

Zhou, M., Guo, J., Cha, J., Chae, M., Chen, S., Barral, J. M., et al. (2013). Non-optimal codon usage affects expression, structure and function of clock protein FRQ. *Nature* 495, 111–115. doi: 10.1038/nature11833

Zhou, M., Wang, T., Fu, J., Xiao, G., and Liu, Y. (2015). Nonoptimal codon usage influences protein structure in intrinsically disordered regions. *Mol. Microbiol.* 97, 974–987. doi: 10.1111/mmi.13079

Zhou, Z., Dang, Y., Zhou, M., Li, L., Yu, C. H., Fu, J., et al. (2016). Codon usage is an important determinant of gene expression levels largely through its effects on transcription. *Proc. Natl. Acad. Sci. U.S.A.* 113, E6117–E6125. doi: 10.1073/pnas.1606724113

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

REVIEW
published: 21 January 2021
doi: 10.3389/fmicb.2020.624830

Coupled Transcription-Translation in Prokaryotes: An Old Couple With New Surprises

Mikel Irastortza-Olaziregi and Orna Amster-Choder**

Department of Microbiology and Molecular Genetics, Faculty of Medicine, IMRIC, The Hebrew University of Jerusalem, Jerusalem, Israel

Coupled transcription-translation (CTT) is a hallmark of prokaryotic gene expression. CTT occurs when ribosomes associate with and initiate translation of mRNAs whose transcription has not yet concluded, therefore forming "RNAP.mRNA.ribosome" complexes. CTT is a well-documented phenomenon that is involved in important gene regulation processes, such as attenuation and operon polarity. Despite the progress in our understanding of the cellular signals that coordinate CTT, certain aspects of its molecular architecture remain controversial. Additionally, new information on the spatial segregation between the transcriptional and the translational machineries in certain species, and on the capability of certain mRNAs to localize translation-independently, questions the unanimous occurrence of CTT. Furthermore, studies where transcription and translation were artificially uncoupled showed that transcription elongation can proceed in a translation-independent manner. Here, we review studies supporting the occurrence of CTT and findings questioning its extent, as well as discuss mechanisms that may explain both coupling and uncoupling, e.g., chromosome relocation and the involvement of cis- or trans-acting elements, such as small RNAs and RNA-binding proteins. These mechanisms impact RNA localization, stability, and translation. Understanding the two options by which genes can be expressed and their consequences should shed light on a new layer of control of bacterial transcripts fate.

Keywords: coupled transcription-translation, uncoupled transcription-translation, subcellular organization of prokaryotes, translation-independent mRNA localization, local translation, expressome

OPEN ACCESS

Edited by:
Hans-Georg Koch,
University of Freiburg, Germany

Reviewed by:
Evgeny Nudler,
New York University, United States
Jared Schrader,
Wayne State University, United States

****Correspondence:***
Mikel Irastortza-Olaziregi
mikel.irastorza@mail.huji.ac.il
Orna Amster-Choder
ornaam@ekmd.huji.ac.il

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *01 November 2020*
Accepted: *18 December 2020*
Published: *21 January 2021*

Citation:
Irastortza-Olaziregi M and Amster-Choder O (2021) Coupled Transcription-Translation in Prokaryotes: An Old Couple With New Surprises. Front. Microbiol. 11:624830. doi: 10.3389/fmicb.2020.624830

COUPLED TRANSCRIPTION-TRANSLATION: A HALLMARK FEATURE OF PROKARYOTIC GENE EXPRESSION

Due to the scarcity of intracellular membrane-delimited compartmentalization, prokaryotic cells have historically been regarded as spatially unorganized. The lack of a nuclear membrane that physically separates the chromosomal DNA from the cytosolic environment led to the well-accepted notion that transcription and translation are spatiotemporally coupled in bacteria and archaea. Coupled transcription-translation (CTT) occurs when ribosomes bind and start to translate nascent mRNAs, whose transcription has not terminated yet, therefore forming an "RNAP·nascent mRNA·ribosome" complex (**Figure 1**).

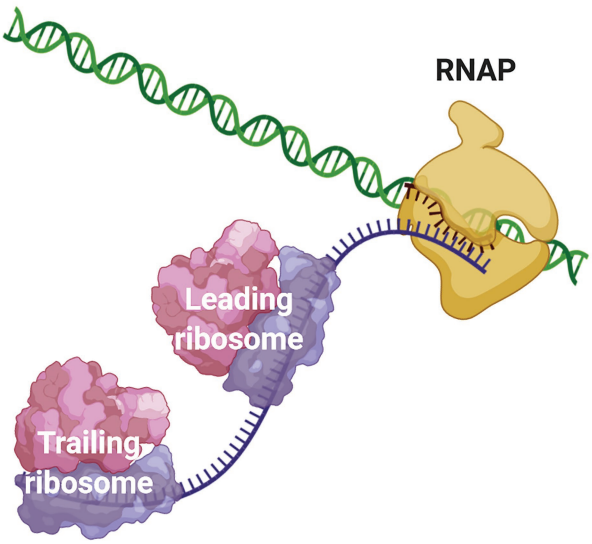


FIGURE 1 | Coupling of transcription and translation in prokaryotes. When the nascent mRNA emerges from the RNAP, the transcript is bound by the leading ribosome forming a transcribing-translating complex. Additional ribosomes can associate with the nascent mRNA to form a convoy of trailing ribosomes on the transcript that is still bound to the transcription machinery. The leading ribosome can physically interact with the RNAP or the two machineries may be connected *via* the transcript.

The fact that transcription and translation could be coupled in prokaryotes was first proposed by Stent in the mid-1960s. He argued that due to the apparent inability to dissociate nascent transcripts from the chromosome *in vitro*, an active force exerted by translating ribosomes could be necessary to release the mRNAs from their templates. Indirectly, he implied that transcription and translation could be spatiotemporally coupled (Stent, 1964). Subsequent *in vitro* work from the Nirenberg lab demonstrated DNA·RNA·ribosome complexes (Byrne et al., 1964; Bladen et al., 1965). In the early 1970s, Miller et al. (1970) published an electron microscopy image, which showed ribosomes strongly associating with and translating nascent mRNA in a concatenated fashion forming polysomes. More recently, similar observations were reported in archaea (French et al., 2007), extending the occurrence of CTT to all prokaryotes. Miller's micrographs have illustrated microbiology textbooks for decades and CTT is nowadays a widely accepted dogma of prokaryotic gene expression.

In this review, we will first highlight the biological significance of maintaining proper CTT and present the current understanding of how transcription and translation are coordinated under different growth conditions. We will then discuss the recent findings that shed light on the mechanistic and molecular details that mediate the physical coupling of these two processes. We will proceed by describing different models explaining where CTT takes place in the context of the prokaryotic cell. We will then introduce several types of evidence that challenge the CTT dogma and suggest that its occurrence may not be as general as currently assumed. Furthermore, we will extensively elaborate on molecular mechanisms that potentially promote the spatiotemporal uncoupling of transcription and translation. We will conclude by discussing open questions and principles emanating from this old but still exciting couple.

COORDINATION OF CTT

In optimal growth conditions, *Escherichia coli* ribosomes translate 14–17 amino acids per second (Young and Bremer, 1976; Proshkin et al., 2010; Zhu et al., 2016), meaning that they translocate about 42–51 nucleotides per second (nt/s) along the mRNA being translated. On the other hand, RNAP synthesizes mRNA at a rate of 42–49 nt/s (Proshkin et al., 2010; Iyer et al., 2018). Thus, mRNA transcription and translation rates are well-matched. Of note, translation and transcription rates vary across different growth conditions, but the rates of both processes remain coordinated (Vogel and Jensen, 1994b; Proshkin et al., 2010; Iyer et al., 2018), suggesting CTT coordination is important.

Indeed, a balanced CTT regime is crucial for the proper function of *E. coli* cells, and the uncoupling of transcription and translation can lead to multiple conflicts that compromise cell viability. Many such conflicts are related to the intimate link between CTT and Rho-mediated premature transcription termination (PTT). It was assumed for years that, in the absence of ribosomes engaged in CTT, Rho utilization (*rut*) sites in the nascent transcripts become exposed and are readily recognized by Rho, which translocates towards RNAP and causes the disassembly of the transcription elongation complex (TEC; Chalissery et al., 2011; Lawson et al., 2018). Recent evidence, on the other hand, supports a different mechanism where, in the absence of a physically coupled ribosome, transcription termination factor Rho associates with RNAP early after transcription initiation (Mooney et al., 2009a) *via* protein-protein interactions with NusA and NusG, rendering the TEC into a moribund pre-termination complex (PTC). This state favors Rho recognition of the *rut* sites in the nascent mRNA, which is followed by the closure of the Rho ring and disassembly of the TEC (Epshtein et al., 2010; Hao et al., 2020; Said et al., 2020). These molecular events lead to PTT and operon polarity (Richardson, 1991). Ribosomes and Rho compete for the same binding interface of NusG and, thus, a coupled ribosome prevents Rho-mediated PTT (Burmann et al., 2010). Coupled ribosomes can additionally antiterminate intrinsic terminators (Li et al., 2015). Moreover, leading ribosomes push stalled RNAPs forward and facilitate elongation over transcriptional roadblocks (Proshkin et al., 2010; Stevenson-Jones et al., 2020). Beyond gene expression, this is especially important for avoiding clashes between replisomes and backtracked RNAPs, which cause double-strand breaks and lead to genome instability (Dutta et al., 2011). Furthermore, cotranscriptional translation prevents reannealing of the nascent RNA to the template DNA strand, which can give way to dangerous R-loops (Gowrishankar and Harinarayanan, 2004), and also protects nascent transcripts from ribonucleolytic attack (Makarova et al., 1995; Deana and Belasco, 2005). In agreement with this, the stability of *lacZ*

transcripts was greatly reduced when transcribed by T7 RNAP (Iost and Dreyfus, 1995), which transcribes about 230 nt/s (Golomb and Chamberlin, 1974) and cannot be closely followed by translating ribosomes. Tight CTT could be of critical importance for the expression of genes whose decay initiates before the transcription of the full mRNA is completed (Chen et al., 2015). For this population of transcripts, post-transcriptional translation appears to be unfeasible and ribosomes should closely follow RNAP before ribonucleases trigger the degradation of the nascent mRNA. Additionally, it is speculated that physically associated ribosomes can perform a pioneering round of translation for mRNA quality control, a well-characterized phenomenon in eukaryotes (Maquat et al., 2010). On the other hand, controlled uncoupling of transcription and translation can also be utilized for gene regulation. A classical example of this is the operon for tryptophan biosynthesis (Yanofsky, 1981), where due to low tryptophan concentrations the leader ribosome lags behind RNAP and allows the formation of an antitermination secondary structure that precludes the formation of the attenuator structure.

rRNA transcription (90 nt/s; Vogel and Jensen, 1994b) exceeds by far the translation elongation rates in *E. coli*, implying that, when transcribing mRNA, RNAP does not work at its maximum biosynthetic capacity, and that mechanisms for equalizing mRNA transcription and translation exist. In this regard, physical CTT offers an elegant model to explain the transcription-translation correlation under different growth conditions. According to this model, translation elongation rates of the leading ribosome, dictate transcription elongation rates (Proshkin et al., 2010). This model is supported by the finding that in strains harboring ribosomal mutations that slow translation, transcription elongation rates decreased accordingly, and that transcription-translation rates of different genes correlated with their rare codon content, highlighting the role of translation in dictating transcription elongation (Proshkin et al., 2010). In this model, the leading ribosome equalizes transcription-translation rates by physically pushing forward the RNAP, which otherwise tends to spontaneous backtracking and/or pauses (Proshkin et al., 2010; Stevenson-Jones et al., 2020), so that CTT remains coordinated and futile transcription is prevented. Of note, this mode of CTT coordination does not necessarily imply persistent RNAP·ribosome interactions or the formation of a stable complex, and it could rather be driven by occasional ribosome-to-RNA pushing contacts at sites where transcription slows down.

Besides CTT coordinated by physical RNAP·ribosome contacts, recent evidence argues in favor of additional mechanisms for CTT coordination. In contrast to common thought, chloramphenicol does not inhibit translation by reducing ribosome elongation rates, but rather by diminishing the population of ribosomes engaged in translation (Dai et al., 2016). In contrast, cells challenged with fusidic acid showed slower translation elongation rates, but strikingly, no reduction in transcription elongation rates (Zhu et al., 2019), suggesting that transcription elongation can be modulated independently to translation. Also, in cells subjected to nitrogen starvation, which is characterized by a slowdown in translation elongation rates, transcription elongation slowed down correspondingly and no hints of PTT were observed (Iyer et al., 2018). Importantly, translation inhibition by chloramphenicol did not decrease transcription elongation rates under these experimental conditions (Iyer et al., 2018), indicating that the RNAP elongates independently of the leading ribosome. Possibly, under conditions where translation slows down below a certain threshold, transcriptional cooperation among RNAPs (see section The CTT Dogma has Been Challenged: Towards Uncoupled Transcription-Translation?) might suffice to maintain transcriptional elongation in a ribosome-independent manner. Additional studies support a model where the occurrence of CTT is a stochastic event that does not depend on the physical contact of both machineries (see section The CTT Dogma has Been Challenged: Towards Uncoupled Transcription-Translation?; Li et al., 2015; Chen and Fredrick, 2018, 2020). Thus, it is reasonable to argue that, besides CTT coordinated by physical contacts between the leading ribosome and RNAP, additional mechanisms ensure the coordination of CTT.

Notably, the alarmone (p)ppGpp has recently emerged as an important player in CTT coordination. Upon amino acid starvation, the ribosome-associated RelA detects uncharged tRNAs at the A-site of the ribosomes and triggers the rapid synthesis of (p)ppGpp (Wendrich et al., 2002), inducing stringent response. Traditionally (p)ppGpp has been linked to transcriptional regulation, primarily by binding to the interface between β and ω subunit of the RNAP (Mechold et al., 2013; Zuo et al., 2013) and negatively regulating rRNA transcription, as well as a transcription of a myriad of coding sequences (Potrykus and Cashel, 2008; Sanchez-Vazquez et al., 2019). In agreement with this, in cells treated with fusidic acid that maintained normal transcription (p)ppGpp accumulation caused a slowdown of transcription elongation in an alarmone dose-dependent manner (Zhu et al., 2019). Alarmone-dependent slowdown of transcription elongation was also observed in cells subjected to amino acid starvation (Vogel et al., 1992; Vogel and Jensen, 1994a) and nitrogen limitation (Iyer et al., 2018).

Recent publications have expanded the (p)ppGpp targetome (Zhang et al., 2018; Wang et al., 2019), and additional alarmone-mediated layers of regulation have been discovered (reviewed in Hauryliuk et al., 2015). For instance (p)ppGpp competes with GTP for binding translation initiation factor 2 and elongation factor G (EF-G; Milon et al., 2006; Mitkevich et al., 2010), inhibiting their activity. Interestingly, under nitrogen starvation, ribosome stalling at glutamine codons diminishes translation elongation, and alarmone-defective strains show even slower translational elongation rates compared to wild-type cells (Li et al., 2018). Apparently, bacteria prevent generalized translation initiation, which could deplete pools of charged tRNAs and lead to incomplete translational rounds, and instead ensure that basal translation of housekeeping and stress-responsive genes remains active (Dai et al., 2016). This phenomenon has been proven essential for survival under oxidative stress, which is also characterized by a dramatic slowdown of translation elongation rates (Zhu and Dai 2019, 2020). It was recently reported that (p)ppGpp-dependent translational selectivity depends on structural motifs formed by the mRNAs (Vinogradova et al., 2020). Thus, upon environmental challenges that reduce translation

elongation rates, cells achieve CTT coordination by slowing down RNAP elongation and by prioritizing translation of a subset of essential genes that avoid alarmone-mediated translation inhibition. Highlighting the importance of this regulation is the finding that disrupting (p)ppGpp-mediated CTT coordination sensitizes cells to nitrogen starvation (Iyer et al., 2018).

Alarmones also coordinate transcription and translation less straightforwardly. It has recently been shown that (p)ppGpp binds inosine-guanosine kinase, which prevents the excessive accumulation of purine nucleotides and leads to pRpp synthesis, a precursor molecule for pyrimidine nucleotides, tryptophan, and histidine (Wang et al., 2020a). Thus, alarmones further coordinate CTT indirectly, by balancing nucleotide stocks required for transcription and ensuring tryptophan and histidine availability to maintain basal translation. Interestingly, the operons for the biosynthesis of these amino acids are regulated by transcriptional attenuation (Yanofsky, 1981), which links their regulation to sensing the CTT status.

The evidence presented above supports a scenario where CTT coordination is preserved by different, although complementary means, depending on the growth conditions (**Figure 2**). When transcription slows down below translation elongation rates, due to spontaneous backtracking or pausing, the leading ribosome physically pushes the RNAP forward in a way that translation rates dictate equal transcription elongation rates (Proshkin et al., 2010; Stevenson-Jones et al., 2020). Under conditions where ribosome elongation lags behind transcription, such as nitrogen limitation or oxidative stress, multilayered regulation *via* (p)ppGpp ensures that RNAP elongation rates decrease accordingly and maintains CTT coordination without the need for physical contacts of the transcription and translation machineries. Considering the ongoing discovery of (p)ppGpp-regulated processes and the tight relationship of alarmones with the metabolic status of the cell, we expect that additional pathways ensuring CTT coordination *via* (p)ppGpp will be discovered. These complementary mechanisms, i.e., physical contacts between RNAP and the leading ribosome and alarmone-mediated coordination, account for the observed correlation between transcription and translation elongation rates under different growth conditions.

THE MOLECULAR ARCHITECTURE OF CTT

How do the transcriptional and translational machineries intimately interact? Several recent publications have shed light on the molecular interactions that mediate the physical coupling of transcription and translation in *E. coli*, although their conclusions were not always in agreement (reviewed by Conn et al., 2019). CTT was initially proposed to be mediated by NusG. Specifically, nuclear magnetic resonance (NMR) experiments showed that the N-terminal domain (NTD) and

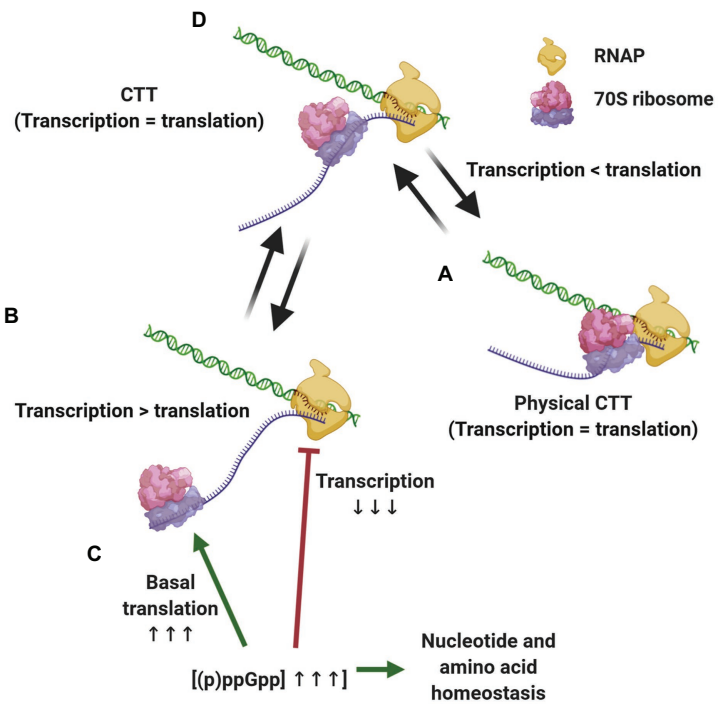


FIGURE 2 | Strategies for maintaining coupled transcription-translation (CTT) coordination. Under conditions that slow down transcription elongation rates below translational rate, such as carbon limitation and RNAP backtracking **(A)**, the leading ribosome catches up with and physically pushes the transcription elongation complex (TEC) forward, equalizing transcription and translation rates. When translation elongation lags behind transcription rate, e.g., upon nitrogen limitation and oxidative stress **(B)** (p)ppGpp concentrations increase, slowing down transcription elongation, promoting basal translation and balancing nucleotide and amino acid stocks **(C)**, thus equalizing transcription and translation elongation rates **(D)**.

C-terminal domain (CTD) of NusG interact with RNAP and NusE, which doubles as S10 protein in the ribosome (Mooney et al., 2009b; Burmann et al., 2010), respectively, thus, bridging the transcriptional and translational machineries. As Rho and the leading ribosome compete for NusG binding, this mode of coupling also explains why TECs engaged in CTT avoid Rho-mediated termination (Burmann et al., 2010). CTT of a subset of horizontally acquired operons was shown to be bridged by the NusG homolog RfaH in a similar fashion (Burmann et al., 2012; Zuber et al., 2019). The NusG-dependent CTT model was further supported by *in vivo* and *in vitro* experiments evidencing that NusG simultaneously interacts with the ribosome and RNAP (Saxena et al., 2018). A cryo-electron microscopy (cryo-EM) structure of a TEC·NusG complex did not show any defined density for the NusG CTD, suggesting that NusG mediates a flexible RNAP-ribosome association (Kang et al., 2018b). This was confirmed by another cryo-EM structure of a ribosome-bound NusG, that showed a defined density for the CTD but not for the NTD (Washburn et al., 2020). This latter publication also reported NMR results supporting the association of an RNAP·NusG complex with S10.

In parallel, a series of publications posited a model where transcription and translation are coupled independent of NusG, *via* direct RNAP·ribosome interactions. Chemical cross-linking experiments demonstrated multiple NusG-independent interactions between RNAP and both 30S and 50S ribosome subunits (Fan et al., 2017), and an RNAP·30S subunit cryo-EM structure (Demo et al., 2017) recapitulated several of the interactions detected by cross-linking. Simultaneously, a cryo-EM structure of a transcribing-translating RNAP·ribosome complex, named expressome, showed that direct interactions between both machineries can mediate the coupling (Kohler et al., 2017). Different from the previous works described above, this structure was produced by colliding a translating ribosome against a transcriptionally stalled RNAP. In this structure, the exit site of RNAP and the mRNA entry site of the ribosome were closely placed, which suggests continuous protection of the nascent mRNA from transcription to translation. This protection also excludes Rho from accessing nascent transcripts and from terminating transcription. Importantly, the NusG binding partners in the CTT complex, i.e., the β'-subunit of RNAP and the S10 protein of the ribosome, were located on opposite sites of the complex and the NusG linker was too short to bridge such distance. Thus, the collided expressome was not compatible with NusG-mediated CTT (Kohler et al., 2017).

How can the divergent views of NusG-dependent and -independent CTT, both emerging from structural studies, be reconciled? Two simultaneous studies presenting cryo-EM structures suggest that these views can co-exist. The first study by the Weixlbaumer lab describes the structures of transcribing-translating complexes assembled on mRNA scaffolds that allow different distances between the RNAP active site and the ribosomal P-site (Webster et al., 2020). One structure showed, for the first time, simultaneous binding of the NTD and CTD of NusG to RNAP and the leading ribosome, correspondingly. In this structure NusG forms a bridge between the ribosome and RNAP, thus stabilizing the interaction interface between the two machineries. Increasing the length of the intervening mRNA resulted in a similar structure, i.e., a NusG-coupled expressome. However, shortening of the intervening mRNA resulted in a structure in which RNAP is located closer to the ribosome entry tunnel. In this collided expressome, RNAP could still bind NusG, but the latter was not able to bridge the distance to S10. Hence, as shown previously (Kohler et al., 2017), the collided expressome is not compatible with NusG-mediated CTT. Of note, a structure nearly identical to this collided expressome was obtained when assembling expressomes on short scaffolds in the absence of NusG (Kohler et al., 2017). The conclusions that emerge from these structures are that coupling *via* NusG restrains RNAP motions, that the length of the intervening mRNA determines whether NusG can be involved, that expressome formation is strictly mRNA-dependent, although RNAP·30S complexes were previously observed (Demo et al., 2017), and that both the NusG-bridged and the collided expressomes are compatible with translation factor binding (Webster et al., 2020).

The second study, published by the Ebright lab, in addition to confirming many of the results presented by Weixlbaumer and coworkers and showing that collided expressomes are incompatible with ribosome·NusG binding, provides insights into the participation of NusA in CTT (Wang et al., 2020b). The researchers' present high-resolution structures for expressomes that involve both NusG and NusA in CTT. In these structures, which accommodated spacer mRNAs of different lengths, NusA promoted expressome assembly by acting as a "coupling pantograph" between RNAP and the S2/S5 protein of the leading ribosome.

Based on molecular modeling, Ebright and co-workers point out several features of the collided expressome that question its capability to promote proper gene expression (Wang et al., 2020b). First, the collided expressome is sterically incompatible with NusA binding (Guo et al., 2018), with the formation of Q-dependent antitermination complexes (Shi et al., 2019; Yin et al., 2019) and with the formation of pause and termination RNA hairpins (Kang et al., 2018a; Roberts, 2019). Regarding translation, the collided expressome is not compatible with the swiveling of the 30S subunit head that takes place during ribosome translocation (Schuwirth et al., 2005; Ratje et al., 2010; Guo and Noller, 2012). Furthermore, collided expressomes lack densities corresponding to the RNAP ω subunit, which assists in TEC assembly and, by binding (p)ppGpp, mediates CTT coordination during stringent response (see section Coordination of CTT; Kurkela et al., 2020). Thus, it is highly unlikely that collided expressomes are responsible for general CTT and, instead, they could be specialized complexes in charge of CTT under specific conditions, or even anomalous complexes resulting from RNAP-ribosome clashes (Wang et al., 2020b). Supporting the latter, RNAP from a collided expressome was shown to reinitiate transcription elongation *in vitro* and detach from the ribosome that is purposely stalled (Stevenson-Jones et al., 2020), suggesting that it is not a stable complex with significant biological functions.

We propose a mechanistic model (**Figure 3**) where different expressomes could come into play sequentially during different stages of *E. coli* CTT. Translation initiation is a relatively lengthy

process (median time of 15–30 s; Siwiak and Zielenkiewicz, 2013; Shaham and Tuller, 2018) compared to transcription elongation (49 nt/s; Iyer et al., 2018), so according to our model, RNAP can elongate several dozens, if not hundreds of nucleotides beyond the start codon translation-independently before the leader ribosome initiates elongating. Once the leading ribosome begins elongating, the two machineries are remotely connected by the nascent mRNA as an uncoupled expressome. The two machineries could remain kinetically coupled and the uncoupled expressome could conduct gene expression under conditions where the non-physical coupling is coordinated *via* (p)ppGpp (see section Coordination of CTT). In other cases, this kinetic coupling precedes the formation of a coupled expressome. Specifically, RNAP tends to pause within the first 100 nt after the start codon (Mooney et al., 2009a; Larson et al., 2014), which can allow the leading ribosome to catch up with the TEC. Several types of data indicate that ribosome translocation along the nascent mRNA to the proximity of the TEC may be a prerequisite for the formation of a physically coupled expressome. Firstly, NusG association with TECs, which occurs only after substantial transcription (Mooney et al., 2009a), is facilitated by cotranscriptional translation (Washburn et al., 2020). Additionally, the intranucleoidal ribosome concentration (2–8 μM; Bakshi et al., 2012; Sanamrad et al., 2014) is substantially lower than the NusG·ribosome dissociation constant (50 μM; Burmann et al., 2010), so the NusG·ribosome association is likely favored by the translocation of the leading ribosome towards the TEC along the nascent mRNA. Thus, the arrangement of the uncoupled expressome promotes interactions of NusG with both machineries and favors the formation of a coupled expressome. This complex could be further stabilized by NusA (Wang et al., 2020b), which associates with elongating TECs early after transcription initiation (Mooney et al., 2009a).

The coupled expressome may proceed with CTT. Yet, transcriptional and translational elongation rates can respond to independent signals, resulting in varying lengths of the nascent mRNA connecting the RNAP and the leading ribosome (Conn et al., 2019). For instance, transcriptional pauses or backtracking reduce the distance between the leading ribosome and RNAP, which can give way to the formation of a collided expressome (Kohler et al., 2017; Wang et al., 2020b; Webster et al., 2020). As the leading ribosome pushes the stalled RNAP forward (Stevenson-Jones et al., 2020), the coupled expressome can be reconfigured. Such "collision-and-reconfiguration" events can also occur in the case of uncoupled expressomes. On the other hand, translational roadblocks will increase the distance between the leading ribosome and RNAP. As the mRNA connecting the two machineries in the coupled expressome is exposed to the solvent (Wang et al., 2020b; Webster et al., 2020), the longer nascent mRNA can be accommodated by looping out from the coupled expressome (Conn et al., 2019).

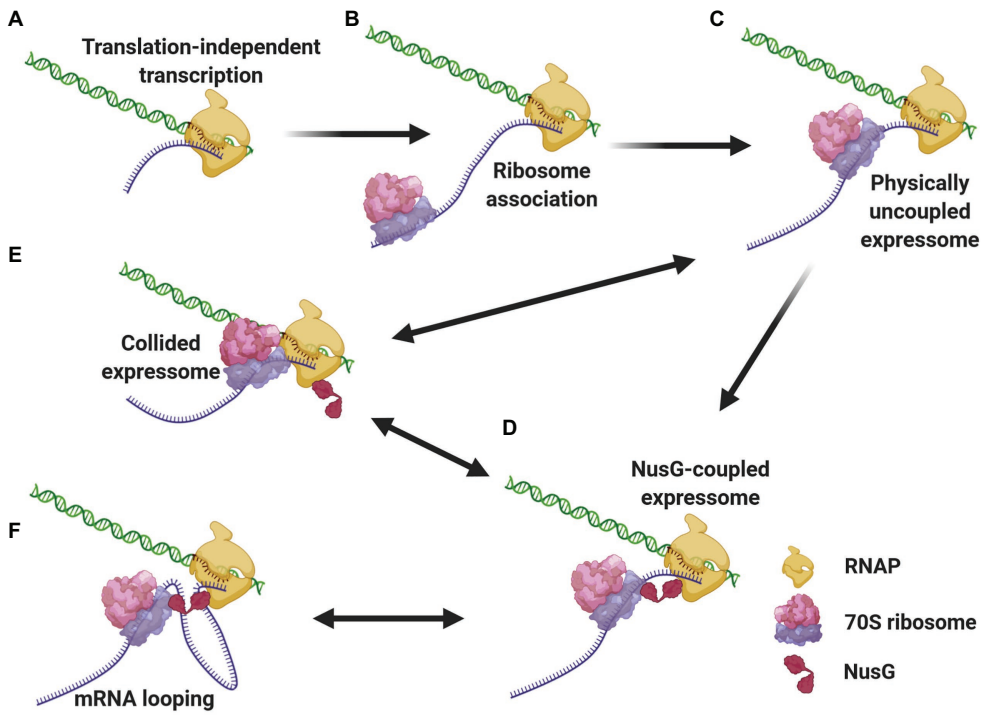


FIGURE 3 | Different expressome configurations suggested to engage in gene expression. Transcription initiates over 5'-proximal coding sequences, eventually forming the TEC **(A)**. After considerable elongation, the leading ribosome associates with the nascent transcript and initiates translation **(B)**. RNAP pauses, allowing the leading ribosome to catch up with the TEC, forming a physically-uncoupled expressome **(C)**. The proximity of the RNAP and the leading ribosome allows for NusG association with both machineries, forming a NusG-bridged coupled expressome **(D)**. Transcriptional roadblocks can reduce the length of the connecting mRNA, giving way to a collided expressome, either with or without NusG **(E)**. Conversely, translational slowdown lengthens the connecting mRNA, which can loop out the complex **(F)**. The two latter events are reversible, giving way to a rearrangement of the coupled expressome.

Decreased transcription elongation and/or increased translation elongation reduce the length of the intervening mRNA and loop the protruding mRNA back into the complex.

The in-cell architectures of *Mycoplasma pneumoniae* expressomes published recently show similarities to the *E. coli* expressomes, as well as differences (O'Reilly et al., 2020). On one hand, structures obtained from pseudouridimycin-halted *M. pneumoniae* TECs, which likely represent collided expressomes, showed direct interactions between RNAP and the leading ribosome, similar to *E. coli* collided expressomes. On the other hand, *M. pneumoniae* elongating expressomes showed neither direct nor NusG-mediated interactions. Instead, the coupling between RNAP and the leading ribosome was mediated solely by NusA (O'Reilly et al., 2020). Thus, although direct RNAP·ribosome interactions, which characterize collided expressomes, seem a conserved phenomenon, the molecular actors and interactions that drive factor-mediated RNAP-ribosome coupling differ in evolutionarily unrelated species.

It should be emphasized that all *E. coli* expressome structures discussed in this section were obtained *in vitro*, and that their existence *in vivo* remains to be investigated. Also, the questions of whether RNAP and the lead ribosome are physically coupled on the nascent mRNA *in vivo* and whether there is a mechanism to ensure or promote this coupling remain open.

THE CELL BIOLOGY OF CTT

Where in the bacterial cell does CTT occur? The subcellular organization of the transcriptional and translational machineries offers clues to answer this question. RNAP spends most of its lifetime bound to DNA, either engaged in transcription or non-specifically searching promoters (Endesfelder et al., 2013; Stracy et al., 2015; Ladouceur et al., 2020), and very few RNAP molecules are observed outside the nucleoid region (Bakshi et al., 2012). In rich media, RNAPs form nucleolus-like clusters engaged in rRNA transcription, but cluster assembly is independent of ongoing transcription (Jin et al., 2013; Gaal et al., 2016; Weng et al., 2019). Recently, it has been shown that RNAPs nucleate and form biomolecular condensates by liquid-liquid phase separation (LLPS; Ladouceur et al., 2020).

Ribosome localization, on the other hand, differs considerably among species. In organisms with a high nucleocytoplasmic (NC) ratio (Gray et al., 2019), such as *Caulobacter crescentus*, the ribosomes, and the nucleoid are homogeneously mixed (Bowman et al., 2010; Montero Llopis et al., 2010; Bayas et al., 2018). This facilitates the encounter of the transcriptional and translational machineries, and the occurrence of CTT can be envisaged fairly intuitively in these organisms. Yet, in many other species with low NC ratio, including *E. coli* (Hobot et al., 1985; Azam et al., 2000; Valencia-Burton et al., 2007; Wang et al., 2011; Bakshi et al., 2012; Chai et al., 2014; Cougot et al., 2014; Mohapatra and Weisshaar, 2018; Zhu et al., 2020), *Bacillus subtilis* (Lewis et al., 2000; Mascarenhas et al., 2001), *Bdellovibrio* (Borgnia et al., 2008), and *Pseudomonas putida* (Kim et al., 2019a), ribosomes and nucleoids are strongly segregated. The nucleoid-excluded localization of *E. coli* factors engaged in translation, such as tRNAs (Plochowietz et al., 2016; Volkov et al., 2018) and translation EFs (Chai et al., 2014; Mohapatra et al., 2017; Mustafi and Weisshaar, 2018), also supports that in these species bulk translation takes place in spatial separation from the genetic material and the transcriptional machinery. Hence, in these species, the transcriptional and translational machineries rarely encounter each other and the occurrence of CTT is less intuitive.

Several biophysical forces cause the subcellular nucleoid-vs-ribosome segregation. By avoiding extensive contacts with the inner membrane, the DNA polymer maximizes its number of available conformational states, i.e., the conformational entropy, and by segregating from the nucleoid, the ribosomes optimize their freedom for motion and the translational entropy (Mondal et al., 2011; Bakshi et al., 2014). This segregation is further accentuated by volume exclusion forces mutually exerted by the nucleoid polymer against the bulky polysomes (Mondal et al., 2011; Bakshi et al., 2014; Castellana and Wingreen, 2016). Lastly, electrostatic repulsion forces between negatively charged nucleic acids, in this case, chromosomal DNA and RNA-rich ribosomes (Joyeux, 2016), and phase separation effects (Joyeux, 2018) could also account for the observed antilocalization of nucleoids and ribosomes.

Considering this subcellular segregation of RNAPs and ribosomes, how does CTT occur in these organisms? One possibility is that CTT occurs at the surface of the nucleoid, where both RNAPs and ribosomes may encounter each other (**Figure 4A**). Indeed, it was shown that highly transcribed gene loci migrate to the nucleoid periphery (Stracy et al., 2015; Weng et al., 2019; Yang et al., 2019). The association of several RNAPs into biomolecular condensates at highly transcribed loci (Ladouceur et al., 2020) could cause the exclusion of these condensates from the nucleoid surface by the biophysical forces described above. Supporting this hypothesis, the cotranscriptional association of bulky ribosomes amplifies loci migration to the nucleoid periphery (Yang et al., 2019).

Gene loci encoding membrane proteins are thought to be expressed by coupled transcription-translation-membrane insertion, a mechanism known as transertion (Woldringh, 2002). This implies that certain gene loci migrate from the nucleoid mesh to the vicinity of the inner membrane and become exposed to ribosome-rich regions, where CTT could readily occur (**Figure 4B**). Although the occurrence of transertion still awaits direct experimental validation (Roggiani and Goulian, 2015), the notion is supported by the demonstration that a small population of ribosomes and RNAPs resides in the membrane vicinity (Herskovits and Bibi, 2000; Bakshi et al., 2012), and by the visualization of induction-dependent membrane relocation of several gene loci coding for membrane proteins (Libby et al., 2012; Kannaiah et al., 2019; Yang et al., 2019). Thus, localizing TECs to ribosome-rich regions *via* transertion may very well promote the encountering of RNAPs and ribosomes and facilitate CTT.

Alternatively, it was shown in *E. coli* that, although 70S ribosomes are segregated from the nucleoid, free 30S and 50S subunits can penetrate the nucleoid mesh (Sanamrad et al., 2014; Mohapatra and Weisshaar, 2018; Zhu et al., 2020). This implies

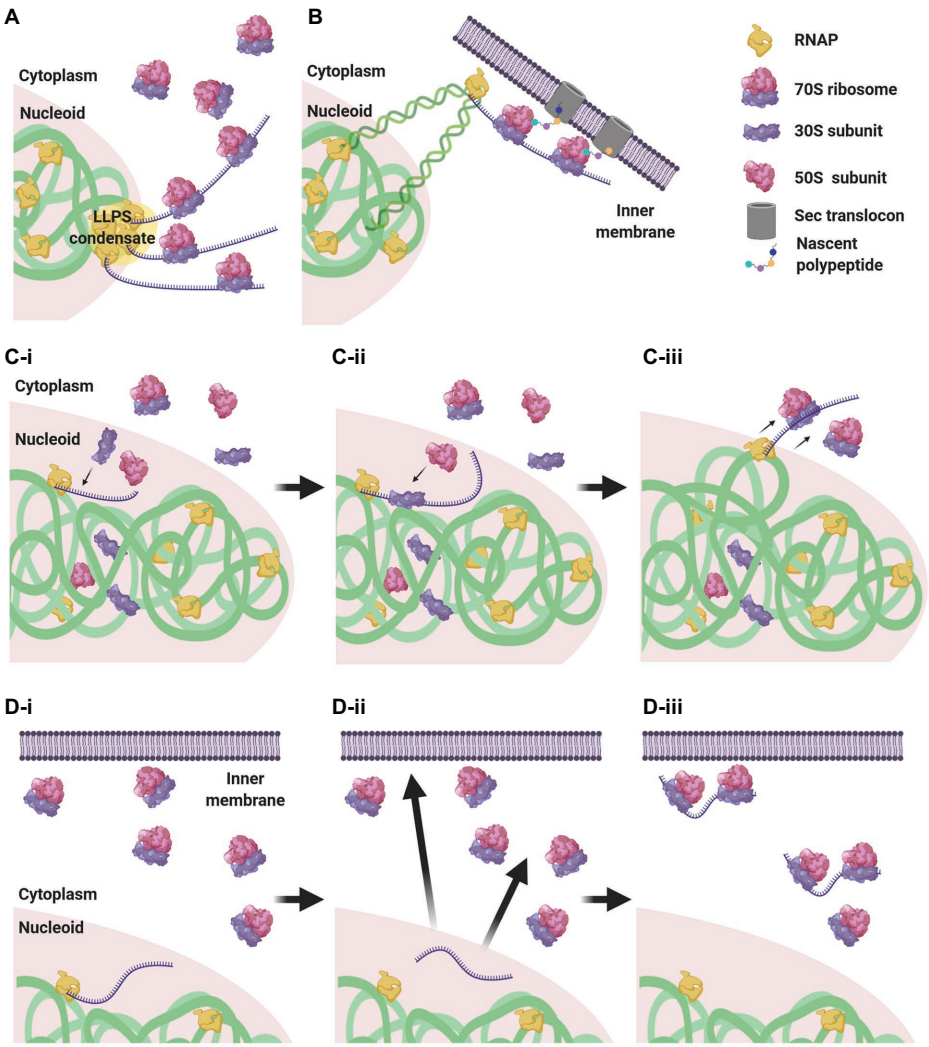


FIGURE 4 | CTT scenarios in low nucleocytoplasmic (NC) ratio species. **(A)** RNAP clusters form condensates by liquid-liquid phase separation (LLPS) that are expelled to the nucleoid periphery, where they encounter ribosomes that engage in CTT. **(B)** *Via* transertion, gene loci encoding membrane proteins emerge from the nucleoid to the inner membrane, where CTT readily occurs. **(C)** Intranucleoidal translation initiation. Free ribosomal subunits penetrate the nucleoid and a 30S subunit associates with the SD sequence of the nascent transcript **(C-i)**. A 50S subunit associates with the pre-initiation complex, forming a 70S ribosome **(C-ii)**. The RNAP·mRNA·ribosome complex is expelled to the nucleoid periphery, where CTT proceeds **(C-iii)**. **(D)** Uncoupled transcription-translation (UTT). Transcription takes place within the nucleoid translation-independently **(D-i)**, and ribosome-free transcripts navigate the cell to their corresponding destination in the cytoplasm or in the membrane **(D-ii)**, where they are locally translated **(D-iii)**.

that canonical translation initiation, which is conducted by a free 30S subunit that recognizes the Shine-Dalgarno (SD) sequence of the nascent transcript, could initiate within the nucleoid in a cotranscriptional manner. Upon assembly of the 70S monosome, the RNAP·nascent mRNA·ribosome complex would be pushed to the surface of the nucleoid by the biophysical forces mentioned above, where additional ribosomes could engage in CTT (**Figure 4C**).

An additional possibility is that transcription and translation are not compulsorily coupled. In this scenario, transcription takes place within the nucleoid, in spatiotemporal separation from translation, and the mRNAs then navigate the cytoplasm to their final destination where they are locally translated (Kannaiah and Amster-Choder, 2014; Irastortza-Olaziregi and Amster-Choder, 2020; **Figure 4D**). We will elaborate on this scenario in section The CTT Dogma has Been Challenged: Towards Uncoupled Transcription-Translation?

THE CTT DOGMA HAS BEEN CHALLENGED: TOWARDS UNCOUPLED TRANSCRIPTION-TRANSLATION?

Coupled transcription-translation is widely accepted by the microbiology community and supported by extensive work. Yet, most knowledge regarding CTT emanates either from *in*

vitro studies or from experiments conducted with only a handful of genes. Importantly, the subcellular segregation of the transcriptional and translational machineries observed in some species raises the possibility that these processes could take place in spatiotemporal separation (see section The Cell Biology of CTT). Indeed, several lines of evidence have recently challenged the classical view that transcription is inherently coupled to translation, and the global occurrence of CTT has been questioned.

The Fredrick lab developed a hammerhead ribozyme-based reporter system that enables measuring and comparing protein synthesis carried out by limited vs. unlimited number of translation rounds (Chen and Fredrick, 2018). They applied this system to study the translation of six adjacent gene-pairs that are cotranscribed in the same operon. In a tight CTT scenario, limiting translation rounds should bring the relative protein amounts for each gene-pair close to 1:1, as protein synthesis of the two co-transcribed genes would presumably be carried out by a leading ribosome physically coupled to RNAP. However, for five out of six gene-pairs, they observed that the relative protein synthesis derived from limited translation rounds of the two genes was not close to 1:1. Rather, the ratio was similar to that measured for unlimited translation rounds. This indicates that the first translational rounds occur independently of transcription, i.e., they are not carried out by a ribosome that is physically bound to RNAP. Importantly, when these experiments were repeated with an RNAP mutant showing reduced transcription elongation rate, which, presumably, facilitates the physical coupling of the leading ribosome with RNAP (see section The Molecular Architecture of CTT), protein production after limited translation rounds was closer to 1:1 for one of the tested gene-pairs (Chen and Fredrick, 2018). These results support a model where the physical coupling of transcription and translation is a stochastic event, which depends on the rates of transcription and translation elongation. These observations indicate that physical association between the leading ribosomes and RNAPs, as well as coordinated elongation by the two machineries, is significantly less common than currently assumed, implying that RNAP often transcribes without a linked ribosome (Chen and Fredrick, 2020). Supporting this notion, pseudouridimycin, which stalls TECs, notably increased the percentage of ribosomes that are physically coupled to RNAP in *M. pneumoniae* (O'Reilly et al., 2020). This indicates that, in the absence of transcription-halting antibiotics, the majority of those ribosomes are engaged in CTT, but that they do not translate in physical association to RNAP.

Similar conclusions emanated from studying the relationship between termination efficiency (TE) at intrinsic terminators and their intragenic position (Li et al., 2015). TE of terminators located in the first 100 nt of ORFs was close to 100%, and it gradually decreased as the distance from the start codon increased. This position-dependent loss of TE was explained by the fact that ribosomes follow and catch up RNAPs closely enough to prevent the formation of terminators (Li et al., 2015). Indirectly, the dependence of TE on the distance from the start codon implies that transcription of the 5'-proximal coding sequences occurs translation-independently (see section The Molecular Architecture of CTT).

Whereas these studies argue against the idea of physical CTT, they are not incompatible with the existence of RNAP-ribosome complexes remotely connected by nascent mRNAs. Yet, further evidence supports the idea that transcription and translation in some cases are completely uncoupled. It is accepted that translation increases mRNA stability by ribosome shielding against ribonucleolytic attack (Iost and Dreyfus, 1995; Makarova et al., 1995; Deana and Belasco, 2005). A reassessment of RNA decay patterns unraveled that, contrary to the widespread idea that RNA decay is exponential, two-thirds of the analyzed transcripts followed a biphasic degradation pattern, with a very steep decay at short post-transcription times followed by an exponential decay at longer times (Deneke et al., 2013). These results suggest that most transcripts spend a minor fraction of their lifetime in a ribosome-free form, where they are highly vulnerable to ribonucleases until ribosomes engage in translation and slowdown mRNA decay. At the same time, similar to what happens in eukaryotic cells, other RNA-binding proteins could be responsible for the protection of the transcripts till they are translated. Whatever the reason is, these results imply that transcription is not tightly coupled to translation for most genes (Deneke et al., 2013).

Furthermore, previous work from our lab and others showed that transcripts can localize to sites overlapping with the localization of their encoded proteins in the cytoplasm, the membrane, or the poles of *E. coli* cells (Nevo-Dinur et al., 2011; Moffitt et al., 2016). Very importantly, these localization patterns were preserved when the translation of the tracked transcripts was inhibited by antibiotics, translational roadblocks, or mutations (Nevo-Dinur et al., 2011). Similarly, a transcript encoding a short membrane protein localized to the membrane even when the SD sequence of the mRNA was deleted (Steinberg et al., 2020). Likewise, translation-independent RNA localization has been observed in cyanobacteria, where transcripts encoding photosystem components localized to thylakoid membranes in the presence of puromycin concentrations that inhibit translation and detach ribosomes from mRNAs (Mahbub et al., 2020). These observations imply that bacterial mRNAs can skip tight CTT, navigate the cytosol as ribosome-free transcripts, and undergo local translation once they reach their corresponding destination (Nevo-Dinur et al., 2011; Kannaiah and Amster-Choder, 2014; Irastortza-Olaziregi and Amster-Choder, 2020). Another recent publication from our lab, which reports the distribution of *E. coli* RNAs between the membrane, cytoplasm, and poles by combining cell fractionation with deep-sequencing, showed that a significant fraction of the *E. coli* transcriptome localizes in a translation-independent manner and challenged the idea that CTT is a general mechanism for gene expression (Kannaiah et al., 2019). Collectively, this evidence supports the notion that CTT is not as predominant as currently assumed, and that spatiotemporally uncoupled transcription-translation (UTT) could be responsible for substantial gene expression.

UTT implies that processive transcription can occur independently of ribosomes that are coupled to RNAP by

physical interaction, *via* protein factors or through mRNAs. For *E. coli* RNAPs, translation-independent transcription has been attributed only to TECs engaged in rRNA transcription, which are modified by the antitermination complex and show fast transcription elongation rates that outpace Rho and avoid termination (Squires and Zaporojets, 2000; Paul et al., 2004). For mRNA-transcribing TECs, the dogma that prevailed was that a coupled ribosome is required for transcription processivity (see section Coordination of CTT). Then again, early evidence already indicated that transcription can be tuned independently of translation. Although mRNA transcription-translation rates change and equalize each other at different growth rates, rRNA transcription rates also vary according to growth rates (Vogel and Jensen, 1994b), suggesting that ribosome-independent mechanisms exist in bacteria for determining transcription elongation rates. As discussed in section Coordination of CTT, the application of sublethal concentrations of fusidic acid, which slows down ribosome translocation, did not affect RNAP elongation rates (Zhu et al., 2019). Besides, when ribosomes stalled at proline-rich sequences of *E. coli* cells deleted for EF-P (Elgamal et al., 2016), or when (p)ppGpp-mediated CTT coordination was disrupted under nitrogen starvation (Iyer et al., 2018), RNAP was still able to elongate. In further agreement with ribosome-independent transcription, when backtracked TECs are pushed and reactivated by the leading ribosome, transcription elongation restarts even when translation elongation is inhibited (Stevenson-Jones et al., 2020). Likewise, in *M. pneumoniae* subjected to chloramphenicol treatment, the RNAP·ribosome association is lost (O'Reilly et al., 2020), which reinforces the idea that RNAP can detach from the expressome and transcription elongation proceeds independently to translation.

How can the transcription processivity of TECs that are engaged in mRNA transcription be maintained in the absence of a coupled ribosome? The trafficking of transcription elongation and termination factors can offer a partial explanation for this. NusG recruitment to *E. coli* TECs occurs after substantial transcription and is assisted by a translationally coupled ribosome (see section The Molecular Architecture of CTT). Although Rho is recruited to TECs early after transcription initiation (Mooney et al., 2009a), it still requires an RNAP-associated NusG for inducing the PTC before triggering termination (Epshtein et al., 2010; Hao et al., 2020; Said et al., 2020). Thus, translation-independent TECs could show less affinity for NusG and, consequently, may be less prone to Rho-mediated termination.

Cooperation among RNAPs (reviewed by Le and Wang, 2018) could further facilitate the processivity of ribosome-independent TECs. For instance, similar to a leading ribosome that pushes a backtracked RNAP forward (see section Coordination of CTT), trailing RNAPs facilitate transcription of the leading RNAP over DNA roadblocks and rescue backtracked TECs by physically pushing the preceding RNAP forward (Epshtein et al., 2003; Epshtein and Nudler, 2003). Effects related to DNA supercoiling offer alternative explanations for this transcriptional cooperation. A mathematical model predicts that the torque created by transcription elongation over the DNA double helix pushes and pulls TECs located in close proximity forward without the mediation of any physical contact among each other (Heberling et al., 2016). Recently, a publication from the Jacobs-Wagner lab evidenced that, as long as gene promoters remain induced and multiple RNAPs initiate transcription of the *lacZ* gene, elongation rates of ongoing transcription are maintained by the mutual cancelation of positive and negative DNA supercoiling upstream and downstream the TEC convoy (Kim et al., 2019b). Regardless of the actual underlying mechanism, we suggest that the cooperation between RNAPs could suffice for maintaining the transcription processivity required for UTT.

In vitro studies showed that the *B. subtilis* RNAP transcribes much faster than the *E. coli* RNAP (Artsimovitch et al., 2000) and, notably, a recent publication demonstrated that transcription elongation in *B. subtilis* outpaces ribosome translocation over nascent transcripts (Johnson et al., 2020). Although, as reported for *E. coli*, leading ribosomes could still physically push and rescue stalled or paused *B. subtilis* TECs (see section Coordination of CTT), this "runaway" transcription creates extensive distance between RNAP and the leading ribosome, supporting the idea that transcription and translation are mostly uncoupled in *B. subtilis*. Accordingly, transcription terminators in *B. subtilis* are located only a few nucleotides downstream of stop codons (Johnson et al., 2020), which would otherwise be masked by a ribosome physically associated with RNAP (Li et al., 2015). A bioinformatic exploration using this short distance between stop codons and intrinsic terminators as a proxy for runaway transcription suggested that this mode of gene expression could be a fairly widespread phenomenon in bacteria (Johnson et al., 2020), once again arguing against the universality of CTT (Wang and Artsimovitch, 2020).

MECHANISMS POTENTIALLY ENABLING UTT

Considering the evidence presented in section The CTT Dogma has Been Challenged: Towards Uncoupled Transcription-Translation, it is plausible that CTT may not be as universal as has been assumed for many years. In other words, transcription may occur translation-independently, and mRNAs could be transcribed and translated in spatiotemporal separation. Yet, for UTT to take place, bacteria would need to face three major challenges: (1) Prevent association between the leading ribosome and the nascent mRNA, once the SD sequence emerges from RNAP. (2) Protect the ribosome-free transcript from ribonucleolytic decay in the cytoplasm. (3) Carry out translation-uncoupled transcription without disruption by intrinsic and Rho-dependent terminators.

Dynamics of transcription elongation over 5'-proximal coding sequences are of special interest for UTT. In *E. coli*, RNAP shows pronounced transcriptional pausing at sites overlapping with SD sequences and start codons (Larson et al., 2014), thus opening a temporal window for regulation over the naked 5'-UTRs by RNA-Binding proteins (RBPs) or RNA folding events prior to ribosome association with the transcript. Furthermore, transcription of 5'-proximal coding sequences is

anyhow expected to occur translation-independently (see section The Molecular Architecture of CTT), so why cannot this type of regulation continue to be exerted over the downstream ribosome-naked transcript before translation begins?

Below, we describe several molecular factors and mechanisms that, when acting in cooperation, potentially promote UTT in a way that the three challenges mentioned above would be satisfactorily resolved (**Figure 5**). Evidence supporting the capacity of these factors and mechanisms to act cotranscriptionally on nascent mRNAs is presented. Furthermore, formation of biomolecular condensates by LLPS and their putative implication in promoting UTT are briefly discussed.

RNA-Binding Proteins

RBPs have gained major interest as a consequence of their capability to regulate transcript fate by affecting mRNA translation and stability (Holmqvist and Vogel, 2018; Richards and Belasco, 2019), and emerging evidence suggest they can promote UTT. For instance, *Synechocystis* RBPs Rbp2 and Rbp3 are required for the translation-independent thylakoid membrane localization of transcripts encoding photosystem proteins (Mahbub et al., 2020), implying that they act cotranscriptionally on mRNA targeting before translation initiates. Likewise, Grad-seq experiments performed in *Salmonella* and *E. coli* showed that several RBPs reviewed here co-sediment with RNAP (Smirnov et al., 2016; Hör et al., 2020), which could indicate a cotranscriptional association of RBPs with nascent RNAs. Considering their transcript fate-determining activities, these RBPs emerge as potential candidates for promoting UTT.

Cold Shock Proteins

Cold Shock Proteins (CSPs) belong to an evolutionarily widespread family of small, acidic proteins that were originally discovered as mediators of cold shock response, but it is currently understood that their function exceeds adaptation to cold (reviewed in Ermolenko and Makhatadze, 2002; Horn et al., 2007; Budkina et al., 2020). Notably, the presence of at least one *csp* gene in the cell is essential for viability in *B. subtilis* (Graumann et al., 1997). The *E. coli* genome encodes nine different CSP genes, from which only four, CspA, CspB, CspG, and CspI, are induced by cold shock (Etchegaray and Inouye, 1999; Wang et al., 1999), and two, CspC and CspE, are constitutively expressed at 37°C (Yamanaka et al., 1994; Bae et al., 1999). Of note, the cold shock domain (CSD) in CSPs, confers a capacity to bind single-stranded nucleic acids (Jiang et al., 1997; Lopez et al., 1999; Phadtare and Inouye, 1999) and melt secondary structures within them (Phadtare et al., 2002a; Phadtare and Severinov, 2005; Rennella et al., 2017). Hence, in addition to mediating UTT as described below, CSPs may aid CTT by ensuring the transfer of unfolded mRNA from the nucleoid to ribosomes (El-Sharoud and Graumann, 2007).

Bioinformatical explorations unraveled a sequence-level bias towards U-richness in transcripts encoding integral membrane proteins in *E. coli* (Prilusky and Bibi, 2009) and *Lactococcus lactis* (van Gijtenbeek et al., 2016). Information published by our lab confirmed that the U-richness in membrane-traversing domains is important for their membrane localization, as predicted bioinformatically (Kannaiah et al., 2019). Interestingly, *E. coli* CspC and CspE preferentially bind U-rich artificial transcripts and endogenous membrane mRNAs (Benhalevy et al., 2015). Moreover, CspE overexpression causes the accumulation of ribosome-free transcripts encoding membrane proteins in the cytoplasm and positively affects their translation in the membrane (Benhalevy et al., 2017). This suggests that, *via* CspE mediation, these transcripts avoid CTT until they reach the membrane, where they are locally translated.

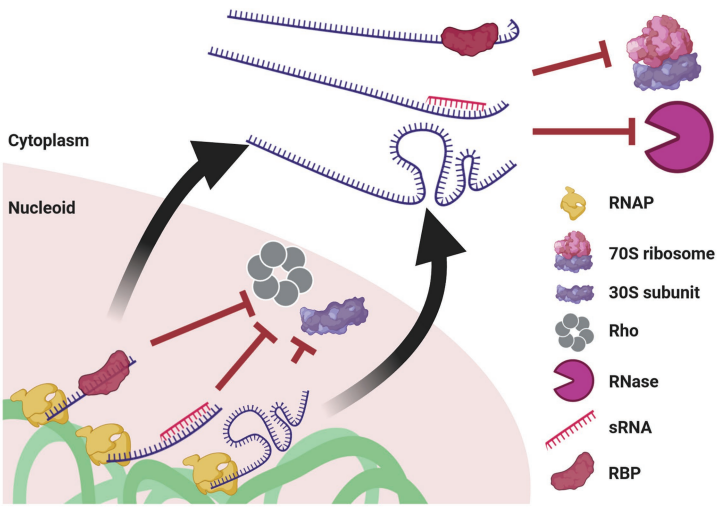


FIGURE 5 | Mechanisms potentially enabling UTT. Cotranscriptional events, such as association with an RBP or with an sRNA, as well as riboswitch formation, prevent transcription termination by Rho and association with the leading ribosome. When ribosome-free transcripts are released to the cytoplasm, these transcript-protecting events counteract the activity of ribonucleases. Of note, although drawn linearly, transcripts supposedly acquire complex secondary and tertiary structures that confer further protection. This protection also prevents mRNA translation until they reach their final destination.

Current evidence suggests that CSPs promotes UTT by protecting nascent transcripts from RNases. For instance, overexpression of CspC and CspE increases the stability of transcripts encoding stress response proteins RpoS and UspA (Phadtare and Inouye, 2001). Whereas CspE levels remain constant through different growth phases, CspC levels increase upon entry into stationary phase, which results in the stabilization of *rpoS* transcripts (Shenhar et al., 2012). These transcript-stabilizing capabilities can be explained by the tendency of CspE to bind poly(A) sequences, counteracting the ribonucleolytic activity of PNPase and RNase E (Feng et al., 2001). Considering that most transcripts undergo polyadenylation (Mohanty and Kushner, 2006), CspE could act as a global molecular shield against the concerted poly(A)-dependent exoribonucleolytic action of PNPase in the 3'-end and the endonucleolytic attack of RNase E in internal cleavage sites of nascent transcripts. Of note, *cspE* mRNA is downregulated in the absence of PNPase (Polissi et al., 2003), suggesting that a regulatory mechanism balances the ribonucleolytic activity of PNPase and the anti-RNase protection by CspE.

For these activities to promote UTT, CSPs should access nascent transcripts as soon as they emerge from RNAP. Indeed, CspE binds nascent RNAs and acts as an antiterminator by melting secondary structures of intrinsic terminators (Hanna and Liu, 1998; Bae et al., 2000; Phadtare et al., 2002a,b; Phadtare and Severinov, 2005). Collectively, these pieces of evidence indicate that CSPs can promote the uncoupling of transcription and translation by counteracting RNase activity over transcripts and promoting translation-independent transcription over intrinsic terminators.

Cold shock proteins show UTT-promoting activities in other species. In *Salmonella*, CspC and CspE bind about 20% of the transcriptome with important regulatory implications for its pathogenicity (Michaux et al., 2017). For example, the *ecnB* mRNA, which is bound by these CSPs, shows lower transcript levels and stability in a *ΔcspCE* background. Levels and stability of this transcript were restored in *ΔcspCE* cells expressing a temperature-sensitive RNase E mutant at the restrictive temperature. Thus, CSPs protect *ecnB* transcripts from RNase E-mediated decay (Michaux et al., 2017). Similarly, CspE binds and stabilizes *yciF* transcripts, conferring *Salmonella* increased resistance to bile salts by impermeabilizing the cell membrane (Ray et al., 2019b). Also, in *B. subtilis*, CspB shows extranucleoidal localization in a transcription-dependent manner (Mascarenhas et al., 2001; Weber et al., 2001), suggesting that CspB binds transcripts upon entry into the cytoplasm.

All in all, CSPs have emerged as RBPs with important roles in gene regulation and, potentially, promoting UTT. Their function resembles that of FRGY2, a CSD-containing protein that binds mRNAs in the nucleous of frog oocytes and protects transcripts from degradation and translation in the cytoplasm until they are released in a regulated manner during oocyte development (Bouvet and Wolffe, 1994). Their multifaceted functions as transcription antiterminators and antiribonucleolytic shields make CSPs promising subjects for future study regarding their putative role as UTT facilitators.

Ribosomal Proteins

In addition to multiple interactions with rRNAs within the ribosome, ribosomal proteins (RPs) perform extraribosomal moonlighting functions as free RBPs, often involved in gene regulation (reviewed in Bhavsar et al., 2010; Aseev and Boni, 2011). Several RPs bind their cognate mRNAs to negatively autoregulate their translation and to keep RP homeostasis. In other cases, RPs play functions that could enable UTT.

One example is the S1 RP in *E. coli*, which binds single-stranded AU-rich sequences located upstream of SD sequences within the 5'-UTRs (Boni et al., 1991; Mogridge and Greenblatt, 1998; Komarova et al., 2002), an activity shown to increase the stability of the *lacZ* mRNA (Komarova et al., 2005), most likely due to the overlap between S1 binding sites and RNase E cleavage sites (Kaberdin and Bläsi, 2006). In agreement with this hypothesis, S1 binds *cspE* and *rpsO* transcripts and protects them against RNase E attack (Delvillani et al., 2011). S1 was further shown to counteract PNPase-mediated mRNA decay (Briani et al., 2008). When overexpressed, S1 binds several mRNAs, including the *pnp* mRNA itself, and increases their stability against PNPase-mediated decay. In line with these results, depletion of S1 leads to the destabilization of these transcripts *in vivo* (Briani et al., 2008). Similar to CspE, S1 binds poly(A) sequences (Kalapos et al., 1997), but this binding does not protect transcripts against PNPase degradation *in vitro* (Feng et al., 2001). Whether S1 counteracts PNPase *in vivo* by promoting CspE-dependent transcript protection (see above) or through an alternative mechanism remains unknown. Additionally, S1 was shown to increase protein secretion that is mediated by the hemolysin signal peptide, and this was accompanied by the stabilization of the corresponding transcripts (Khosa et al., 2018). If S1 can promote UTT, these RNA-protecting activities should be implemented as soon as the nascent transcript emerges from RNAP. In this regard, S1 associates with RNAP and promotes its transcriptional processivity (Sukhodolets and Garges, 2003; Sukhodolets et al., 2006). Additionally, S1 binds RNAP indirectly forming a ribonucleoprotein (RNP) with IsrA sRNA (van Nues et al., 2015). Thus, it is plausible that S1 binds nascent transcripts cotranscriptionally and protects them from ribonucleolytic attack, thus favoring the occurrence of UTT.

Another example is the S4 RP, which negatively autoregulates its translation by binding its cognate cistron within the α-operon and creating a pseudoknot that leads to the entrapment of an inactive translation initiation complex (Spedding and Draper, 1993; Schlax et al., 2001). This autoregulation also represses the translation of several other RPs cotranscribed within the same operon. However, the regulation of the *rpoA* gene, which encodes the α-subunit of RNAP and is also cotranscribed within the α-operon, is not subjected to this negative regulation (Thomas et al., 1987). Although the mechanism of this exclusion remains unknown, it implies that the α-operon can be subjected to partial disruption of CTT; i.e., whereas the α-subunit of RNAP is cotranscriptionally translated, the genes encoding ribosomal proteins are cotranscribed but translationally repressed by S4. To exert such CTT disruption, S4 would need to act cotranscriptionally. Indeed, S4 can bind RNAP and antiterminate

Rho-dependent terminators (Torres et al., 2001), thus protecting ribosome-free nascent transcripts from PTT. Therefore, S4 plays a double role in promoting UTT: disruption of the transcription-translation coupling in the α-operon and preventing Rho-mediated termination of TECs that are not physically coupled to a translating ribosome.

The L4 RP also shows multiple activities that could promote UTT. L4 binds the regulatory CTD of RNase E and inhibits its activity, leading to stabilization of a subset of stress-related transcripts crucial for cell survival (Singh et al., 2009). Alleviation of the ribonucleolytic pressure on these mRNAs could allow a less tight CTT regime for these transcripts. Among the L4-stabilized transcripts is that of the triptophanase (*tna*) operon, which is subjected to an additional layer of regulation by L4 that is degradosome-independent. Specifically, L4 overexpression increases the stability of the *tnaCAB* transcript but causes translational repression of the TnaA protein by binding to the spacer between *tnaC-tnaA* (Singh et al., 2020). This could lead to the partial disruption of CTT in this operon, i.e., *tnaC* can be expressed by CTT, whereas *tnaA* translation is repressed despite the increase in its mRNA levels. Interestingly, this translational repression of *tnaA* takes place at early stationary phase to prevent the degradation of tryptophan, which is required for long-term survival through deep stationary phase (Singh et al., 2020), highlighting that CTT disruptions can have important physiological implications. Importantly, L4 binds its cognate transcript to attenuate its transcription (Lindahl et al., 1983). Thus, L4 could act cotranscriptionally as an antitranslation agent also over other nascent transcripts. Further investigation of extraribosomal functions of RPs should lead to a better understanding of these activities that are potentially involved in UTT.

Hfq

Besides its highly studied role as an sRNA-mRNA matchmaker (see below), Hfq exerts post-transcriptional regulation in *E. coli* by directly binding to transcripts and affecting their fate (reviewed in Kavita et al., 2018). Hfq was shown to bind the 5'-UTRs of its cognate transcript (Večerek et al., 2005) and of *cirA* (Salvail et al., 2013) and *mutS* mRNAs (Chen and Gottesman, 2017), as well as the ribosome binding site (RBS) of the Tn10 transposase mRNA (Ellis et al., 2015). In all these cases Hfq binding leads to translational repression of the target transcripts. These activities can be explained by the observation that in enterohemorrhagic *E. coli*, Hfq shows a preference for binding ARN triplets located in the proximity of RBSs (Tree et al., 2014), which could preclude translation initiation. Likewise, translational repression by Hfq plays a central role in the catabolic repression of the opportunistic pathogen *Pseudomonas aeruginosa*. With the assistance of the catabolic repression control protein Crc, Hfq binds specific A-rich sequences in the proximity of translation initiation regions to suppress translation of catabolite repressed genes (Sonnleitner and Bläsi, 2014; Pei et al., 2019).

Hfq was also shown to preferentially bind intrinsic terminator sequences in *Salmonella* (Holmqvist et al., 2016). Correspondingly, Hfq binds PAP I and promotes the synthesis of poly(A) tails at intrinsic terminators in *E. coli* (Hajnsdorf and Régnier, 2000; Le Derout et al., 2003; Mohanty et al., 2004). Analogously to CspE (see above), Hfq binds to poly(A) sequences overlapping with RNase E cleavage sites and confers protection against exoribonucleolytic decay by PNPase and RNase II (Folichon et al., 2003; Moll et al., 2003; Zhang et al., 2003). Collectively, this evidence indicates that Hfq binds intrinsic termination sites in order to promote polyadenylation and, by binding to these poly(A) sequences, protects 3'-UTRs and upstream sequences from ribonucleolytic decay.

Importantly, Hfq interacts with RNAP to promote transcription (Sukhodolets and Garges, 2003). Furthermore, Hfq shares topological similarities with YaeO, the only so-far discovered protein that binds and inhibits Rho in *E. coli* (Pichoff et al., 1998; Gutiérrez et al., 2007) and *Vibrio cholerae* (Pal et al., 2019). Accordingly, Hfq suppresses Rho-dependent termination by simultaneously binding Rho and AU-rich sequences located upstream *rut* sites (Rabhi et al., 2011). Recently, it has been shown that Hfq pervasively binds nascent transcripts in *E. coli* (Persson et al., 2013; Sedlyarova et al., 2016) and *P. aeruginosa* (Kambara et al., 2018; Gebhardt et al., 2020). The helix-like localization of Hfq observed in *E. coli* under certain growth conditions (Taghbalout et al., 2014; Malabirade et al., 2017; Kannaiah et al., 2019) resembles the helical-ellipsoidal conformation of the nucleoid (Fisher et al., 2013), reinforcing the idea that Hfq could indirectly associate with the chromosome by binding nascent transcripts. Similar to S1, Hfq forms an RNP with IsrA sRNA that associates with RNAP (van Nues et al., 2015). Therefore, Hfq emerges as a potential UTT-promoting factor by antiterminating ribosome-free transcription that could otherwise be terminated by Rho. The fact that Hfq can interact with nascent transcripts implies that the antitranslation and antiribonuclease roles of Hfq described here could come into play cotranscriptionally to disrupt CTT and protect nascent mRNAs from RNases.

CsrA/RsmA

Although initially discovered as a regulator of glycogen biosynthesis upon entry into stationary phase (Romeo et al., 1993), CsrA has emerged as a global post-transcriptional regulator implicated in multiple cellular functions (reviewed in Vakulskas et al., 2015; Romeo and Babitzke, 2018). CsrA has been shown to bind hundreds of transcripts in *E. coli* (Edwards et al., 2011; Potts et al., 2017), *Campylobacter* (Dugar et al., 2016), *Salmonella* (Holmqvist et al., 2016), and *Legionella* (Sahr et al., 2017) affecting their post-transcriptional fate and potentially promoting UTT.

The main regulatory activity of CsrA is *via* translational repression (Dugar et al., 2016; Potts et al., 2017). For example, CsrA binds the SD sequence of the *hfq* mRNA and inhibits its translation (Baker et al., 2007). Similarly, CsrA binds at two sites in the 5'-UTR, including the SD sequence, of transcripts encoding the transcriptional regulator NhaR, which responses to high sodium concentrations and alkaline pH (Pannuri et al., 2012). Likewise, CsrA inhibits translation of the iron storage *dsp* mRNA by binding the 5'-UTR region of the transcript (Potts et al., 2017; Pourciau et al., 2019).

In enteropathogenic *E. coli* (EPEC), CsrA represses translation of the virulence effector NleA (Katsowich et al., 2017). Upon contact with the host, EPEC injects effectors to the host cell through a type III secretion system with the assistance of the effector-bound chaperone CesT. After releasing the effector, free CesT binds and inhibits CsrA, leading to derepression of NleA and its subsequent translation and translocation to the host cell (Katsowich et al., 2017). The CsrA homolog RsmA binds the 5'-UTR of *psl* mRNA in charge of the biosynthesis of a structural polysaccharide in *P. aeruginosa* biofilms and inhibits its translation by refolding the transcript structure so that the SD sequence is not accessible to ribosomes (Irie et al., 2010). In all examples discussed here, translation inhibition was not accompanied by decreased transcript stability. Consequently, this can lead to the accumulation of ribosome-free transcripts in the cytoplasm, resembling the phenomenon observed upon CspE and S1 overexpression (see above). Interestingly, the CsrA/RsmA-regulated transcripts discussed here encode proteins responding to environmental stimuli. Thus, it is tempting to speculate that CsrA/RsmA promote UTT and the accumulation of untranslated pools of mRNAs whose expression can be rapidly derepressed upon environmental challenges without transcriptional delay.

Furthermore, CsrA can also act as an anti-RNase shield (Esquerré et al., 2016; Potts et al., 2017). For instance, CsrA was shown to be essential for *E. coli* motility by regulating the *fhl* operon, the master operon for flagellum biosynthesis, by increasing the stability of *fhlDC* transcripts (Wei et al., 2001), and this is achieved by protecting these mRNAs from RNase E cleavage (Yakhnin et al., 2013). Likewise, CsrA stabilizes the *Legionella* iron uptake regulator *fur* mRNA by binding a specific site in the proximity of a potential RNase E site (Sahr et al., 2017). CsrA and RsmA were also shown to increase the stability of the STM3611 transcript in *Salmonella* (Jonas et al., 2010) and of the *hrpG* transcript, the master regulator of T3SS genes in *Xanthomonas*, respectively (Andrade et al., 2014).

CsrA also displays transcription-related activities, e.g., binding to the *gap* operon transcript in *Legionella* cotranscriptionally to counteract Rho-dependent transcription termination (Sahr et al., 2017) and to RNAP as an RNP with IsrA sRNA in *E. coli* (van Nues et al., 2015). Furthermore, in *P. aeruginosa*, RsmA was shown to bind over 500 nascent transcripts cotranscriptionally (Gebhardt et al., 2020). Collectively, these publications indicate that CsrA/RsmA are major gene regulators that act over nascent transcripts. Thus, it is plausible that the antitranslation, anti-RNase, and antitermination activities of CsrA/RsmA favor the occurrence of UTT. Further investigating the cotranscriptional activities of CsrA/RsmA should shed light on the detailed involvement of CsrA/RsmA in promoting UTT.

ProQ

ProQ, a FinO-domain protein that acts as an sRNA-mRNA matchmaker (Smirnov et al., 2017; Westermann et al., 2019; Melamed et al., 2020), has recently emerged as an important RBP, which binds a substantial fraction of the *E. coli* and *Salmonella* transcriptomes (Smirnov et al., 2016; Holmqvist et al., 2018), suggesting that it may be involved in UTT implementation.

ProQ recognizes structural features present in sRNAs and 3'-UTRs of mRNAs and increases their stability upon binding (Smirnov et al., 2016; Holmqvist et al., 2018; Bauriedl et al., 2020; Stein et al., 2020). Specifically, ProQ was shown to stabilize *cspE* mRNA by binding its 3'-UTR and preventing RNase II-mediated exoribonucleolytic attack (Holmqvist et al., 2018). Stability of *cspC*, *cspD*, and *ompD* was also reduced in *ΔproQ* background (Holmqvist et al., 2018), but whether ProQ stabilizes these transcripts by antiribonuclease protection awaits experimental confirmation. Furthermore, about a third of ProQ binding events take place in sites overlapping with RNase E cleavage sites (Chao et al., 2017; Holmqvist et al., 2018). Collectively, these anti-RNase activities resemble those displayed by FinO, which recognizes and binds a similar structural feature of its only target FinP antisense RNA and exerts anti-RNase E protection (Jerome et al., 1999; Arthur et al., 2011). Importantly, ProQ associates with RNAP *via* an RNP formed with IsrA sRNA (van Nues et al., 2015). Thus, it is plausible that ProQ accesses and binds nascent transcripts cotranscriptionally, promoting UTT by protecting untranslated transcripts against ribonucleolytic decay.

Nus Factors

Nus factors regulate transcription elongation by affecting RNAP pausing and promoting transcription termination/antitermination (Santangelo and Artsimovitch, 2011; Sen et al., 2014). NusA is an essential component of the antitermination complex that, together with NusB, NusE, NusG, and several ribosomal proteins, enhances rRNA transcription rates (Squires and Zaporojets, 2000; Paul et al., 2004). Besides, NusA promotes RNAP pausing, which favors the formation, stability, and efficiency of intrinsic terminators (Farnham et al., 1982; Schmidt and Chamberlin, 1987; Toulokhonov et al., 2001). As discussed in section Coordination of CTT, NusA also mediates transcription termination by Rho (Epshtein et al., 2010; Hao et al., 2020; Said et al., 2020). Yet, NusA could possibly counteract Rho-dependent termination in certain instances. Specifically, NusA mutants with increased affinity for binding NusA utilization (*nut*) sites decreased Rho-dependent termination at specific cases where *nut* and *rut* sites overlap (Qayyum et al., 2016). Thus, regarding UTT, NusA could facilitate translation-independent transcription of mRNAs by interfering with certain Rho-dependent termination events on nascent ribosome-free transcripts.

Other Nus factors are involved in processive antitermination (PA) mechanisms. Opposite to dedicated antiterminators, PA factors associate with and modify TECs to promote transcriptional readthrough over multiple transcription terminators distally located in the operons under their regulation (Goodson and Winkler, 2018). For example, the NusG paralogue LoaP regulates transcription through termination sites located within two antibiotic biosynthesis operons in *Firmicutes*, *Actinobacteria*, and *Spirochaetes* (Goodson et al., 2017). Deletion of *loaP* led to a reduction in transcript levels of LoaP regulons. LoaP is

thought to processively antiterminate intrinsic terminators, an activity that requires the 5' leader sequence of the transcripts under its regulation (Goodson et al., 2017). All in all, the PA activity of such Nus paralogs on their specific target operons could favor UTT by alleviating the need for a translationally coupled ribosome for counteracting intrinsic and Rho-dependent terminators.

Cis-Acting RNA Elements

Riboswitches are regulatory elements located in the 5'-UTR of mRNAs that are comprised of two modules: a structurally complex aptamer that binds a ligand and an expression platform that is regulated by the aptamer structural folding. Ligand binding causes structural refolding of the aptamer, which affects the expression of the downstream expression platform in an ON/OFF manner (Sherwood and Henkin, 2016). Some riboswitches act by translationally repressing the ORF under their regulation (Breaker, 2018). For example, in cobalamin riboswitches, the SD sequence is sequestered within the aptamer structure (Johnson et al., 2012), which becomes accessible for ribosomes only upon ligand binding. In thiamin pyrophosphate riboswitches, on the other hand, an anti-SD sequence within the aptamer structure anneals to and folds over the SD sequence located in the expression platform and inhibits translation initiation (Winkler et al., 2002). Again, ligand binding causes the refolding of the aptamer and liberates the SD sequence for ribosome binding. Similar to riboswitches, RNA thermometers are 5'-UTR elements that undergo structural refolding and affect downstream gene expression, but their refolding is caused by changes in temperature rather than ligand binding (Kortmann and Narberhaus, 2012). For example, the transcription of the cold-induced *cspA* gene takes place at all temperatures, but the *cspA* transcript is highly unstable at 37°C due to RNase E-mediated decay (Fang et al., 1997). Upon cold shock, the *cspA* transcript undergoes substantial refolding and is stabilized in a more RNase-resistant folding that allows translation of the protein (Giuliodori et al., 2010). Oppositely, the translation of *rpoH* transcript, coding for the heat shock sigma factor, is repressed at physiological temperatures by a secondary structure that sequesters the SD sequence. Upon temperature upshift, this structure refolds in a way that exposes the SD sequence for translation initiation (Morita et al., 1999). Likewise, several virulence factors are regulated by RNA thermometers that enable translation at 37°C upon entry in warm-blooded mammalian hosts (Johansson, 2009). Importantly, riboswitch structures are folded cotranscriptionally (Mironov et al., 2002; Frieda and Block, 2012; Watters et al., 2016; Uhm et al., 2017; Ray et al., 2019a). Considering that RNAP pauses at SD sequences and start codons (Larson et al., 2014), it is likely that riboswitch regulation comes into action shortly after they emerge from RNAP and before the translation initiation. Hence, it is reasonable to argue that riboswitches and RNA thermometers can cotranscriptionally block translation and protect transcripts from ribonucleolytic attack, which would promote UTT. Such inert transcripts would be activated upon environmental changes that induce their translation.

Similarly, *cis*-acting RNA elements act as translational repressors of type I toxin-antitoxin (TA) systems (Masachis and Darfeuille, 2018). In many type I TA systems the primary toxin transcript is translationally inert, and this is achieved by sequestering the SD sequence in a secondary structure formed with an anti-SD sequence located upstream in the transcript (Gultyaev et al., 1997; Darfeuille et al., 2007; Shokeen et al., 2008; Kristiansen et al., 2016; Wen et al., 2017). In other cases, the translation of the toxin is repressed by structures formed by interactions between the 5'- and 3'-ends of the full-length mRNA (Thisted et al., 1995; Franch and Gerdes, 1996; Gultyaev et al., 1997). The high structural complexity of toxin mRNAs is also thought to prevent the interaction of the nascent toxin transcript with the template DNA, which avoids the formation of deleterious R-loops, and to confer increased antiribonuclease resistance (Masachis and Darfeuille, 2018). Thus, structural features of toxin mRNAs ensure their translationally-inert transcription and protection against RNases until downstream activation events trigger post-transcriptional translation of these transcripts.

Some *cis*-acting RNA elements could facilitate UTT in a more indirect manner. For example, inhibitory RNA aptamers (iRAPs) interact with RNAP and facilitate Rho-dependent termination (Sedlyarova et al., 2017). Interestingly, many iRAPs map to the antisense strand and curb antisense transcription, which reduces transcriptional interference and favors sense transcription (Sedlyarova et al., 2017; Magán et al., 2019). Considering translation-independent TECs are less likely to reinitiate transcription after clashing with an antisense TEC (Hoffmann et al., 2019), diminishing antisense transcription by antisense iRAPs can alleviate the requirement of a leading ribosome that supports the TEC conducting sense transcription.

Other *cis*-acting RNA elements not only have the potential to attenuate CTT but enforce its disruption. For instance, the *Salmonella* virulence *mgtCBR* operon harbors a leader region that acts as a Rho-antagonizing RNA element (RARE) and a *rut* site that is necessary for Rho-mediated transcription termination of this operon (Sevostyanova and Groisman, 2015). The RARE counteracts termination by trapping Rho in a termination-defective conformation. Importantly, translation of the *mgtCBR* transcript sequesters the RARE in a stem-loop (Sevostyanova and Groisman, 2015). Thus, the only manner to express this operon is by UTT, in a way that the RARE inhibits Rho-mediated termination and allows translation-uncoupled transcription of the full mRNA, which would be translated post-transcriptionally. The *corA* operon of *Salmonella* is subjected to a very similar regulation and its leader region is highly conserved in enterobacteria (Kriner and Groisman, 2015), so such strictly UTT-dependent gene expression could be relatively widespread in these species.

Lastly, besides elements that act against individual termination sites, other *cis*-acting RNA sequences are involved in PA (see above Goodson and Winkler, 2018). For example, in *B. subtilis* the *eps* operon, encoding exopolysaccharide biosynthesis proteins, is regulated by an *eps*-associated RNA (EAR) sequence, located in the intergenic region between the second and third genes of the operon (Irnov and Winkler, 2010). The EAR sequence

is necessary for transcription of the entire operon, as it antiterminates several intrinsic terminators located in distal sites within the operon. The EAR sequence is also able to antiterminate heterologous terminators, supporting the hypothesis that it acts as a PA (Irnov and Winkler, 2010). Such *cis*-acting RNA elements with PA activities could promote UTT in the operons that they regulate since they would alleviate the requirement of a cotranscriptionally coupled ribosome to counteract terminators.

Trans-Acting RNA Elements

sRNAs are the archetypical example of *trans*-acting RNAs in bacterial gene regulation. They are typically 50–300 nt long transcripts that determine transcript fate by imperfect base-pairing with target mRNAs (Wagner and Romby, 2015). This process is often facilitated by the mRNA-sRNA matchmakers Hfq (Updegrove et al., 2016) and ProQ (Holmqvist et al., 2020). More recently, matchmaking activity has been reported for CsrA (Müller et al., 2019). sRNAs can affect the stability and/or translation of target mRNAs either positively or negatively, and these activities can potentially promote UTT.

In *Salmonella*, the glucose-responsive sRNA SgrS stabilizes the *pdlB-yigL* bicistronic transcript by preventing RNase E-mediated cleavage (Papenfort et al., 2013). Similarly, the sRNA RydC basepairs with the 5'-UTR of the *cfa* mRNA and stabilizes the longer isoform of this transcript by counteracting RNase E attack (Fröhlich et al., 2013). In *B. subtilis*, the sRNA RoxS binds the 5'-end of the *yflS* mRNA and prevents exoribonucleolytic decay by RNase J1 (Durand et al., 2017). Besides anti-RNase protection, the *B. subtilis* sRNA SR1 blocks translation binds in the 5'-UTR of its target *ahrC* mRNA and blocks its translation (Heidrich et al., 2006, 2007). Likewise, in *Legionella pneumophila*, the competence operon is translationally repressed by the sRNA RocR (Attaiech et al., 2016). sRNAs can further promote translational repression by recruiting Hfq to binding sites that prevent ribosomes association with SD sequences of their mRNA targets (Desnoyers and Massé, 2012; Azam and Vanderpool, 2018).

As sRNAs are relatively short and not translated by ribosomes they can penetrate the nucleoid mesh (Sheng et al., 2017) and putatively engage in cotranscriptional processes. Indeed, it was recently discovered that the sRNAs DsrA, ArcZ, and RprA, although induced by different stresses, bind the 5'-UTR of nascent *rpoS* transcripts to suppress Rho-dependent termination and allow expression of the stationary phase sigma factor (Sedlyarova et al., 2016). The expression of Rho itself is subjected to similar regulation, as the sRNA SraL basepairs with the 5'-UTR of the *rho* transcript to antiterminate its transcription (Silva et al., 2019). Thus, sRNAs could promote UTT by counteracting pervasive Rho-mediated transcription antitermination of ribosome-free mRNAs. The cotranscriptional association of sRNAs with mRNAs also suggests that they could exert their antitranslation and antiribonucleolytic regulation over nascent transcripts, facilitating the occurrence of UTT.

Interestingly, the *E. coli* sRNA IsrA associates with RNAP and forms RNPs together with important transcript fate-determining proteins Hfq, S1, CsrA, ProQ, and PNPase (van Nues et al., 2015). The association of these RNPs with RNAP and their corresponding regulatory outputs await further characterization. Yet, besides directly acting as UTT-inducing factors, sRNAs could promote UTT by serving as landing and integration platforms to enable cotranscriptional action of proteins with antitranslation, anti-RNase, and antitermination functions.

Bacterial RNP Bodies

Biomolecular condensates formed by multivalent interactions among proteins and nucleic acids have recently gained special interest. These condensates assemble and dissolve according to LLPS principles and they are involved in important cellular functions, including RNA metabolism (Banani et al., 2017). For example, processing bodies (P-bodies) and stress granules have been shown to sequester poorly translated transcripts for decay or storage in eukaryotes (Decker and Parker, 2012; Khong and Parker, 2020).

Biomolecular condensates show selective permeability and concentrate biomolecules and processes to discreet subcellular regions (Banani et al., 2017). Thus, they are attractive tools for prokaryotes to gain spatial complexity in their cytoplasms, which are generally devoid of membrane-bound organelles. Indeed, the existence of Bacterial RNP bodies (BR-bodies) in bacteria has been recently reported (Al-Husini et al., 2018; Muthunayake et al., 2020). In *C. crescentus*, RNase E assembles into BR-bodies together with the degradosome components and poorly translated RNAs, creating P-body-like condensates engaged in RNA degradation (Al-Husini et al., 2018, 2020). These RNase E condensates associate with the *C. crescentus* nucleoid in the proximity of rDNA loci (Bayas et al., 2018). Interestingly, the *E. coli* RNase E, which unlike in *C. crescentus* localizes to the inner membrane together with the other degradosome components, forms clusters that show RNA-dependent assembly and dynamics (Strahl et al., 2015), resembling the *C. crescentus* BR-bodies, thus suggesting that *E. coli* degradosomes may form condensates by LLPS in the inner membrane. Endoribonucleolytic attack by RNase E is believed to be the initial and the rate-limiting step of RNA degradation, so it is reasonable to argue that proximity to these RNase E condensates may increase the likelihood of a transcript to undergo decay. In agreement with this, *E. coli* membrane-localizing transcripts show a lower average stability, and artificially targeting cytoplasmic mRNAs to the membrane increases their degradation rate (Moffitt et al., 2016). Additionally, detaching RNase E from the membrane sensitized cytoplasmic ribosome-free transcripts (Hadjeras et al., 2019). Collectively, this evidence indicates that RNase activity is highly localized within bacterial cells by the mediation of BR-bodies, and that regions distant to RNA-degrading condensates may not be subjected to intensive ribonucleolytic pressure. This would certainly support UTT, especially in those organisms where the core ribonucleolytic machinery localizes in the membrane, as nascent ribosome-free transcripts would not be reachable for degradosomes and, hence, substantial antiribonucleolytic protection would not be required.

Furthermore, the bacterial cytoplasm shows glass-like properties and its fluidity varies according to the physiological state of the cell, where higher metabolic activity correlates with higher fluidity of the cytoplasm and vice versa (Parry et al., 2014). These biophysical properties may affect the function of biomolecular condensates, i.e., more fluid, liquid-like condensates allow higher motion and enzymatic activities,

whereas more solid, aggregate-like entities specialize in storage (Banani et al., 2017). Bacteria undergo a notable metabolic slowdown upon entry into stationary phase or stress. This leads to the glassification of the cytoplasm (Parry et al., 2014), which could convert BR-bodies into RNA storage condensates rather than RNA processing bodies (Muthunayake et al., 2020). Such RNA-storing BR-bodies could further support UTT by accumulating and protecting untranslated transcripts until favorable conditions permit translation reinitiation.

FINAL REMARKS

Although the individual processes leading to gene expression are subjected to extensive study, less is known about how these processes affect each other and how living cells spatiotemporally arrange the machineries that execute these functions. Yet, it is now acknowledged that these processes regulate each other and that their correct interplay is crucial for cell viability (Dahan et al., 2011). In this regard, prokaryotic CTT remains a paradigmatic example of such crosstalk. A proper CTT regime is of crucial importance for the overall cellular function, and is maintained over different growth conditions. Thus, upon environmental challenges that independently affect transcription or translation, bacteria quickly coordinate the kinetics of the unaffected process accordingly (see section Coordination of CTT). The radically different CTT regimes showed by different organisms (Johnson et al., 2020; O'Reilly et al., 2020; Wang et al., 2020b; Webster et al., 2020) reflect the underlying diversity in gene expression regulatory mechanisms among different species. Hence, further studying the physiological and molecular factors that mediate and regulate CTT is paramount. Of note, a transcriptome-wide assessment of the actual occurrence of CTT remains pending.

As a prokaryotic-specific phenomenon, CTT arises as an interesting target for the development of antimicrobial therapeutics and, for instance, targeting (p)ppGpp-mediated CTT regulation shows great potential. Alarmone-deficient strains are metabolically compromised but viable, so, unlike compounds that target essential targets, targeting (p)ppGpp metabolism may greatly reduce cell viability without exerting the selective pressure that leads to the arousal of antibiotic resistance (Hauryliuk et al., 2015; Syal et al., 2017).

Then again, studying evolutionarily shaped mechanisms that promote UTT (section Mechanisms Potentially Enabling UTT) can offer hints of how to purposely disrupt CTT under conditions that it is essential. Most likely, these mechanisms act in cooperation. In some cases, simultaneous action of UTT-promoting factors may be required, as suggested by the existence of RNA-mediated Hfq-CsrA complexes in *E. coli* (Caillet et al., 2019) and the partial overlap of Hfq and RsmA targetomes in *P. aeruginosa* (Gebhardt et al., 2020). Likewise, overexpression of S1 leads to the accumulation of ribosome-free *cspE* transcripts with increased stability, indicating that this population is subjected to simultaneous anti-translation and anti-RNase protection (Delvillani et al., 2011). RNPs that associate with RNAP, such as those mediated by IsrA sRNA (van Nues et al., 2015), could provide proper ground for such cooperation. Alternatively, UTT could be the result of tandem action of these factors, as in the case of the *L. pneumophila* RocC, a FinO-domain protein that protects the RocR sRNA from ribonucleolytic degradation, so that RocR subsequently represses translation of the competence operon (Attaiech et al., 2016). Similarly, *cspE* and *cspC* mRNAs are among the transcripts stabilized by the L4-mediated inhibition of RNase E (Singh et al., 2009), and *cspE* transcripts are also stabilized by S1 (Briani et al., 2008). Besides, ProQ also protects *csp* transcripts from RNases (Holmqvist et al., 2018). Thus, RPs and ProQ appear to be upstream activators of CSP-mediated UTT.

By now, the notion that prokaryotes have intricate intracellular organization is well-acknowledged (Rudner and Losick, 2010; Govindarajan and Amster-Choder, 2016; Surovtsev and Jacobs-Wagner, 2018). Consequently, prokaryotic UTT implies that the spatiotemporal separation between transcription and translation emerged earlier than the arousal of the eukaryotic cell. This is supported by the partial conservation of Rbp2 and Rbp3, RBPs mediating translation-independent RNA localization in cyanobacteria (Mahbub et al., 2020), in the eukaryotic algae *Chlamydomonas*, where they also bind and localize mRNAs to specific sites where local translation takes place (Uniacke and Zerges, 2009). Hence, it is likely that certain RNA-localizing pathways operating in eukaryotes originated from prokaryotic CTT-disrupting mechanisms. Further investigating these events in prokaryotes may unmask novel mechanisms operating at a suborganellar scale in eukaryotes.

In the past, researchers have drawn a clear line between transcriptional and post-transcriptional regulatory mechanisms in bacteria. In the light of the emerging evidence, we posit that many post-transcriptional regulatory mechanisms come into action on nascent transcripts and that their cotranscriptionality has so far been ignored. As already recognized in eukaryotes (Choder, 2011), factors acting cotranscriptionally can determine transcript fate in prokaryotes as well. Half a century after Miller's famous micrographs (Miller et al., 1970), we foresee that the study of prokaryotic CTT and UTT will doubtlessly produce novel insights that will reshape our understanding of prokaryotic gene expression and subcellular organization.

AUTHOR CONTRIBUTIONS

MIO and OAC wrote the original draft and edited the manuscript. OAC coordinated the work and acquired funding. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the Israel Science Foundation founded by the Israel Academy of Sciences and Humanities (1274/19) and the Deutsch–Israeli Project Cooperation (AM 441/1-1 SO 568/1-1).

ACKNOWLEDGMENTS

We appreciate helpful discussions with members of the OAC group. Figures were generated with BioRender.

REFERENCES

Al-Husini, N., Tomares, D. T., Bitar, O., Childers, W. S., and Schrader, J. M. (2018). α-Proteobacterial RNA degradosomes assemble liquid-liquid phase-separated RNP bodies. *Mol. Cell* 71, 1027–1039.e14. doi: 10.1016/j.molcel.2018.08.003

Al-Husini, N., Tomares, D. T., Pfaffenberger, Z. J., Muthunayake, N. S., Samad, M. A., Zuo, T., et al. (2020). BR-bodies provide selectively permeable condensates that stimulate mRNA decay and prevent release of decay intermediates. *Mol. Cell* 78, 670–682.e8. doi: 10.1016/j.molcel.2020.04.001

Andrade, M. O., Farah, C. S., and Wang, N. (2014). The post-transcriptional regulator rsmA/csrA activates T3SS by stabilizing the 5' UTR of hrpG, the master regulator of hrp/hrc genes, in Xanthomonas. *PLoS Pathog.* 10:e1003945. doi: 10.1371/journal.ppat.1003945

Arthur, D. C., Edwards, R. A., Tsutakawa, S., Tainer, J. A., Frost, L. S., and Glover, J. N. M. (2011). Mapping interactions between the RNA chaperone FinO and its RNA targets. *Nucleic Acids Res.* 39, 4450–4463. doi: 10.1093/nar/gkr025

Artsimovitch, I., Svetlov, V., Anthony, L., Burgess, R. R., and Landick, R. (2000). RNA polymerases from *Bacillus subtilis* and *Escherichia coli* differ in recognition of regulatory signals *in vitro*. *J. Bacteriol.* 182, 6027–6035. doi: 10.1128/JB.182.21.6027-6035.2000

Aseev, L. V., and Boni, I. V. (2011). Extraribosomal functions of bacterial ribosomal proteins. *Mol. Biol.* 45, 739–750. doi: 10.1134/S0026893311050025

Attaiech, L., Boughammoura, A., Brochier-Armanet, C., Allatif, O., Peillard-Fiorente, F., Edwards, R. A., et al. (2016). Silencing of natural transformation by an RNA chaperone and a multitarget small RNA. *Proc. Natl. Acad. Sci. U. S. A.* 113, 8813–8818. doi: 10.1073/pnas.1601626113

Azam, T. A., Hiraga, S., and Ishihama, A. (2000). Two types of localization of the DNA-binding proteins within the *Escherichia coli* nucleoid. *Genes Cells* 5, 613–626. doi: 10.1046/j.1365-2443.2000.00350.x

Azam, M. S., and Vanderpool, C. K. (2018). Translational regulation by bacterial small RNAs *via* an unusual Hfq-dependent mechanism. *Nucleic Acids Res.* 46, 2585–2599. doi: 10.1093/nar/gkx1286

Bae, W., Phadtare, S., Severinov, K., and Inouye, M. (1999). Characterization of *Escherichia coli* cspE, whose product negatively regulates transcription of cspA, the gene for the major cold shock protein. *Mol. Microbiol.* 31, 1429–1441. doi: 10.1046/j.1365-2958.1999.01284.x

Bae, W., Xia, B., Inouye, M., and Severinov, K. (2000). *Escherichia coli* CspA-family RNA chaperones are transcription antiterminators. *Proc. Natl. Acad. Sci. U. S. A.* 97, 7784–7789. doi: 10.1073/pnas.97.14.7784

Baker, C. S., Eöry, L. A., Yakhnin, H., Mercante, J., Romeo, T., and Babitzke, P. (2007). CsrA inhibits translation initiation of *Escherichia coli* Hfq by binding to a single site overlapping the Shine-Dalgarno sequence. *J. Bacteriol.* 189, 5472–5481. doi: 10.1128/JB.00529-07

Bakshi, S., Choi, H., Mondal, J., and Weisshaar, J. C. (2014). Time-dependent effects of transcription- and translation-halting drugs on the spatial distributions of the *Escherichia coli* chromosome and ribosomes. *Mol. Microbiol.* 94, 871–887. doi: 10.1111/mmi.12805

Bakshi, S., Siryaporn, A., Goulian, M., and Weisshaar, J. C. (2012). Superresolution imaging of ribosomes and RNA polymerase in live *Escherichia coli* cells. *Mol. Microbiol.* 85, 21–38. doi: 10.1111/j.1365-2958.2012.08081.x

Banani, S. F., Lee, H. O., Hyman, A. A., and Rosen, M. K. (2017). Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* 18, 285–298. doi: 10.1038/nrm.2017.7

Bauriedl, S., Gerovac, M., Heidrich, N., Bischler, T., Barquist, L., Vogel, J., et al. (2020). The minimal meningococcal ProQ protein has an intrinsic capacity for structure-based global RNA recognition. *Nat. Commun.* 11:2823. doi: 10.1038/s41467-020-16650-6

Bayas, C. A., Wang, J., Lee, M. K., Schrader, J. M., Shapiro, L., and Moerner, W. E. (2018). Spatial organization and dynamics of RNase E and ribosomes in *Caulobacter crescentus*. *Proc. Natl. Acad. Sci. U. S. A.* 115, E3721–E3721. doi: 10.1073/pnas.1721648115

Benhalevy, D., Biran, I., Bochkareva, E. S., Sorek, R., and Bibi, E. (2017). Evidence for a cytoplasmic pool of ribosome-free mRNAs encoding inner membrane proteins in *Escherichia coli*. *PLoS One* 12:e0183862. doi: 10.1371/journal.pone.0183862

Benhalevy, D., Bochkareva, E. S., Biran, I., and Bibi, E. (2015). Model uracil-rich RNAs and membrane protein mRNAs interact specifically with cold shock proteins in *Escherichia coli*. *PLoS One* 10:e0134413. doi: 10.1371/journal.pone.0134413

Bhavsar, R. B., Makley, L. N., and Tsonis, P. A. (2010). The other lives of ribosomal proteins. *Hum. Genomics* 4, 327–344. doi: 10.1186/1479-7364-4-5-327

Bladen, H. A., Byrne, R., Levin, J. G., and Nirenberg, M. W. (1965). An electron microscopic study of a DNA-ribosome complex formed *in vitro*. *J. Mol. Biol.* 11, 78–83. doi: 10.1016/S0022-2836(65)80172-3

Boni, I. V., Lsaeva, D. M., Musychenko, M. L., and Tzareva, N. V. (1991). Ribosome-messenger recognition: MRNA target sites for ribosomal protein S1. *Nucleic Acids Res.* 19, 155–162. doi: 10.1093/nar/19.1.155

Borgnia, M. J., Subramaniam, S., and Milne, J. L. S. (2008). Three-dimensional imaging of the highly bent architecture of *Bdellovibrio* bacteriovorus by using cryo-electron tomography. *J. Bacteriol.* 190, 2588–2596. doi: 10.1128/JB.01538-07

Bouvet, P., and Wolffe, A. P. (1994). A role for transcription and FRGY2 in masking maternal mRNA within Xenopus oocytes. *Cell* 77, 931–941. doi: 10.1016/0092-8674(94)90141-4

Bowman, G. R., Comolli, L. R., Gaietta, G. M., Fero, M., Hong, S. H., Jones, Y., et al. (2010). Caulobacter PopZ forms a polar subdomain dictating sequential changes in pole composition and function. *Mol. Microbiol.* 76, 173–189. doi: 10.1111/j.1365-2958.2010.07088.x

Breaker, R. R. (2018). Riboswitches and translation control. *Cold Spring Harb. Perspect. Biol.* 10:a032797. doi: 10.1101/cshperspect.a032797

Briani, F., Curti, S., Rossi, F., Carzaniga, T., Mauri, P., and Dehò, G. (2008). Polynucleotide phosphorylase hinders mRNA degradation upon ribosomal protein S1 overexpression in *Escherichia coli*. *RNA* 14, 2417–2429. doi: 10.1261/rna.1123908

Budkina, K. S., Zlobin, N. E., Kononova, S. V., Ovchinnikov, L. P., and Babakov, A. V. (2020). Cold shock domain proteins: structure and interaction with nucleic acids. *Biochemistry (Mosc)* 85, S1–S19. doi: 10.1134/S0006297920140011

Burmann, B. M., Knauer, S. H., Sevostyanova, A., Schweimer, K., Mooney, R. A., Landick, R., et al. (2012). An α helix to β barrel domain switch transforms the transcription factor RfaH into a translation factor. *Cell* 150, 291–303. doi: 10.1016/j.cell.2012.05.042

Burmann, B. M., Schweimer, K., Luo, X., Wahl, M. C., Stitt, B. L., Gottesman, M. E., et al. (2010). A NusE:NusG complex links transcription and translation. *Science* 328, 501–504. doi: 10.1126/science.1184953

Byrne, R., Levin, J. G., Bladen, H. A., and Nirenberg, M. W. (1964). The *in vitro* formation of a DNA-ribosome complex. *Proc. Natl. Acad. Sci. U. S. A.* 52, 140–148. doi: 10.1073/pnas.52.1.140

Caillet, J., Baron, B., Boni, I. V., Caillet-Saguy, C., and Hajnsdorf, E. (2019). Identification of protein-protein and ribonucleoprotein complexes containing Hfq. *Sci. Rep.* 9:14054. doi: 10.1038/s41598-019-50562-w

Castellana, M., and Wingreen, N. S. (2016). Spatial organization of bacterial transcription and translation. *Proc. Natl. Acad. Sci. U. S. A.* 113, 9286–9291. doi: 10.1073/pnas.1604995113

Chai, Q., Singh, B., Peisker, K., Metzendorf, N., Ge, X., Dasgupta, S., et al. (2014). Organization of ribosomes and nucleoids in *Escherichia coli* cells during growth and in quiescence. *J. Biol. Chem.* 289, 11342–11352. doi: 10.1074/jbc.M114.557348

Chalissery, J., Muteeb, G., Kalarickal, N. C., Mohan, S., Jisha, V., and Sen, R. (2011). Interaction surface of the transcription terminator rho required to form a complex with the C-terminal domain of the antiterminator NusG. *J. Mol. Biol.* 405, 49–64. doi: 10.1016/j.jmb.2010.10.044

Chao, Y., Li, L., Girodat, D., Förstner, K. U., Said, N., Corcoran, C., et al. (2017). *In vivo* cleavage map illuminates the central role of RNase E in coding and non-coding RNA pathways. *Mol. Cell* 65, 39–51. doi: 10.1016/j.molcel.2016.11.002

Chen, M., and Fredrick, K. (2018). Measures of single-versus multiple-round translation argue against a mechanism to ensure coupling of transcription and translation. *Proc. Natl. Acad. Sci. U. S. A.* 115, 10774–10779. doi: 10.1073/pnas.1812940115

Chen, M., and Fredrick, K. (2020). RNA polymerase's relationship with the ribosome: not so physical, most of the time. *J. Mol. Biol.* 432, 3981–3986. doi: 10.1016/j.jmb.2020.03.018

Chen, J., and Gottesman, S. (2017). Hfq links translation repression to stress-induced mutagenesis in *E. coli*. *Genes Dev.* 31, 1382–1395. doi: 10.1101/gad.302547.117

Chen, H., Shiroguchi, K., Ge, H., and Xie, X. S. (2015). Genome-wide study of mRNA degradation and transcript elongation in *Escherichia coli*. *Mol. Syst. Biol.* 11:781. doi: 10.15252/msb.20145794

Choder, M. (2011). mRNA imprinting: additional level in the regulation of gene expression. *Cell Logist.* 1, 37–40. doi: 10.4161/cl.1.1.14465

Conn, A. B., Diggs, S., Tam, T. K., and Blaha, G. M. (2019). Two old dogs, one new trick: a review of rna polymerase and ribosome interactions during transcription-translation coupling. *Int. J. Mol. Sci.* 20:2595. doi: 10.3390/ijms20102595

Cougot, N., Molza, A. E., Delesques, J., Giudice, E., Cavalier, A., Rolland, J. P., et al. (2014). Visualizing compaction of polysomes in bacteria. *J. Mol. Biol.* 426, 377–388. doi: 10.1016/j.jmb.2013.09.035

Dahan, O., Gingold, H., and Pilpel, Y. (2011). Regulatory mechanisms and networks couple the different phases of gene expression. *Trends Genet.* 27, 316–322. doi: 10.1016/j.tig.2011.05.008

Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Patsalo, V., Okano, H., et al. (2016). Reduction of translating ribosomes enables *Escherichia coli* to maintain elongation rates during slow growth. *Nat. Microbiol.* 2:16231. doi: 10.1038/nmicrobiol.2016.231

Darfeuille, F., Unoson, C., Vogel, J., and Wagner, E. G. H. (2007). An antisense RNA inhibits translation by competing with standby ribosomes. *Mol. Cell* 26, 381–392. doi: 10.1016/j.molcel.2007.04.003

Deana, A., and Belasco, J. G. (2005). Lost in translation: the influence of ribosomes on bacterial mRNA decay. *Genes Dev.* 19, 2526–2533. doi: 10.1101/gad.1348805

Decker, C. J., and Parker, R. (2012). P-bodies and stress granules: possible roles in the control of translation and mRNA degradation. *Cold Spring Harb. Perspect. Biol.* 4:a012286. doi: 10.1101/cshperspect.a012286

Delvillani, F., Papiani, G., Dehò, G., and Briani, F. (2011). S1 ribosomal protein and the interplay between translation and mRNA decay. *Nucleic Acids Res.* 39, 7702–7715. doi: 10.1093/nar

Demo, G., Rasouly, A., Vasilyev, N., Svetlov, V., Loveland, A. B., Diaz-Avalos, R., et al. (2017). Structure of RNA polymerase bound to ribosomal 30S subunit. *elife* 6:e28560. doi: 10.7554/eLife.28560

Deneke, C., Lipowsky, R., and Valleriani, A. (2013). Effect of ribosome shielding on mRNA stability. *Phys. Biol.* 10:46008. doi: 10.1088/1478-3975/10/4/046008

Desnoyers, G., and Massé, E. (2012). Noncanonical repression of translation initiation through small RNA recruitment of the RNA chaperone Hfq. *Genes Dev.* 26, 726–739. doi: 10.1101/gad.182493.111

Dugar, G., Svensson, S. L., Bischler, T., Wäldchen, S., Reinhardt, R., Sauer, M., et al. (2016). The CsrA-FliW network controls polar localization of the dual-function flagellin mRNA in *Campylobacter jejuni*. *Nat. Commun.* 7:11667. doi: 10.1038/ncomms11667

Durand, S., Braun, F., Helfer, A. C., Romby, P., and Condon, C. (2017). sRNA-mediated activation of gene expression by inhibition of 5'-3' exonucleolytic mRNA degradation. *elife* 6:e23602. doi: 10.7554/eLife.23602

Dutta, D., Shatalin, K., Epshtein, V., Gottesman, M. E., and Nudler, E. (2011). Linking RNA polymerase backtracking to genome instability in *E. coli*. *Cell* 146, 533–543. doi: 10.1016/j.cell.2011.07.034

Edwards, A. N., Patterson-Fortin, L. M., Vakulskas, C. A., Mercante, J. W., Potrykus, K., Vinella, D., et al. (2011). Circuitry linking the Csr and stringent response global regulatory systems. *Mol. Microbiol.* 80, 1561–1580. doi: 10.1111/j.1365-2958.2011.07663.x

Elgamal, S., Artsimovitch, I., and Ibba, M. (2016). Maintenance of transcription-translation coupling by elongation factor P. *mBio2* 7:e01373-16. doi: 10.1101/gad.196741.112

Ellis, M. J., Trussler, R. S., and Haniford, D. B. (2015). Hfq binds directly to the ribosome-binding site of IS10 transposase mRNA to inhibit translation. *Mol. Microbiol.* 96, 633–650. doi: 10.1111/mmi.12961

El-Sharoud, W. M., and Graumann, P. L. (2007). Cold shock proteins aid coupling of transcription and translation in bacteria. *Sci. Prog.* 90, 15–27. doi: 10.3184/003685007780440549

Endesfelder, U., Finan, K., Holden, S. J., Cook, P. R., Kapanidis, A. N., and Heilemann, M. (2013). Multiscale spatial organization of RNA polymerase in *Escherichia coli*. *Biophys. J.* 105, 172–181. doi: 10.1016/j.bpj.2013.05.048

Epshtein, V., Dutta, D., Wade, J., and Nudler, E. (2010). An allosteric mechanism of rho-dependent transcription termination. *Nature* 463, 245–249. doi: 10.1038/nature08669

Epshtein, V., and Nudler, E. (2003). Cooperation between RNA polymerase molecules in transcription elongation. *Science* 300, 801–805. doi: 10.1126/science.1083219

Epshtein, V., Toulmé, F., Rachid Rahmouni, A., Borukhov, S., and Nudler, E. (2003). Transcription through the roadblocks: the role of RNA polymerase cooperation. *EMBO J.* 22, 4719–4727. doi: 10.1093/emboj/cdg452

Ermolenko, D. N., and Makhatadze, G. I. (2002). Bacterial cold-shock proteins. *Cell. Mol. Life Sci.* 59, 1902–1913. doi: 10.1007/PL00012513

Esquerré, T., Bouvier, M., Turlan, C., Carpousis, A. J., Girbal, L., and Cocaign-Bousquet, M. (2016). The Csr system regulates genome-wide mRNA stability and transcription and thus gene expression in *Escherichia coli*. *Sci. Rep.* 6:25057. doi: 10.1038/srep25057

Etchegaray, J. P., and Inouye, M. (1999). CspA, CspB, and CspG, major cold shock proteins of *Escherichia coli*, are induced at low temperature under conditions that completely block protein synthesis. *J. Bacteriol.* 181, 1827–1830. doi: 10.1128/jb.181.6.1827-1830.1999

Fan, H., Conn, A. B., Williams, P. B., Diggs, S., Hahm, J., Gamper, H. B., et al. (2017). Transcription–translation coupling: direct interactions of RNA polymerase with ribosomes and ribosomal subunits. *Nucleic Acids Res.* 37, 5578–5588. doi: 10.1093/nar/gkx719

Fang, L., Jiang, W., Bae, W., and Inouye, M. (1997). Promoter-independent cold-shock induction of cspA and its derepression at 37°C by mRNA stabilization. *Mol. Microbiol.* 23, 355–364. doi: 10.1046/j.1365-2958.1997.2351592.x

Farnham, P. J., Greenblatt, J., and Platt, T. (1982). Effects of NusA protein on transcription termination in the tryptophan operon of *Escherichia coli*. *Cell* 29, 945–951. doi: 10.1016/0092-8674(82)90457-3

Feng, Y., Huang, H., Liao, J., and Cohen, S. N. (2001). *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, 31651–31656. doi: 10.1074/jbc.M102855200

Fisher, J. K., Bourniquel, A., Witz, G., Weiner, B., Prentiss, M., and Kleckner, N. (2013). Four-dimensional imaging of *E. coli* nucleoid organization and dynamics in living cells. *Cell* 153, 882–895. doi: 10.1016/j.cell.2013.04.006

Folichon, M., Arluison, V., Pellegrini, O., Huntzinger, E., Régnier, P., and Hajnsdorf, E. (2003). The poly(A) binding protein Hfq protects RNA from RNase E and exoribonucleolytic degradation. *Nucleic Acids Res.* 31, 7302–7310. doi: 10.1093/nar/gkg915

Franch, T., and Gerdes, K. (1996). Programmed cell death in bacteria: translational repression by mRNA end-pairing. *Mol. Microbiol.* 21, 1049–1060. doi: 10.1046/j.1365-2958.1996.771431.x

French, S. L., Santangelo, T. J., Beyer, A. L., and Reeve, J. N. (2007). Transcription and translation are coupled in Archaea. *Mol. Biol. Evol.* 24, 893–895. doi: 10.1093/molbev/msm007

Frieda, K. L., and Block, S. M. (2012). Direct observation of cotranscriptional folding in an adenine riboswitch. *Science* 338, 397–400. doi: 10.1126/science.1225722

Fröhlich, K. S., Papenfort, K., Fekete, A., and Vogel, J. (2013). A small RNA activates CFA synthase by isoform-specific mRNA stabilization. *EMBO J.* 32, 2963–2979. doi: 10.1038/emboj.2013.222

Gaal, T., Bratton, B. P., Sanchez-Vazquez, P., Sliwicki, A., Sliwicki, K., Vegel, A., et al. (2016). Colocalization of distant chromosomal loci in space in *E. coli*: a bacterial nucleolus. *Genes Dev.* 30, 2272–2285. doi: 10.1101/gad.290312.116

Gebhardt, M. J., Kambara, T. K., Ramsey, K. M., and Dove, S. L. (2020). Widespread targeting of nascent transcripts by RsmA in *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. U. S. A.* 117, 10520–10529. doi: 10.1073/pnas.1917587117

Giuliodori, A. M., Di Pietro, F., Marzi, S., Masquida, B., Wagner, R., Romby, P., et al. (2010). The cspA mRNA is a thermosensor that modulates translation of the cold-shock protein CspA. *Mol. Cell* 37, 21–33. doi: 10.1016/j.molcel.2009.11.033

Golomb, M., and Chamberlin, M. (1974). Characterization of T7-specific ribonucleic acid polymerase. IV. Resolution of the major *in vitro* transcripts by gel electrophoresis. *J. Biol. Chem.* 249, 2858–2863.

Goodson, J. R., Klupt, S., Zhang, C., Straight, P., and Winkler, W. C. (2017). LoaP is a broadly conserved antiterminator protein that regulates antibiotic gene clusters in *Bacillus amyloliquefaciens*. *Nat. Microbiol.* 2:17003. doi: 10.1038/nmicrobiol.2017.3

Goodson, J. R., and Winkler, W. C. (2018). Processive antitermination. *Microbiol. Spectr.* 6, 1–18. doi: 10.1128/microbiolspec.rwr-0031-2018

Govindarajan, S., and Amster-Choder, O. (2016). Where are things inside a bacterial cell? *Curr. Opin. Microbiol.* 33, 83–90. doi: 10.1016/j.mib.2016.07.003

Gowrishankar, J., and Harinarayanan, R. (2004). Why is transcription coupled to translation in bacteria? *Mol. Microbiol.* 54, 598–603. doi: 10.1111/j.1365-2958.2004.04289.x

Graumann, P., Wendrich, T. M., Weber, M. H. W., Schröder, K., and Marahiel, M. A. (1997). A family of cold shock proteins in *Bacillus subtilis* is essential for cellular growth and for efficient protein synthesis at optimal and low temperatures. *Mol. Microbiol.* 25, 741–756. doi: 10.1046/j.1365-2958.1997.5121878.x

Gray, W. T., Govers, S. K., Xiang, Y., Parry, B. R., Campos, M., Kim, S., et al. (2019). Nucleoid size scaling and intracellular organization of translation across bacteria. *Cell* 177, 1632–1648.e20. doi: 10.1016/j.cell.2019.05.017

Gultyaev, A. P., Franch, T., and Gerdes, K. (1997). Programmed cell death by hok/sok of plasmid R1: coupled nucleotide covariations reveal a phylogenetically conserved folding pathway in the hok family of mRNAs. *J. Mol. Biol.* 273, 26–37. doi: 10.1006/jmbi.1997.1295

Guo, X., Myasnikov, A. G., Chen, J., Crucifix, C., Papai, G., Takacs, M., et al. (2018). Structural basis for NusA stabilized transcriptional pausing. *Mol. Cell* 69, 816–827.e4. doi: 10.1016/j.molcel.2018.02.008

Guo, Z., and Noller, H. F. (2012). Rotation of the head of the 30S ribosomal subunit during mRNA translocation. *Proc. Natl. Acad. Sci. U. S. A.* 109, 20391–20394. doi: 10.1073/pnas.1218999109

Gutiérrez, P., Kozlov, G., Gabrielli, L., Elias, D., Osborne, M. J., Gallouzi, I. E., et al. (2007). Solution structure of YaeO, a rho-specific inhibitor of transcription termination. *J. Biol. Chem.* 282, 23348–23353. doi: 10.1074/jbc.M702010200

Hadjeras, L., Poljak, L., Bouvier, M., Morin-Ogier, Q., Canal, I., Cocaign-Bousquet, M., et al. (2019). Detachment of the RNA degradosome from the inner membrane of *Escherichia coli* results in a global slowdown of mRNA degradation, proteolysis of RNase E and increased turnover of ribosome-free transcripts. *Mol. Microbiol.* 111, 1715–1731. doi: 10.1111/mmi.14248

Hajnsdorf, E., and Régnier, P. (2000). Host factor Hfq of *Escherichia coli* stimulates elongation of poly(a) tails by poly(A) polymerase I. *Proc. Natl. Acad. Sci. U. S. A.* 97, 1501–1505. doi: 10.1073/pnas.040549897

Hanna, M. M., and Liu, K. (1998). Nascent RNA in transcription complexes interacts with CspE, a small protein in *E. coli* implicated in chromatin condensation. *J. Mol. Biol.* 282, 227–239. doi: 10.1006/jmbi.1998.2005

Hao, Z., Epshtein, V., Kim, K. H., Proshkin, S., Svetlov, V., Kamarthapu, V., et al. (2020). Pre-termination transcription complex: structure and function. *Mol. Cell* 81, 1–12. doi: 10.1016/j.molcel.2020.11.013

Hauryliuk, V., Atkinson, G. C., Murakami, K. S., Tenson, T., and Gerdes, K. (2015). Recent functional insights into the role of (p)ppGpp in bacterial physiology. *Nat. Rev. Microbiol.* 13, 298–309. doi: 10.1038/nrmicro3448

Heberling, T., Davis, L., Gedeon, J., Morgan, C., and Gedeon, T. (2016). A mechanistic model for cooperative behavior of co-transcribing RNA polymerases. *PLoS Comput. Biol.* 12:e1005069. doi: 10.1371/journal.pcbi.1005069

Heidrich, N., Chinali, A., Gerth, U., and Brantl, S. (2006). The small untranslated RNA SR1 from the *Bacillus subtilis* genome is involved in the regulation of arginine catabolism. *Mol. Microbiol.* 62, 520–536. doi: 10.1111/j.1365-2958.2006.05384.x

Heidrich, N., Moll, I., and Brantl, S. (2007). *In vitro* analysis of the interaction between the small RNA SR1 and its primary target ahrC mRNA. *Nucleic Acids Res.* 35, 4331–4346. doi: 10.1093/nar/gkm439

Herskovits, A. A., and Bibi, E. (2000). Association of *Escherichia coli* ribosomes with the inner membrane requires the signal recognition particle receptor but is independent of the signal recognition particle. *Proc. Natl. Acad. Sci. U. S. A.* 97, 4621–4626. doi: 10.1073/pnas.080077197

Hobot, J. A., Villiger, W., Escaig, J., Maeder, M., Ryter, A., and Kellenberger, E. (1985). Shape and fine structure of nucleoids observed on sections of ultrarapidly frozen and cryosubstituted bacteria. *J. Bacteriol.* 162, 960–971. doi: 10.1128/JB.162.3.960-971.1985

Hoffmann, S. A., Hao, N., Shearwin, K. E., and Arndt, K. M. (2019). Characterizing transcriptional interference between converging genes in bacteria. *ACS Synth. Biol.* 8, 466–473. doi: 10.1021/acssynbio.8b00477

Holmqvist, E., Berggren, S., and Rizvanovic, A. (2020). RNA-binding activity and regulatory functions of the emerging sRNA-binding protein ProQ. *Biochim. Biophys. Acta – Gene Regul. Mech.* 1863:194596. doi: 10.1016/j.bbagrm.2020.194596

Holmqvist, E., Li, L., Bischler, T., Barquist, L., and Vogel, J. (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., and Vogel, J. (2018). RNA-binding proteins in bacteria. *Nat. Rev. Microbiol.* 16, 601–615. doi: 10.1038/s41579-018-0049-5

Holmqvist, E., Wright, P. R., Li, L., Bischler, T., Barquist, L., Reinhardt, R., 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

Hör, J., Di Giorgio, S., Gerovac, M., Venturini, E., Förstner, K. U., and Vogel, J. (2020). Grad-seq shines light on unrecognized RNA and protein complexes in the model bacterium *Escherichia coli*. *Nucleic Acids Res.* 48, 9301–9319. doi: 10.1093/nar/gkaa676

Horn, G., Hofweber, R., Kremer, W., and Kalbitzer, H. R. (2007). Structure and function of bacterial cold shock proteins. *Cell. Mol. Life Sci.* 64, 1457–1470. doi: 10.1007/s00018-007-6388-4

Iost, I., and Dreyfus, M. (1995). The stability of *Escherichia coli* lacZ mRNA depends upon the simultaneity of its synthesis and translation. *EMBO J.* 14, 3252–3261. doi: 10.1002/j.1460-2075.1995.tb07328.x

Irastortza-Olaziregi, M., and Amster-Choder, O. (2020). RNA localization in prokaryotes: where, when, how, and why. *Wiley Interdiscip. Rev. RNA*, e1615. doi: 10.1002/wrna.1615

Irie, Y., Starkey, M., Edwards, A. N., Wozniak, D. J., Romeo, T., and Parsek, M. R. (2010). *Pseudomonas aeruginosa* biofilm matrix polysaccharide Psl is regulated transcriptionally by RpoS and post-transcriptionally by RsmA. *Mol. Microbiol.* 78, 158–172. doi: 10.1111/j.1365-2958.2010.07320.x

Irnov, I., and Winkler, W. C. (2010). A regulatory RNA required for antitermination of biofilm and capsular polysaccharide operons in Bacillales. *Mol. Microbiol.* 76, 559–575. doi: 10.1111/j.1365-2958.2010.07131.x

Iyer, S., Le, D., Park, B. R., and Kim, M. (2018). Distinct mechanisms coordinate transcription and translation under carbon and nitrogen starvation in *Escherichia coli*. *Nat. Microbiol.* 3, 741–748. doi: 10.1038/s41564-018-0161-3

Jerome, L. J., van Biesen, T., and Frost, L. S. (1999). Degradation of FinP antisense RNA from F-like plasmids: the RNA-binding protein, FinO, protects FinP from ribonuclease E. *J. Mol. Biol.* 285, 1457–1473. doi: 10.1006/jmbi.1998.2404

Jiang, W., Hou, Y., and Inouye, M. (1997). CspA, the major cold-shock protein of *Escherichia coli*, is an RNA chaperone. *J. Biol. Chem.* 272, 196–202. doi: 10.1074/jbc.272.1.196

Jin, D. J., Cagliero, C., and Zhou, Y. N. (2013). Role of RNA polymerase and transcription in the organization of the bacterial nucleoid. *Chem. Rev.* 113, 8662–8682. doi: 10.1021/cr4001429

Johansson, J. (2009). RNA thermosensors in bacterial pathogens. *Contrib. Microbiol.* 16, 150–160. doi: 10.1159/000219378

Johnson, G. E., Lalanne, J. -B., Peters, M. L., and Li, G. -W. (2020). Functionally uncoupled transcription–translation in *Bacillus subtilis*. *Nature* 585, 124–128. doi: 10.1038/s41586-020-2638-5

Johnson, J. E., Reyes, F. E., Polaski, J. T., and Batey, R. T. (2012). B12 cofactors directly stabilize an mRNA regulatory switch. *Nature* 492, 133–137. doi: 10.1038/nature11607

Jonas, K., Edwards, A. N., Ahmad, I., Romeo, T., Römling, U., and Melefors, Ö. (2010). Complex regulatory network encompassing the Csr, c-di-GMP and motility systems of *Salmonella typhimurium*. *Environ. Microbiol.* 12, 524–540. doi: 10.1111/j.1462-2920.2009.02097.x

Joyeux, M. (2016). *In vivo* compaction dynamics of bacterial DNA: a fingerprint of DNA/RNA demixing? *Curr. Opin. Colloid Interface Sci.* 26, 17–27. doi: 10.1016/j.cocis.2016.08.005

Joyeux, M. (2018). A segregative phase separation scenario of the formation of the bacterial nucleoid. *Soft Matter* 14, 7368–7381. doi: 10.1039/c8sm01205a

Kaberdin, V. R., and Bläsi, U. (2006). Translation initiation and the fate of bacterial mRNAs. *FEMS Microbiol. Rev.* 30, 967–979. doi: 10.1111/j.1574-6976.2006.00043.x

Kalapos, M. P., Paulus, H., and Sarkar, N. (1997). Identification of ribosomal protein S1 as a poly(a) binding protein in *Escherichia coli*. *Biochimie* 79, 493–502. doi: 10.1016/S0300-9084(97)82741-1

Kambara, T. K., Ramsey, K. M., and Dove, S. L. (2018). Pervasive targeting of nascent transcripts by Hfq. *Cell Rep.* 23, 1543–1552. doi: 10.1016/j.celrep.2018.03.134

Kang, J. Y., Mishanina, T. V., Bellecourt, M. J., Mooney, R. A., Darst, S. A., and Landick, R. (2018a). RNA polymerase accommodates a pause RNA hairpin by global conformational rearrangements that prolong pausing. *Mol. Cell* 69, 802–815.e1. doi: 10.1016/j.molcel.2018.01.018

Kang, J. Y., Mooney, R. A., Nedialkov, Y., Saba, J., Mishanina, T. V., Artsimovitch, I., et al. (2018b). Structural basis for transcript elongation control by NusG family universal regulators. *Cell* 173, 1650–1662. doi: 10.1016/j.cell.2018.05.017

Kannaiah, S., and Amster-Choder, O. (2014). Protein targeting *via* mRNA in bacteria. *Biochim. Biophys. Acta, Mol. Cell Res.* 1843, 1457–1465. doi: 10.1016/j.bbamcr.2013.11.004

Kannaiah, S., Livny, J., and Amster-Choder, O. (2019). Spatiotemporal organization of the *E. coli* transcriptome: translation independence and engagement in regulation. *Mol. Cell* 76, 574–589.e7. doi: 10.1016/j.molcel.2019.08.013

Katsowich, N., Elbaz, N., Pal, R. R., Mills, E., Kobi, S., Kahan, T., et al. (2017). Host cell attachment elicits posttranscriptional regulation in infecting enteropathogenic bacteria. *Science* 355, 735–739. doi: 10.1126/science.aah4886

Kavita, K., de Mets, F., and 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

Khong, A., and Parker, R. (2020). The landscape of eukaryotic mRNPs. *RNA* 26, 229–239. doi: 10.1261/rna.073601.119

Khosa, S., Scholz, R., Schwarz, C., Trilling, M., Hengel, H., Jaeger, K. E., et al. (2018). An A/U-rich enhancer region is required for high-level protein secretion through the HlyA type I secretion system. *Appl. Environ. Microbiol.* 84:e01163-17. doi: 10.1128/AEM.01163-17

Kim, S., Beltran, B., Irnov, I., and Jacobs-Wagner, C. (2019b). Long-distance cooperative and antagonistic RNA polymerase dynamics *via* DNA supercoiling. *Cell* 179, 106–119.e16. doi: 10.1016/j.cell.2019.08.033

Kim, J., Goñi-Moreno, A., Calles, B., and de Lorenzo, V. (2019a). Spatial organization of the gene expression hardware in *Pseudomonas putida*. *Environ. Microbiol.* 21, 1645–1658. doi: 10.1111/1462-2920.14544

Kohler, R., Mooney, R. A., Mills, D. J., Landick, R., and Cramer, P. (2017). Architecture of a transcribing-translating expressome. *Science* 356, 194–197. doi: 10.1126/science.aal3059

Komarova, A. V., Tchufistova, L. S., Dreyfus, M., and Boni, I. V. (2005). AU-rich sequences within 5' untranslated leaders enhance translation and stabilize mRNA in *Escherichia coli*. *J. Bacteriol.* 187, 1344–1349. doi: 10.1128/JB.187.4.1344-1349.2005

Komarova, A. V., Tchufistova, L. S., Supina, E. V., and Boni, I. V. (2002). Protein S1 counteracts the inhibitory effect of the extended Shine-Dalgarno sequence on translation. *RNA* 8, 1137–1147. doi: 10.1017/S1355838202029990

Kortmann, J., and Narberhaus, F. (2012). Bacterial RNA thermometers: molecular zippers and switches. *Nat. Rev. Microbiol.* 10, 255–265. doi: 10.1038/nrmicro2730

Kriner, M. A., and Groisman, E. A. (2015). The bacterial transcription termination factor rho coordinates Mg_{2+} homeostasis with translational signals. *J. Mol. Biol.* 427, 3834–3849. doi: 10.1016/j.jmb.2015.10.020

Kristiansen, K. I., Weel-Sneve, R., Booth, J. A., and Bjørås, M. (2016). Mutually exclusive RNA secondary structures regulate translation initiation of DinQ in *Escherichia coli*. *RNA* 22, 1739–1749. doi: 10.1261/rna.058461.116

Kurkela, J., Fredman, J., Salminen, T. A., and Tyystjärvi, T. (2020). Revealing secrets of the enigmatic omega subunit of bacterial RNA polymerase. *Mol. Microbiol.* mmi.14603. doi: 10.1111/mmi.14603

Ladouceur, A. -M., Parmar, B., Biedzinski, S., Wall, J., Tope, S. G., Cohn, D., et al. (2020). Clusters of bacterial RNA polymerase are biomolecular condensates that assemble through liquid-liquid phase separation. *Proc. Natl. Acad. Sci. U. S. A.* 117, 18540–18549. doi: 10.1101/2020.03.16.994491

Larson, M. H., Mooney, R. A., Peters, J. M., Windgassen, T., Nayak, D., Gross, C. A., et al. (2014). A pause sequence enriched at translation start sites drives transcription dynamics *in vivo*. *Science* 344, 1042–1047. doi: 10.1126/science.1251871

Lawson, M. R., Ma, W., Bellecourt, M. J., Artsimovitch, I., Martin, A., Landick, R., et al. (2018). Mechanism for the regulated control of bacterial transcription termination by a universal adaptor protein. *Mol. Cell* 71, 911–922.e4. doi: 10.1016/j.molcel.2018.07.014

Le Derout, J., Folichon, M., Briani, F., Dehò, G., Régnier, P., and Hajnsdorf, E. (2003). Hfq affects the length and the frequency of short oligo(A) tails at the 3' end of *Escherchia coli* rpsO mRNAs. *Nucleic Acids Res.* 31, 4017–4023. doi: 10.1093/nar/gkg456

Le, T. T., and Wang, M. D. (2018). Molecular highways—navigating collisions of DNA motor proteins. *J. Mol. Biol.* 430, 4513–4524. doi: 10.1016/j.jmb.2018.08.006

Lewis, P. J., Thaker, S. D., and Errington, J. (2000). Compartmentalization of transcription and translation in *Bacillus subtilis*. *EMBO J.* 19, 710–718. doi: 10.1093/emboj/19.4.710

Li, S. H. J., Li, Z., Park, J. O., King, C. G., Rabinowitz, J. D., Wingreen, N. S., et al. (2018). *Escherichia coli* translation strategies differ across carbon, nitrogen and phosphorus limitation conditions. *Nat. Microbiol.* 3, 939–947. doi: 10.1038/s41564-018-0199-2

Li, R., Zhang, Q., Li, J., and Shi, H. (2015). Effects of cooperation between translating ribosome and RNA polymerase on termination efficiency of the rho-independent terminator. *Nucleic Acids Res.* 44, 2554–2563. doi: 10.1093/nar/gkv1285

Libby, E. A., Roggiani, M., and Goulian, M. (2012). Membrane protein expression triggers chromosomal locus repositioning in bacteria. *Proc. Natl. Acad. Sci. U. S. A.* 109, 7445–7450. doi: 10.1073/pnas.1109479109

Lindahl, L., Archer, R., and Zengel, J. M. (1983). Transcription of the s10 ribosomal protein operon is regulated by an attenuator in the leader. *Cell* 33, 241–248. doi: 10.1016/0092-8674(83)90353-7

Lopez, M. M., Yutani, K., and Makhatadze, G. I. (1999). Interactions of the major cold shock protein of *Bacillus subtilis* CspB with single-stranded DNA templates of different base composition. *J. Biol. Chem.* 274, 33601–33608. doi: 10.1074/jbc.274.47.33601

Magán, A., Amman, F., El-Isa, F., Hartl, N., Shamovsky, I., Nudler, E., et al. (2019). iRAPs curb antisense transcription in *E. coli*. *Nucleic Acids Res.* 47, 10894–10905. doi: 10.1093/nar/gkz791

Mahbub, M., Hemm, L., Yang, Y., Kaur, R., Carmen, H., Engl, C., et al. (2020). mRNA localization, reaction centre biogenesis and thylakoid membrane targeting in cyanobacteria. *Nat. Plants* 6, 1179–1191. doi: 10.1038/s41477-020-00764-2

Makarova, O. V., Makarov, E. M., Sousa, R., and Dreyfus, M. (1995). Transcribing of *Escherichia coli* genes with mutant T7 RNA polymerases: stability of lacZ mRNA inversely correlates with polymerase speed. *Proc. Natl. Acad. Sci. U. S. A.* 92, 12250–12254. doi: 10.1073/pnas.92.26.12250

Malabirade, A., Morgado-Brajones, J., Trépout, S., Wien, F., Marquez, I., Seguin, J., et al. (2017). Membrane association of the bacterial riboregulator Hfq and functional perspectives. *Sci. Rep.* 7:10724. doi: 10.1038/s41598-017-11157-5

Maquat, L. E., Tarn, W. Y., and Isken, O. (2010). The pioneer round of translation: features and functions. *Cell* 142, 368–374. doi: 10.1016/j.cell.2010.07.022

Masachis, S., and Darfeuille, F. (2018). Type I toxin-antitoxin systems: regulating toxin expression *via* Shine-Dalgarno sequence sequestration and small RNA binding. *Microbiol. Spectr.* 6, 1–18. doi: 10.1128/microbiolspec.rwr-0030-2018

Mascarenhas, J., Weber, M. H. W., and Graumann, P. L. (2001). Specific polar localization of ribosomes in *Bacillus subtilis* depends on active transcription. *EMBO Rep.* 2, 685–689. doi: 10.1093/embo-reports/kve160

Mechold, U., Potrykus, K., Murphy, H., Murakami, K. S., and Cashel, M. (2013). Differential regulation by ppGpp versus pppGpp in *Escherichia coli*. *Nucleic Acids Res.* 41, 6175–6189. doi: 10.1093/nar/gkt302

Melamed, S., Adams, P. P., Zhang, A., Zhang, H., and Storz, G. (2020). RNA-RNA interactomes of ProQ and Hfq reveal overlapping and competing roles. *Mol. Cell* 77, 411–425.e7. doi: 10.1016/j.molcel.2019.10.022

Michaux, C., Holmqvist, E., Vasicek, E., Sharan, M., Barquist, L., Westermann, A. J., et al. (2017). 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. doi: 10.1073/pnas.1620772114

Miller, O. L., Hamkalo, B. A., and Thomas, C. A. (1970). Visualization of bacterial genes in action. *Science* 169, 392–395. doi: 10.1126/science.169.3943.392

Milon, P., Tischenko, E., Tomšic, J., Caserta, E., Folkers, G., La Teana, A., et al. (2006). The nucleotide-binding site of bacterial translation initiation factor 2 (IF2) as a metabolic sensor. *Proc. Natl. Acad. Sci. U. S. A.* 103, 13962–13967. doi: 10.1073/pnas.0606384103

Mironov, A. S., Gusarov, I., Rafikov, R., Lopez, L. E., Shatalin, K., Kreneva, R. A., et al. (2002). Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. *Cell* 111, 747–756. doi: 10.1016/S0092-8674(02)01134-0

Mitkevich, V. A., Ermakov, A., Kulikova, A. A., Tankov, S., Shyp, V., Soosaar, A., et al. (2010). Thermodynamic characterization of ppGpp binding to EF-G

or IF2 and of initiator tRNA binding to free IF2 in the presence of GDP, GTP, or ppGpp. *J. Mol. Biol.* 402, 838–846. doi: 10.1016/j.jmb.2010.08.016

Moffitt, J. R., Pandey, S., Boettiger, A. N., Wang, S., and Zhuang, X. (2016). Spatial organization shapes the turnover of a bacterial transcriptome. *elife* 5:e13065. doi: 10.7554/eLife.13065

Mogridge, J., and Greenblatt, J. (1998). Specific binding of *Escherichia coli* ribosomal protein S1 to boxA transcriptional antiterminator RNA. *J. Bacteriol.* 180, 2248–2252. doi: 10.1128/jb.180.8.2248-2252.1998

Mohanty, B. K., and Kushner, S. R. (2006). The majority of *Escherichia coli* mRNAs undergo post-transcriptional modification in exponentially growing cells. *Nucleic Acids Res.* 34, 5695–5704. doi: 10.1093/nar/gkl684

Mohanty, B. K., Maples, V. F., and Kushner, S. R. (2004). The Sm-like protein Hfq regulates polyadenylation dependent mRNA decay in *Escherichia coli*. *Mol. Microbiol.* 54, 905–920. doi: 10.1111/j.1365-2958.2004.04337.x

Mohapatra, S., Choi, H., Ge, X., Sanyal, S., and Weisshaar, J. C. (2017). Spatial distribution and ribosome-binding dynamics of EF-P in live *Escherichia coli*. *MBio* 8:e00300-17. doi: 10.1128/mbio.00300-17

Mohapatra, S., and Weisshaar, J. C. (2018). Functional mapping of the *E. coli* translational machinery using single-molecule tracking. *Mol. Microbiol.* 110, 262–282. doi: 10.1111/mmi.14103

Moll, I., Afonyushkin, T., Vytvytska, O., Kaberdin, V. R., and Bläsi, U. (2003). Coincident Hfq binding and RNase E cleavage sites on mRNA and small regulatory RNAs. *RNA* 9, 1308–1314. doi: 10.1261/rna.5850703

Mondal, J., Bratton, B. P., Li, Y., Yethiraj, A., and Weisshaar, J. C. (2011). Entropy-based mechanism of ribosome-nucleoid segregation in *E. coli* cells. *Biophys. J.* 100, 2605–2613. doi: 10.1016/j.bpj.2011.04.030

Montero Llopis, P., Jackson, A. F., Sliusarenko, O., Surovtsev, I., Heinritz, J., Emonet, T., et al. (2010). Spatial organization of the flow of genetic information in bacteria. *Nature* 466, 77–81. doi: 10.1038/nature09152

Mooney, R. A., Davis, S. E., Peters, J. M., Rowland, J. L., Ansari, A. Z., and Landick, R. (2009a). Regulator trafficking on bacterial transcription units *in vivo*. *Mol. Cell* 33, 97–108. doi: 10.1016/j.molcel.2008.12.021

Mooney, R. A., Schweimer, K., Rösch, P., Gottesman, M., and Landick, R. (2009b). Two structurally independent domains of *E. coli* NusG create regulatory plasticity *via* distinct interactions with RNA polymerase and regulators. *J. Mol. Biol.* 391, 341–358. doi: 10.1016/j.jmb.2009.05.078

Morita, M. T., Tanaka, Y., Kodama, T. S., Kyogoku, Y., Yanagi, H., and Yura, T. (1999). Translational induction of heat shock transcription factor σ32: evidence for a built-in RNA thermosensor. *Genes Dev.* 13, 655–665. doi: 10.1101/gad.13.6.655

Müller, P., Gimpel, M., Wildenhain, T., and Brantl, S. (2019). A new role for CsrA: promotion of complex formation between an sRNA and its mRNA target in *Bacillus subtilis*. *RNA Biol.* 16, 972–987. doi: 10.1080/15476286.2019.1605811

Mustafi, M., and Weisshaar, J. C. (2018). Simultaneous binding of multiple EF-Tu copies to translating ribosomes in live *Escherichia coli*. *MBio* 9:e02143-17. doi: 10.1128/mBio.02143-17

Muthunayake, N. S., Tomares, D. T., Childers, W. S., and Schrader, J. M. (2020). Phase-separated bacterial ribonucleoprotein bodies organize mRNA decay. *Wiley Interdiscip. Rev. RNA* 11:e1599. doi: 10.1002/wrna.1599

Nevo-Dinur, K., Nussbaum-Shochat, A., Ben-Yehuda, S., and Amster-Choder, O. (2011). Translation-independent localization of mRNA in *E. coli*. *Science* 331, 1081–1084. doi: 10.1126/science.1195691

O'Reilly, F. J., Xue, L., Graziadei, A., Sinn, L., Lenz, S., Tegunov, D., et al. (2020). In-cell architecture of an actively transcribing-translating expressome. *Science* 369, 554–557. doi: 10.1101/2020.02.28.970111

Pal, K., Yadav, M., Jain, S., Ghosh, B., Sen, R., and Sen, U. (2019). *Vibrio cholerae* YaeO is a structural homologue of RNA chaperone Hfq that inhibits rho-dependent transcription termination by dissociating its hexameric state. *J. Mol. Biol.* 431, 4749–4766. doi: 10.1016/j.jmb.2019.09.019

Pannuri, A., Yakhnin, H., Vakulskas, C. A., Edwards, A. N., Babitzke, P., and Romeo, T. (2012). Translational repression of NhaR, a novel pathway for multi-tier regulation of biofilm circuitry by CsrA. *J. Bacteriol.* 194, 79–89. doi: 10.1128/JB.06209-11

Papenfort, K., Sun, Y., Miyakoshi, M., Vanderpool, C. K., and Vogel, J. (2013). Small RNA-mediated activation of sugar phosphatase mRNA regulates glucose homeostasis. *Cell* 153, 426–437. doi: 10.1016/j.cell.2013.03.003

Parry, B. R., Surovtsev, I. V., Cabeen, M. T., O'Hern, C. S., Dufresne, E. R., and Jacobs-Wagner, C. (2014). The bacterial cytoplasm has glass-like properties and is fluidized by metabolic activity. *Cell* 156, 183–194. doi: 10.1016/j.cell.2013.11.028

Paul, B. J., Ross, W., Gaal, T., and Gourse, R. L. (2004). rRNA transcription in *Escherichia coli*. *Annu. Rev. Genet.* 38, 749–770. doi: 10.1146/annurev.genet.38.072902.091347

Pei, X. Y., Dendooven, T., Sonnleitner, E., Chen, S., Bläsi, U., and Luisi, B. F. (2019). Architectural principles for hfq/crc-mediated regulation of gene expression. *elife* 8:e43158. doi: 10.7554/eLife.43158

Persson, F., Lindén, M., Unoson, C., and Elf, J. (2013). Extracting intracellular diffusive states and transition rates from single-molecule tracking data. *Nat. Methods* 10, 265–269. doi: 10.1038/nmeth.2367

Phadtare, S., and Inouye, M. (1999). Sequence-selective interactions with RNA by CspB, CspC and CspE, members of the CspA family of *Escherichia coli*. *Mol. Microbiol.* 33, 1004–1014. doi: 10.1046/j.1365-2958.1999.01541.x

Phadtare, S., and Inouye, M. (2001). Role of CspC and CspE in regulation of expression of RpoS and UspA, the stress response proteins in *Escherichia coli*. *J. Bacteriol.* 183, 1205–1214. doi: 10.1128/JB.183.4.1205-1214.2001

Phadtare, S., Inouye, M., and Severinov, K. (2002a). The nucleic acid melting activity of *Escherichia coli* CspE is critical for transcription antitermination and cold acclimation of cells. *J. Biol. Chem.* 277, 7239–7245. doi: 10.1074/jbc.M111496200

Phadtare, S., and Severinov, K. (2005). Nucleic acid melting by *Escherichia coli* CspE. *Nucleic Acids Res.* 33, 5583–5590. doi: 10.1093/nar/gki859

Phadtare, S., Tyagi, S., Inouye, M., and Severinov, K. (2002b). Three amino acids in *Escherichia coli* CspE surface-exposed aromatic patch are critical for nucleic acid melting activity leading to transcription antitermination and cold acclimation of cells. *J. Biol. Chem.* 277, 46706–46711. doi: 10.1074/jbc.M208118200

Pichoff, S., Alibaud, L., Guédant, A., Castanié, M. P., and Bouché, J. P. (1998). An *Escherichia coli* gene (yaeO) suppresses temperature-sensitive mutations in essential genes by modulating rho-dependent transcription termination. *Mol. Microbiol.* 29, 859–869. doi: 10.1046/j.1365-2958.1998.00981.x

Plochowietz, A., Farrell, I., Smilansky, Z., Cooperman, B. S., and Kapanidis, A. N. (2016). *In vivo* single-RNA tracking shows that most tRNA diffuses freely in live bacteria. *Nucleic Acids Res.* 45, 926–937. doi: 10.1093/nar/gkw787

Polissi, A., De Laurentis, W., Zangrossi, S., Briani, F., Longhi, V., Pesole, G., et al. (2003). Changes in *Escherichia coli* transcriptome during acclimatization at low temperature. *Res. Microbiol.* 154, 573–580. doi: 10.1016/S0923-2508(03)00167-0

Potrykus, K., and Cashel, M. (2008). (p)ppGpp: still magical? *Annu. Rev. Microbiol.* 62, 35–51. doi: 10.1146/annurev.micro.62.081307.162903

Potts, A. H., Vakulskas, C. A., Pannuri, A., Yakhnin, H., Babitzke, P., and Romeo, T. (2017). Global role of the bacterial post-transcriptional regulator CsrA revealed by integrated transcriptomics. *Nat. Commun.* 8:1596. doi: 10.1038/s41467-017-01613-1

Pourciau, C., Pannuri, A., Potts, A., Yakhnin, H., Babitzke, P., and Romeo, T. (2019). Regulation of iron storage by CsrA supports exponential growth of *Escherichia coli*. *MBio* 10:e01034-19. doi: 10.1128/mBio.01034-19

Prilusky, J., and Bibi, E. (2009). Studying membrane proteins through the eyes of the genetic code revealed a strong uracil bias in their coding mRNAs. *Proc. Natl. Acad. Sci. U. S. A.* 106, 6662–6666. doi: 10.1073/pnas.0902029106

Proshkin, S., Rachid Rahmouni, A., Mironov, A., and Nudler, E. (2010). Cooperation between translating ribosomes and RNA polymerase in transcription elongation. *Science* 328, 504–508. doi: 10.1126/science.1184939

Qayyum, M. Z., Dey, D., and Sen, R. (2016). Transcription elongation factor NusA is a general antagonist of rho-dependent termination in *Escherichia coli*. *J. Biol. Chem.* 291, 8090–8108. doi: 10.1074/jbc.M115.701268

Rabhi, M., Espéli, O., Schwartz, A., Cayrol, B., Rahmouni, A. R., Arluison, V., et al. (2011). The Sm-like RNA chaperone Hfq mediates transcription antitermination at rho-dependent terminators. *EMBO J.* 30, 2805–2816. doi: 10.1038/emboj.2011.192

Ratje, A. H., Loerke, J., Mikolajka, A., Brünner, M., Hildebrand, P. W., Starosta, A. L., et al. (2010). Head swivel on the ribosome facilitates translocation by means of intra-subunit tRNA hybrid sites. *Nature* 468, 713–716. doi: 10.1038/nature09547

Ray, S., Chauvier, A., and Walter, N. G. (2019a). Kinetics coming into focus: single-molecule microscopy of riboswitch dynamics. *RNA Biol.* 16, 1077–1085. doi: 10.1080/15476286.2018.1536594

Ray, S., Costa, R.Da, Das, M., and Nandi, D. (2019b). Interplay of cold shock protein E with an uncharacterized protein, YciF, lowers porin expression and enhances bile resistance in *Salmonella Typhimurium*. *J. Biol. Chem.* 294, 9084–9099. doi:10.1074/jbc.RA119.008209

Rennella, E., Sára, T., Juen, M., Wunderlich, C., Imbert, L., Solyom, Z., et al. (2017). RNA binding and chaperone activity of the *E. coli* cold-shock protein CspA. *Nucleic Acids Res.* 45, 4255–4268. doi: 10.1093/nar/gkx044

Richards, J., and Belasco, J. G. (2019). Obstacles to scanning by RNase E govern bacterial mRNA lifetimes by hindering access to distal cleavage sites. *Mol. Cell* 74, 284–295.e5. doi: 10.1016/j.molcel.2019.01.044

Richardson, J. P. (1991). Preventing the synthesis of unused transcripts by rho factor. *Cell* 64, 1047–1049. doi: 10.1016/0092-8674(91)90257-Y

Roberts, J. W. (2019). Mechanisms of bacterial transcription termination. *J. Mol. Biol.* 431, 4030–4039. doi: 10.1016/j.jmb.2019.04.003

Roggiani, M., and Goulian, M. (2015). Chromosome-membrane interactions in bacteria. *Annu. Rev. Genet.* 49, 115–129. doi: 10.1146/annurev-genet-112414-054958

Romeo, T., and Babitzke, P. (2018). Global regulation by CsrA and its RNA antagonists. *Microbiol. Spectr.* 6, 341-354. doi: 10.1128/microbiolspec.rwr-0009-2017

Romeo, T., Gong, M., Liu, M. Y., and Brun-Zinkernagel, A. M. (1993). 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, 4744–4755. doi: 10.1128/jb.175.15.4744-4755.1993

Rudner, D. Z., and Losick, R. (2010). Protein subcellular localization in bacteria. *Cold Spring Harb. Perspect. Biol.* 2:a000307. doi: 10.1101/cshperspect.a000307

Sahr, T., Rusniok, C., Impens, F., Oliva, G., Sismeiro, O., Coppée, J. Y., et al. (2017). The *Legionella pneumophila* genome evolved to accommodate multiple regulatory mechanisms controlled by the CsrA-system. *PLoS Genet.* 13:e1006629. doi: 10.1371/journal.pgen.1006629

Said, N., Hilal, T., Sunday, N. D., Khatri, A., Bürger, J., Mielke, T., et al. (2020). Steps toward translocation-independent RNA polymerase inactivation by terminator ATPase ρ. *Science* 371:eabd1673. doi: 10.1126/science.abd1673

Salvail, H., Caron, M. P., Bélanger, J., and Massé, E. (2013). Antagonistic functions between the RNA chaperone Hfq and an sRNA regulate sensitivity to the antibiotic colicin. *EMBO J.* 32, 2764–2778. doi: 10.1038/emboj.2013.205

Sanamrad, A., Persson, F., Lundius, E. G., Fange, D., Gynnå, A. H., and Elf, J. (2014). Single-particle tracking reveals that free ribosomal subunits are not excluded from the *Escherichia coli* nucleoid. *Proc. Natl. Acad. Sci. U. S. A.* 111, 11413–11418. doi: 10.1073/pnas.1411558111

Sanchez-Vazquez, P., Dewey, C. N., Kitten, N., Ross, W., and Gourse, R. L. (2019). Genome-wide effects on *Escherichia coli* transcription from ppGpp binding to its two sites on RNA polymerase. *Proc. Natl. Acad. Sci. U. S. A.* 116, 8310–8319. doi: 10.1073/pnas.1819682116

Santangelo, T. J., and Artsimovitch, I. (2011). Termination and antitermination: RNA polymerase runs a stop sign. *Nat. Rev. Microbiol.* 9, 319–329. doi: 10.1038/nrmicro2560

Saxena, S., Myka, K. K., Washburn, R., Costantino, N., Court, D. L., and Gottesman, M. E. (2018). *Escherichia coli* transcription factor NusG binds to 70S ribosomes. *Mol. Microbiol.* 108, 495–504. doi: 10.1111/mmi.13953

Schlax, P. J., Xavier, K. A., Gluick, T. C., and Draper, D. E. (2001). Translational repression of the *Escherichia coli* α operon mRNA. Importance of an mRNA conformational switch and a ternary entrapment complex. *J. Biol. Chem.* 276, 38494–38501. doi: 10.1074/jbc.M106934200

Schmidt, M. C., and Chamberlin, M. J. (1987). nusA protein of *Escherichia coli* is an efficient transcription termination factor for certain terminator sites. *J. Mol. Biol.* 195, 809–818. doi: 10.1016/0022-2836(87)90486-4

Schuwirth, B. S., Borovinskaya, M. A., Hau, C. W., Zhang, W., Vila-Sanjurjo, A., Holton, J. M., et al. (2005). Structures of the bacterial ribosome at 3.5 Å resolution. *Science* 310, 827–834. doi: 10.1126/science.1117230

Sedlyarova, N., Rescheneder, P., Magán, A., Popitsch, N., Rziha, N., Bilusic, I., et al. (2017). Natural RNA polymerase aptamers regulate transcription in *E. coli*. *Mol. Cell* 67, 30–43.e6. doi: 10.1016/j.molcel.2017.05.025

Sedlyarova, N., Shamovsky, I., Bharati, B. K., Epshtein, V., Chen, J., Gottesman, S., et al. (2016). sRNA-mediated control of transcription termination in *E. coli*. *Cell* 167, 111–121.e13. doi: 10.1016/j.cell.2016.09.004

Sen, R., Chalissery, J., Qayyum, M. Z., Vishalini, V., and Muteeb, G. (2014). Nus factors of *Escherichia coli*. *EcoSal Plus* 6, 1–16. doi: 10.1128/ecosalplus.esp-0008-2013

Sevostyanova, A., and Groisman, E. A. (2015). An RNA motif advances transcription by preventing rho-dependent termination. *Proc. Natl. Acad. Sci. U. S. A.* 112, E6835–E6843. doi: 10.1073/pnas.1515383112

Shaham, G., and Tuller, T. (2018). Genome scale analysis of *Escherichia coli* with a comprehensive prokaryotic sequence-based biophysical model of translation initiation and elongation. *DNA Res.* 25, 195–205. doi: 10.1093/dnares/dsx049

Sheng, H., Stauffer, W. T., Hussein, R., Lin, C., and Lim, H. N. (2017). Nucleoid and cytoplasmic localization of small RNAs in *Escherichia coli*. *Nucleic Acids Res.* 45, 2919–2934. doi: 10.1093/nar/gkx023

Shenhar, Y., Biran, D., and Ron, E. Z. (2012). Resistance to environmental stress requires the RNA chaperones CspC and CspE. *Environ. Microbiol. Rep.* 4, 532–539. doi: 10.1111/j.1758-2229.2012.00358.x

Sherwood, A. V., and Henkin, T. M. (2016). Riboswitch-mediated gene regulation: novel RNA architectures dictate gene expression responses. *Annu. Rev. Microbiol.* 70, 361–374. doi: 10.1146/annurev-micro-091014-104306

Shi, J., Gao, X., Tian, T., Yu, Z., Gao, B., Wen, A., et al. (2019). Structural basis of Q-dependent transcription antitermination. *Nat. Commun.* 10:2925. doi: 10.1038/s41467-019-10958-8

Shokeen, S., Patel, S., Greenfield, T. J., Brinkman, C., and Weaver, K. E. (2008). Translational regulation by an intramolecular stem-loop is required for intermolecular RNA regulation of the par addiction module. *J. Bacteriol.* 190, 6076–6083. doi: 10.1128/JB.00660-08

Silva, I. J., Barahona, S., Eyraud, A., Lalaouna, D., Figueroa-Bossi, N., Massé, E., et al. (2019). SraL sRNA interaction regulates the terminator by preventing premature transcription termination of rho mRNA. *Proc. Natl. Acad. Sci. U. S. A.* 116, 3042–3051. doi: 10.1073/pnas.1811589116

Singh, D., Chang, S. J., Lin, P. H., Averina, O. V., Kaberdin, V. R., and Lin-Chao, S. (2009). Regulation of ribonuclease E activity by the L4 ribosomal protein of *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 106, 864–869. doi: 10.1073/pnas.0810205106

Singh, D., Murashko, O. N., and Lin-Chao, S. (2020). Posttranscriptional regulation of tnaA by protein-RNA interaction mediated by ribosomal protein L4 in *Escherichia coli*. *J. Bacteriol.* 202:e00799-19. doi: 10.1128/jb.00799-19

Siwiak, M., and Zielenkiewicz, P. (2013). Transimulation – protein biosynthesis web service. *PLoS One* 8:e73943. doi: 10.1371/journal.pone.0073943

Smirnov, A., Förstner, K. U., Holmqvist, E., Otto, A., Günster, R., Becher, D., et al. (2016). Grad-seq guides the discovery of ProQ as a major small RNA-binding protein. *Proc. Natl. Acad. Sci. U. S. A.* 113, 11591–11596. doi: 10.1073/pnas.1609981113

Smirnov, A., Wang, C., Drewry, L. L., and Vogel, J. (2017). Molecular mechanism of mRNA repression in trans by a ProQ-dependent small RNA. *EMBO J.* 36, 1029–1045. doi: 10.15252/embj.201696127

Sonnleitner, E., and Bläsi, U. (2014). Regulation of Hfq by the RNA CrcZ in *Pseudomonas aeruginosa* carbon catabolite repression. *PLoS Genet.* 10:e1004440. doi: 10.1371/journal.pgen.1004440

Spedding, G., and Draper, D. E. (1993). Allosteric mechanism for translational repression in the *Escherichia coli* α operon. *Proc. Natl. Acad. Sci. U. S. A.* 90, 4399–4403. doi: 10.1073/pnas.90.10.4399

Squires, C. L., and Zaporojets, D. (2000). Proteins shared by the transcription and translation machines. *Annu. Rev. Microbiol.* 54, 775–798. doi: 10.1146/annurev.micro.54.1.775

Stein, E. M., Kwiatkowska, J., Basczok, M. M., Gravel, C. M., Berry, K. E., and Olejniczak, M. (2020). Determinants of RNA recognition by the FinO domain of the *Escherichia coli* ProQ protein. *Nucleic Acids Res.* 48, 7502–7519. doi: 10.1093/nar/gkaa497

Steinberg, R., Origi, A., Natriashvili, A., Sarmah, P., Licheva, M., Walker, P. M., et al. (2020). Posttranslational insertion of small membrane proteins by the bacterial signal recognition particle. *PLoS Biol.* 18:e3000874. doi: 10.1371/journal.pbio.3000874

Stent, G. S. (1964). The operon: on its third anniversary – modulation of transfer RNA species can provide a workable model of an operator-less operon. *Science* 144, 816–820. doi: 10.1126/science.144.3620.816

Stevenson-Jones, F., Woodgate, J., Castro-Roa, D., and Zenkin, N. (2020). Ribosome reactivates transcription by physically pushing RNA polymerase out of transcription arrest. *Proc. Natl. Acad. Sci.* 117, 8462–8467. doi: 10.1073/pnas.1919985117

Stracy, M., Lesterlin, C., Garza de Leon, F., Uphoff, S., Zawadzki, P., and Kapanidis, A. N. (2015). Live-cell superresolution microscopy reveals the organization of RNA polymerase in the bacterial nucleoid. *Proc. Natl. Acad. Sci. U. S. A.* 112, E4390–E4399. doi: 10.1073/pnas.1507592112

Strahl, H., Turlan, C., Khalid, S., Bond, P. J., Kebalo, J. M., Peyron, P., et al. (2015). Membrane recognition and dynamics of the RNA degradosome. *PLoS Genet.* 11:e1004961. doi: 10.1371/journal.pgen.1004961

Sukhodolets, M. V., and Garges, S. (2003). Interaction of *Escherichia coli* RNA polymerase with the ribosomal protein S1 and the Sm-like ATPase Hfq. *Biochemistry* 42, 8022–8034. doi: 10.1021/bi020638i

Sukhodolets, M. V., Garges, S., and Adhya, S. (2006). Ribosomal protein S1 promotes transcriptional cycling. *RNA* 12, 1505–1513. doi: 10.1261/rna.2321606

Surovtsev, I. V., and Jacobs-Wagner, C. (2018). Subcellular organization: a critical feature of bacterial cell replication. *Cell* 172, 1271–1293. doi: 10.1016/j.cell.2018.01.014

Syal, K., Flentie, K., Bhardwaj, N., Maiti, K., Jayaraman, N., Stallings, C. L., et al. (2017). Synthetic (p)ppGpp analogue is an inhibitor of stringent response in mycobacteria. *Antimicrob. Agents Chemother.* 61:e00443-17. doi: 10.1128/AAC.00443-17

Taghbalout, A., Yang, Q., and Arluison, V. (2014). The *Escherichia coli* RNA processing and degradation machinery is compartmentalized within an organized cellular network. *Biochem. J.* 458, 11–22. doi: 10.1042/BJ20131287

Thisted, T., Sørensen, N. S., and Gerdes, K. (1995). Mechanism of post-segregational killing: secondary structure analysis of the entire Hok mRNA from plasmid R1 suggests a fold-back structure that prevents translation and antisense RNA binding. *J. Mol. Biol.* 247, 859–873. doi: 10.1006/jmbi.1995.0186

Thomas, M. S., Bedwell, D. M., and Nomura, M. (1987). Regulation of α operon gene expression in *Escherichia coli*. A novel form of translational coupling. *J. Mol. Biol.* 196, 333–345. doi: 10.1016/0022-2836(87)90694-2

Torres, M., Condon, C., Balada, J. -M., Squires, C., and Squires, C. L. (2001). Ribosomal protein S4 is a transcription factor with properties remarkably similar to NusA, a protein involved in both non-ribosomal and ribosomal RNA antitermination. *EMBO J.* 20, 3811–3820. doi: 10.1093/emboj

Toulokhonov, I., Artsimovitch, I., and Landick, R. (2001). Allosteric control of RNA polymerase by a site that contacts nascent RNA hairpins. *Science* 292, 730–733. doi: 10.1126/science.1057738

Tree, J. J., Granneman, S., McAteer, S. P., Tollervey, D., and Gally, D. L. (2014). Identification of bacteriophage-encoded anti-sRNAs in pathogenic *Escherichia coli*. *Mol. Cell* 55, 199–213. doi: 10.1016/j.molcel.2014.05.006

Uhm, H., Kang, W., Ha, K. S., Kang, C., and Hohng, S. (2017). Single-molecule FRET studies on the cotranscriptional folding of a thiamine pyrophosphate riboswitch. *Proc. Natl. Acad. Sci. U. S. A.* 115, 331–336. doi: 10.1073/pnas.1712983115

Uniacke, J., and Zerges, W. (2009). Chloroplast protein targeting involves localized translation in Chlamydomonas. *Proc. Natl. Acad. Sci. U. S. A.* 106, 1439–1444. doi: 10.1073/pnas.0811268106

Updegrove, T. B., Zhang, A., and Storz, G. (2016). Hfq: the flexible RNA matchmaker. *Curr. Opin. Microbiol.* 30, 133–138. doi: 10.1016/j.mib.2016.02.003

Vakulskas, C. A., Potts, A. H., Babitzke, P., Ahmer, B. M. M., and Romeo, T. (2015). Regulation of bacterial virulence by Csr (Rsm) systems. *Microbiol. Mol. Biol. Rev.* 79, 193–224. doi: 10.1128/mmbr.00052-14

Valencia-Burton, M., McCullough, R. M., Cantor, C. R., and Broude, N. E. (2007). RNA visualization in live bacterial cells using fluorescent protein complementation. *Nat. Methods* 4, 421–427. doi: 10.1038/nmeth1023

van Gijtenbeek, L. A., Robinson, A., van Oijen, A. M., Poolman, B., and Kok, J. (2016). On the spatial organization of mRNA, plasmids, and ribosomes in a bacterial host overexpressing membrane proteins. *PLoS Genet.* 12:e1006523. doi: 10.1371/journal.pgen.1006523

van Nues, R. W., Castro-Roa, D., Yuzenkova, Y., and Zenkin, N. (2015). Ribonucleoprotein particles of bacterial small non-coding RNA IsrA (IS61 or McaS) and its interaction with RNA polymerase core may link transcription to mRNA fate. *Nucleic Acids Res.* 44, 2577–2592. doi: 10.1093/nar/gkv1302

Večerek, B., Moll, I., and Bläsi, U. (2005). Translational autocontrol of the *Escherichia coli* hfq RNA chaperone gene. *RNA* 11, 976–984. doi: 10.1261/rna.2360205

Vinogradova, D. S., Zegarra, V., Maksimova, E., Nakamoto, J. A., Kasatsky, P., Paleskava, A., et al. (2020). How the initiating ribosome copes with ppGpp to translate mRNAs. *PLoS Biol.* 18:e3000593. doi: 10.1371/journal.pbio.3000593

Vogel, U., and Jensen, K. F. (1994a). Effects of guanosine 3',5'-bisdiphosphate (ppGpp) on rate of transcription elongation in isoleucine-starved *Escherichia coli*. *J. Biol. Chem.* 269, 16236–16241.

Vogel, U., and Jensen, K. F. (1994b). The RNA chain elongation rate in *Escherichia coli* depends on the growth rate. *J. Bacteriol.* 176, 2807–2813. doi: 10.1128/jb.176.10.2807-2813.1994

Vogel, U., Sørensen, M., Pedersen, S., Jensen, K. F., and Kilstrup, M. (1992). Decreasing transcription elongation rate in *Escherichia coli* exposed to amino acid starvation. *Mol. Microbiol.* 6, 2191–2200. doi: 10.1111/j.1365-2958.1992.tb01393.x

Volkov, I. L., Lindén, M., Aguirre Rivera, J., Ieong, K. -W., Metelev, M., Elf, J., et al. (2018). tRNA tracking for direct measurements of protein synthesis kinetics in live cells. *Nat. Chem. Biol.* 14, 618–626. doi: 10.1038/s41589-018-0063-y

Wagner, E. G. H., and Romby, P. (2015). Small RNAs in bacteria and archaea: who they are, what they do, and how they do it. *Adv. Genet.* 90, 133–208. doi: 10.1016/bs.adgen.2015.05.001

Wang, B., and Artsimovitch, I. (2020). A growing gap between the RNAP and the lead ribosome. *Trends Microbiol.* 29, 4–5. doi: 10.1016/j.tim.2020.09.011

Wang, B., Dai, P., Ding, D., Del Rosario, A., Grant, R. A., Pentelute, B. L., et al. (2019). Affinity-based capture and identification of protein effectors of the growth regulator ppGpp. *Nat. Chem. Biol.* 15, 141–150. doi: 10.1038/s41589-018-0183-4

Wang, B., Grant, R. A., and Laub, M. T. (2020a). ppGpp coordinates nucleotide and amino-acid synthesis in *E. coli* during starvation. *Mol. Cell* 80, 29–42. e10. doi: 10.1016/j.molcel.2020.08.005

Wang, W., Li, G. W., Chen, C., Xie, X. S., and Zhuang, X. (2011). Chromosome organization by a nucleoid-associated protein in live bacteria. *Science* 333, 1445–1449. doi: 10.1126/science.1204697

Wang, C., Molodtsov, V., Firlar, E., Kaelber, J. T., Blaha, G., Su, M., et al. (2020b). Structural basis of transcription-translation coupling. *Science* 369, 1359–1365. doi: 10.1126/science.abb5317

Wang, N., Yamanaka, K., and Inouye, M. (1999). CspI, the ninth member of the CspA family of *Escherichia coli*, is induced upon cold shock. *J. Bacteriol.* 181, 1603–1609. doi: 10.1128/jb.181.5.1603-1609.1999

Washburn, R. S., Zuber, P. K., Sun, M., Hashem, Y., Shen, B., Li, W., et al. (2020). *Escherichia coli* NusG links the lead ribosome with the transcription elongation complex. *iScience* 23:101352. doi: 10.1016/j.isci.2020.101352

Watters, K. E., Strobel, E. J., Yu, A. M., Lis, J. T., and Lucks, J. B. (2016). Cotranscriptional folding of a riboswitch at nucleotide resolution. *Nat. Struct. Mol. Biol.* 23, 1124–1131. doi: 10.1038/nsmb.3316

Weber, M. H. W., Volkov, A. V., Fricke, I., Marahiel, M. A., and Graumann, P. L. (2001). Localization of cold shock proteins to cytosolic spaces surrounding nucleoids in *Bacillus subtilis* depends on active transcription. *J. Bacteriol.* 183, 6435–6443. doi: 10.1128/JB.183.21.6435-6443.2001

Webster, M. W., Takacs, M., Zhu, C., Vidmar, V., Eduljee, A., Abdelkareem, M., et al. (2020). Structural basis of transcription-translation coupling and collision in bacteria. *Science* 369, 1355–1359. doi: 10.1126/science.abb5036

Wei, B. L., Brun-Zinkernagel, A. M., Simecka, J. W., Prüß, B. M., Babitzke, P., and Romeo, T. (2001). Positive regulation of motility and flhDC expression by the RNA-binding protein CsrA of *Escherichia coli*. *Mol. Microbiol.* 40, 245–256. doi: 10.1046/j.1365-2958.2001.02380.x

Wen, J., Harp, J. R., and Fozo, E. M. (2017). The 5', UTR of the type i toxin ZorO can both inhibit and enhance translation. *Nucleic Acids Res.* 45, 4006–4020. doi: 10.1093/nar/gkw1172

Wendrich, T. M., Blaha, G., Wilson, D. N., Marahiel, M. A., and Nierhaus, K. H. (2002). Dissection of the mechanism for the stringent factor RelA. *Mol. Cell* 10, 779–788. doi: 10.1016/S1097-2765(02)00656-1

Weng, X., Bohrer, C. H., Bettridge, K., Lagda, A. C., Cagliero, C., Jin, D. J., et al. (2019). Spatial organization of RNA polymerase and its relationship with transcription in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 116, 20115–20123. doi: 10.1073/pnas.1903968116

Westermann, A. J., Venturini, E., Sellin, M. E., Förstner, K. U., Hardt, W. D., and Vogel, J. (2019). The major RNA-binding protein ProQ impacts virulence gene expression in *Salmonella enterica* serovar typhimurium. *MBio* 10:e02504-18. doi: 10.1128/mBio.02504-18

Winkler, W., Nahvi, A., and Breaker, R. R. (2002). Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* 419, 952–956. doi: 10.1038/nature01145

Woldringh, C. L. (2002). The role of co-transcriptional translation and protein translocation (transertion) in bacterial chromosome segregation. *Mol. Microbiol.* 45, 17–29. doi: 10.1046/j.1365-2958.2002.02993.x

Yakhnin, A. V., Baker, C. S., Vakulskas, C. A., Yakhnin, H., Berezin, I., Romeo, T., et al. (2013). CsrA activates flhDC expression by protecting flhDC mRNA from RNase E-mediated cleavage. *Mol. Microbiol.* 87, 851–866. doi: 10.1111/mmi.12136

Yamanaka, K., Mitani, T., Ogura, T., Niki, H., and Hiraga, S. (1994). Cloning, sequencing, and characterization of multicopy suppressors of a mukB mutation in *Escherichia coli*. *Mol. Microbiol.* 13, 301–312. doi: 10.1111/j.1365-2958.1994.tb00424.x

Yang, S., Kim, S., Kim, D. K., Jeon An, H., Bae Son, J., Hedén Gynnå, A., et al. (2019). Transcription and translation contribute to gene locus relocation to the nucleoid periphery in *E. coli*. *Nat. Commun.* 10:5131. doi: 10.1038/s41467-019-13152-y

Yanofsky, C. (1981). Attenuation in the control of expression of bacterial operons. *Nature* 289, 751–758. doi: 10.1038/289751a0

Yin, Z., Kaelber, J. T., and Ebright, R. H. (2019). Structural basis of Q-dependent antitermination. *Proc. Natl. Acad. Sci. U. S. A.* 116, 18384–18390. doi: 10.1073/pnas.1909801116

Young, R., and Bremer, H. (1976). Polypeptide chain elongation rate in *Escherichia coli* B/r as a function of growth rate. *Biochem. J.* 160, 185–194. doi: 10.1042/bj1600185

Zhang, A., Wassarman, K. M., Rosenow, C., Tjaden, B. C., Storz, G., and Gottesman, S. (2003). Global analysis of small RNA and mRNA targets of Hfq. *Mol. Microbiol.* 50, 1111–1124. doi: 10.1046/j.1365-2958.2003.03734.x

Zhang, Y., Zborníková, E., Rejman, D., and Gerdes, K. (2018). Novel (p)ppGpp binding and metabolizing proteins of *Escherichia coli*. *MBio* 9:e02188–17. doi: 10.1128/mBio.02188-17

Zhu, M., and Dai, X. (2019). Maintenance of translational elongation rate underlies the survival of *Escherichia coli* during oxidative stress. *Nucleic Acids Res.* 47, 7592–7604. doi: 10.1093/nar/gkz467

Zhu, M., and Dai, X. (2020). Bacterial stress defense: the crucial role of ribosome speed. *Cell. Mol. Life Sci.* 77, 853–858. doi: 10.1007/s00018-019-03304-0

Zhu, M., Dai, X., and Wang, Y. -P. (2016). Real time determination of bacterial *in vivo* ribosome translation elongation speed based on LacZα complementation system. *Nucleic Acids Res.* 44:e155. doi: 10.1093/nar/gkw698

Zhu, M., Mori, M., Hwa, T., and Dai, X. (2019). Disruption of transcription–translation coordination in *Escherichia coli* leads to premature transcriptional termination. *Nat. Microbiol.* 4, 2347–2356. doi: 10.1038/s41564-019-0543-1

Zhu, Y., Mustafi, M., and Weisshaar, J. C. (2020). Biophysical properties of *Escherichia coli* cytoplasm in stationary phase by superresolution fluorescence microscopy. *MBio* 11:e00143-20. doi: 10.1128/mBio.00143-20

Zuber, P. K., Schweimer, K., Rösch, P., Artsimovitch, I., and Knauer, S. H. (2019). Reversible fold-switching controls the functional cycle of the antitermination factor RfaH. *Nat. Commun.* 10:702. doi: 10.1038/s41467-019-08567-6

Zuo, Y., Wang, Y., and Steitz, T. A. (2013). The mechanism of *E. coli* RNA polymerase regulation by ppGpp is suggested by the structure of their complex. *Mol. Cell* 50, 430–436. doi: 10.1016/j.molcel.2013.03.020

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

MINI REVIEW
published: 11 March 2021
doi: 10.3389/fmicb.2021.634004

tRNAs as a Driving Force of Genome Evolution in Yeast

Ana Rita Guimarães, Inês Correia, Inês Sousa, Carla Oliveira, Gabriela Moura, Ana Rita Bezerra and Manuel A. S. Santos**

Department of Medical Sciences, Institute of Biomedicine – iBiMED, University of Aveiro, Aveiro, Portugal

Transfer RNAs (tRNAs) are widely known for their roles in the decoding of the linear mRNA information into amino acid sequences of proteins. They are also multifunctional platforms in the translation process and have other roles beyond translation, including sensing amino acid abundance, interacting with the general stress response machinery, and modulating cellular adaptation, survival, and death. In this mini-review, we focus on the emerging role of tRNA genes in the organization and modification of the genomic architecture of yeast and the role of tRNA misexpression and decoding infidelity in genome stability, evolution, and adaption. We discuss published work showing how quickly tRNA genes can mutate to meet novel translational demands, how tRNAs speed up genome evolution, and how tRNA genes can be sites of genomic instability. We highlight recent works showing that loss of tRNA decoding fidelity and small alterations in tRNA expression have unexpected and profound impacts on genome stability. By dissecting these recent evidence, we hope to lay the groundwork that prompts future investigations on the mechanistic interplay between tRNAs and genome modification that likely triggers genome evolution.

Keywords: tRNA, yeast, chromatin structure, chromosome architecture, genome evolution, genomic instability

OPEN ACCESS

Edited by:
Omar Orellana, University of Chile, Chile

Reviewed by:
Francisco A. Cubillos, University of Santiago, Chile
Cécile Neuvéglise, Institut National de la Recherche Agronomique, Centre Montpellier, France

****Correspondence:***
Ana Rita Bezerra
armbezerra@ua.pt
Manuel A. S. Santos
msantos@ua.pt

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *26 November 2020*
Accepted: *01 February 2021*
Published: *11 March 2021*

Citation:
Guimarães AR, Correia I, Sousa I, Oliveira C, Moura G, Bezerra AR and Santos MAS (2021) tRNAs as a Driving Force of Genome Evolution in Yeast. Front. Microbiol. 12:634004. doi: 10.3389/fmicb.2021.634004

INTRODUCTION

Transfer RNAs (tRNAs) are short non-coding RNAs, approximately 70 to 100 bases long, that play essential roles in translation by linking mRNA codons to their corresponding amino acids, following a set of decoding rules established by the genetic code. They do so by base pairing their anticodon triplets with mRNA codon triplets in the ribosome decoding center and transferring the amino acid attached to its 3′-end in the ribosomal peptidyl transferase center (Phizicky and Hopper, 2010). This is a critical cellular process that requires tight control of tRNA gene expression, tRNA maturation, tRNA charging, and turnover (Chan and Lowe, 2016). In actively dividing yeast cells, tRNAs represent approximately 15% of total RNA (Warner, 1999), indicating that their genes (tDNAs) are highly transcribed. In general, tDNAs are nucleosome free and are flanked by strongly positioned nucleosomes (Yuan et al., 2005; Cole et al., 2012). Their transcription is mediated by RNA polymerase III (Pol III) upon recruitment to the promoter by the transcription factors TFIIIC and TFIIIB. TFIIIC binds to the internal A-box and B-box promoter elements and helps recruit the multi-subunit factor TFIIIB to AT-rich sequences upstream of the transcription start site, forming a

highly stable TFIIIB–DNA complex that participates in multiple rounds of Pol III recruitment and initiation (Schramm and Hernandez, 2002). In *Saccharomyces cerevisiae*, the cellular concentration of each tRNA is directly proportional to its gene copy number (Percudani et al., 1997). This is particularly important because translation efficiency is described as the degree to which the tRNA pool can accommodate the transcriptome, thus affecting protein production and accuracy (dos Reis et al., 2004). This interplay is fine-tuned by codon usage, which is under selective pressure and show variation across budding yeast species (LaBella et al., 2019). Yet tRNAs have other non-canonical roles in the biological theater beyond their role as adaptors in protein synthesis (reviewed in Raina and Ibba, 2014 and Su et al., 2020). For example, tDNAs have roles in chromatin organization and gene regulation and are sites for binding of numerous chromatin proteins, including the architectural structural maintenance of chromosomes (SMC) proteins, nuclear pore proteins, chromatin remodelers, and histone modifiers (D'Ambrosio et al., 2008; Su et al., 2020).

The 275 tDNAs present in the yeast *S. cerevisiae* genome are dispersed throughout the linear maps of the 16 chromosomes. Fluorescence *in situ* hybridization microscopy (FISH) showed that tDNAs cluster at the outer periphery of the nucleolus in a microtubule-dependent manner and or adjacent to centromeres (Thompson et al., 2003). This happens with the assistance of condensing complexes bound at each tDNA gene locus (Haeusler et al., 2008) and requires substantial rearrangements of the genome topology. Whether individual tDNA associations play a role in genome organization is still poorly understood. We review below recent works on how tDNAs and related Pol III promoter elements function as boundary elements that limit chromatin domains (Donze and Kamakaka, 2001), how they work as barriers to DNA replication fork progression, and how they contribute to the formation of genomic fragile sites (Pryce et al., 2009). Beyond their role in the three-dimensional and functional organization of the genome, this review also describes how changes in the tRNA pool can drive genome evolution in fungi.

ROLES OF tRNA GENES IN CHROMATIN REMODELING AND GENOME ORGANIZATION

The three-dimensional organization of the genome can promote long-range genomic rearrangements between interacting loci whose associated chromatin and transcriptional states can be selected through evolution (Bagadia et al., 2016). In yeast, tRNA genes have been implicated in the spatial organization of the genome by acting as barrier elements and by regulating chromatin structure (Donze and Kamakaka, 2001; Noma et al., 2006; Simms et al., 2008; Iwasaki et al., 2010; Hamdani et al., 2019). Evidence that tDNAs can hamper silenced chromatin domains from invading active domains was first obtained in *S. cerevisiae* (Donze and Kamakaka, 2001; Simms et al., 2004), where the deletion of a Thr-tRNA$_{AGU}$ gene at the transcriptionally silent HMR mating-type locus resulted in the spread of silencing and consequent repression of the *GIT1* gene on chromosome III (Donze and Kamakaka, 2001). A Gln-tRNA$_{UUG}$ gene has also been shown to block silencing at the *S. cerevisiae* rDNA locus (RDN1) (Biswas et al., 2009). Insulator activity was similarly shown in *Schizosaccharomyces pombe*, where deletion of a centromeric Ala-tRNA gene led to the spread of pericentromeric heterochromatin and gene silencing (Scott et al., 2006). The precise mechanisms by which tDNAs exert their barrier function remain largely unexplored; however, the assembly of the complete Pol III transcription apparatus does seem to be required for barrier function (Donze and Kamakaka, 2001; Scott et al., 2006; Biswas et al., 2009). Mutations in internal Thr-tRNA promoter elements, A-box or B-box, at the HMR locus led to deficiencies of TFIIIC and TFIIIB assembly, resulting in the loss of barrier function in *S. cerevisiae* (Donze and Kamakaka, 2001). Furthermore, yeast cohesin complex mutants (Δ*smc1* and Δ*smc3*) have impaired tDNA-mediated insulator function (Donze et al., 1999).

A study by Duan et al. (2010) mapped *cis*- and *trans*-interactions across the entire genome in *S. cerevisiae* and showed that physical interactions among tDNAs are significantly enriched and that they largely co-localize into clusters associated with the nucleolus or centromeres (Duan et al., 2010). Other studies, using DNA FISH, also showed that some tRNA genes cluster together near centromeres (Thompson et al., 2003). Furthermore, microscopic observations and genome-wide mapping of physical interactions show the co-localization of TFIIIC, cohesins, and other structural proteins at tDNA physical domain borders, suggesting that these insulators are critical players in chromosome folding and organization in the yeast nucleus. Recently, Hamdani et al. (2019) devised a strategy to tackle this topic. They eliminated the internal promoter elements (A-box and B-box) of two tDNAs on the left arm and eight tDNAs on the right arm of chromosome III in *S. cerevisiae* to generate a "tDNA-less" chromosome where binding of transcription factors TFIIIC and TFIIIB and chromatin proteins was abrogated. This allowed the detailed characterization of chromatin packaging, folding, and nuclear dynamics of chromosome III. Using various approaches, such as MNase-seq, ChIP-seq, RNA-seq, and fluorescence microscopy co-localization analysis, authors showed that (1) tDNA loss affects chromatin structure by disrupting the precise nucleosome positioning outside tDNAs; (2) tDNAs are essential to recruitment of cohesins and condensins; and (3) tDNAs influence centromere clustering, which in turn affects nuclear architecture. Lastly, as in previous studies (Donze and Kamakaka, 2001; Simms et al., 2004; Biswas et al., 2009), loss of tDNAs alters the long-range interactions of the silenced HML and HMR loci of chromosome III, leading to alterations in gene silencing (Hamdani et al., 2019).

The discovery of tDNA insulator function in yeast along with the recent advances in uncovering their involvement in the functional and spatial organization of the genome is particularly relevant because they provide a framework for future studies in this field. Furthermore, tDNA insulator functions seem to be conserved from yeast to humans (Raab et al., 2012), and their activities appear to be associated with a significant number

of protein complexes whose actions and regulation remain to be determined.

tRNA GENES, R-LOOPS, AND TRANSPOSABLE ELEMENTS

tDNAs are often located near naturally occurring genomic fragile sites, and genome-wide studies in *S. cerevisiae* have detected R-loops at tRNA genes, along with other Pol III transcribed genes (Chan et al., 2014; El Hage et al., 2014; Wahba et al., 2016; Yeung and Smith, 2020). The replication machinery naturally slows down at tDNAs, and DNA helicases must take action to promote the progression of the replication fork (Ivessa et al., 2003). However, when the direction of DNA replication conflicts with the direction of the tDNA transcription, it leads to replication-fork pausing (Osmundson et al., 2017; Yeung and Smith, 2020). Head-on replication-fork pausing promotes DNA damage by R-loop formation (Tran et al., 2017). R-loops are stable DNA:RNA hybrid structures with an unpaired DNA strand that naturally blocks replication but can also generate genomic instability (Santos-Pereira and Aguilera, 2015). If left unresolved, R-loops can create replication–transcription conflicts and lead to double-strand breaks which potentially increase DNA recombination (Hegazy et al., 2020). Tran et al. (2017) showed that tDNAs represent sites of double-strand breaks and of increased recombination events in a series of helicase mutants. Moreover, this phenomenon is also intimately connected with high expression levels of tRNAs and with the fact that tDNAs are usually associated with the pre-initiation complex, *i.e.*, at a ready transcription state, which is a stable multiprotein complex consisting of a constant passage barrier for helicases (Arimbasseri et al., 2014).

Comparative genomics of 11 evolutionary-related yeast species showed a prevalence of tRNA genes at DNA breakpoints, which have also been linked to sites of genomic rearrangement (Gordon et al., 2009). One of the aspects that could underlie this observation is the preferential integration of transposable elements (TEs) at the proximity of tDNAs (Hani and Feldmann, 1998). TEs are mobile self-replicating elements that can integrate themselves in new genomic sites, being a potential source of mutations. There is an underlying assumption that TE insertions are deleterious, and indeed, they are a potential threat to genome integrity. *S. cerevisiae* has five families of TEs classified as long terminal repeat (LTR) retrotransposons, *Ty1* to *Ty5*. The most abundant and active ones are the Ty1 and Ty2, which, apart from their ORFs, share a high sequence similarity between their LTR sequences (Carr et al., 2012). These long and near-identical sequences scattered in the genome are prone to recombination, particularly ectopic recombination, which allows for an array of rearrangements like deletions, duplications, inversions, and translocations (Mieczkowski et al., 2006). It is therefore important that a tight control of retrotransposons' expression is maintained. Ty1 mobility is regulated by a retrograde mechanism where Ty1 self-encoded elements, like p22, inhibits Ty1's mobility when an elevated number of copies are present (Saha et al., 2015). This ability in *S. cerevisiae* was acquired by horizontal transfer from *Saccharomyces paradoxus* (Czaja et al., 2020). Nevertheless, comparative studies have shown that a large percentage of TEs are fixed in the genome (Bensasson, 2011), although there is also evidence for recent *Ty* insertions at a high rate (Carr et al., 2012), which can be seen as a source of genomic diversity and evolution. Increased retromobility has been observed upon exposure to several stress conditions like UV light (Bradshaw and McEntee, 1989) and adenine starvation (Todeschini et al., 2005). Furthermore, in physiological conditions like aging, retromobility has been observed in several species (Dennis et al., 2012; De Cecco et al., 2013; Li et al., 2013). In yeast, Maxwell et al. (2011) reported that during chronological aging, there is an association between Ty1 mobility and the observed genomic instability, particularly in loss of heterozygosity (LOH) events. Important adaptative roles for transposons have also been reported in experimentally evolved yeast. Dunham et al. (2002) studied the recombination events in evolved strains under glucose limitation and found that almost all detected rearrangements could be traced to ectopic rearrangement between transposons, transposon fragments, or tRNAs. In an experimental evolution study of cells expressing a mutant Ser-tRNA (see below), there were large chromosomal rearrangements mediated by homologous recombination between transposons (Kalapis et al., 2015). In a large timescale, phylogenetic studies on genome evolution identified tRNAs and transposons at genome breakpoints and rearrangement sites (Fischer et al., 2000; Kellis et al., 2003; Gordon et al., 2009). Interestingly, in a comparative evolutionary study between *S. cerevisiae* and its related species *S. paradoxus*, *S. mikatae*, and *S. uvarum*, all the inversions identified were flanked by tDNAs in an opposite transcriptional orientation (Kellis et al., 2003). Thus, it is apparent that tDNAs, or their flaking regions, play an important role in genome innovation and evolution. Although it was not acknowledged in any study (at least to our understanding), it is possible that tDNAs (and their vicinity) represent "silent hotspots" for recombination that, when a particular condition is prolonged, become sites for "rapid" adaptive genomic alterations.

GENOMIC CHANGES ASSOCIATED WITH tRNA MISEXPRESSION

Upon environmental challenges, the organism quickly needs a particular set of defenses to survive. There is a body of evidence on how transcription changes in response to several stresses in yeast (Gasch et al., 2000; Gutin et al., 2019). It is, therefore, reasonable to assume that tRNA expression and abundance must also be tuned to follow these changes. Indeed, the tRNA pool dynamically changes to facilitate selective and faster translation of stress-related transcripts (Torrent et al., 2018). The tRNA pool is composed of various tRNA isoacceptor families, each encoded by tDNAs with different copy numbers. tRNA gene families with more copies of the same tDNA decode more frequently used codons, while tRNAs with one gene copy decode rarely used codons, which correlates with the codon usage of protein genes. This establishes the adequate balance between tRNA availability and the usage of its corresponding codon (Percudani et al., 1997).

Curiously, not all copies of the same tRNA species contribute equally to the tRNA pool, and the loss of a particular copy can have different physiological consequences (Bloom-Ackermann et al., 2014). Thus, the multiplicity of copies enables higher expression of tRNAs in high translational demand and enables the tRNA pool to be dynamic enough to allow the dispensability of a particular copy to further expand the tRNA repertoire. This concept was explored by Yona et al. (2013) in a yeast strain with a deletion in the single-copy tDNA *tR(CCU)J* (Bloom-Ackermann et al., 2014), thus eliminating the only cognate tRNA for the AGG codon. Experimental evolution revealed that 200 generations were sufficient for cells to overcome the translational defect. Translational equilibrium was restored by mutating the anticodon of one of the 11 copies of Arg-tRNA$_{UCU}$ from UCU

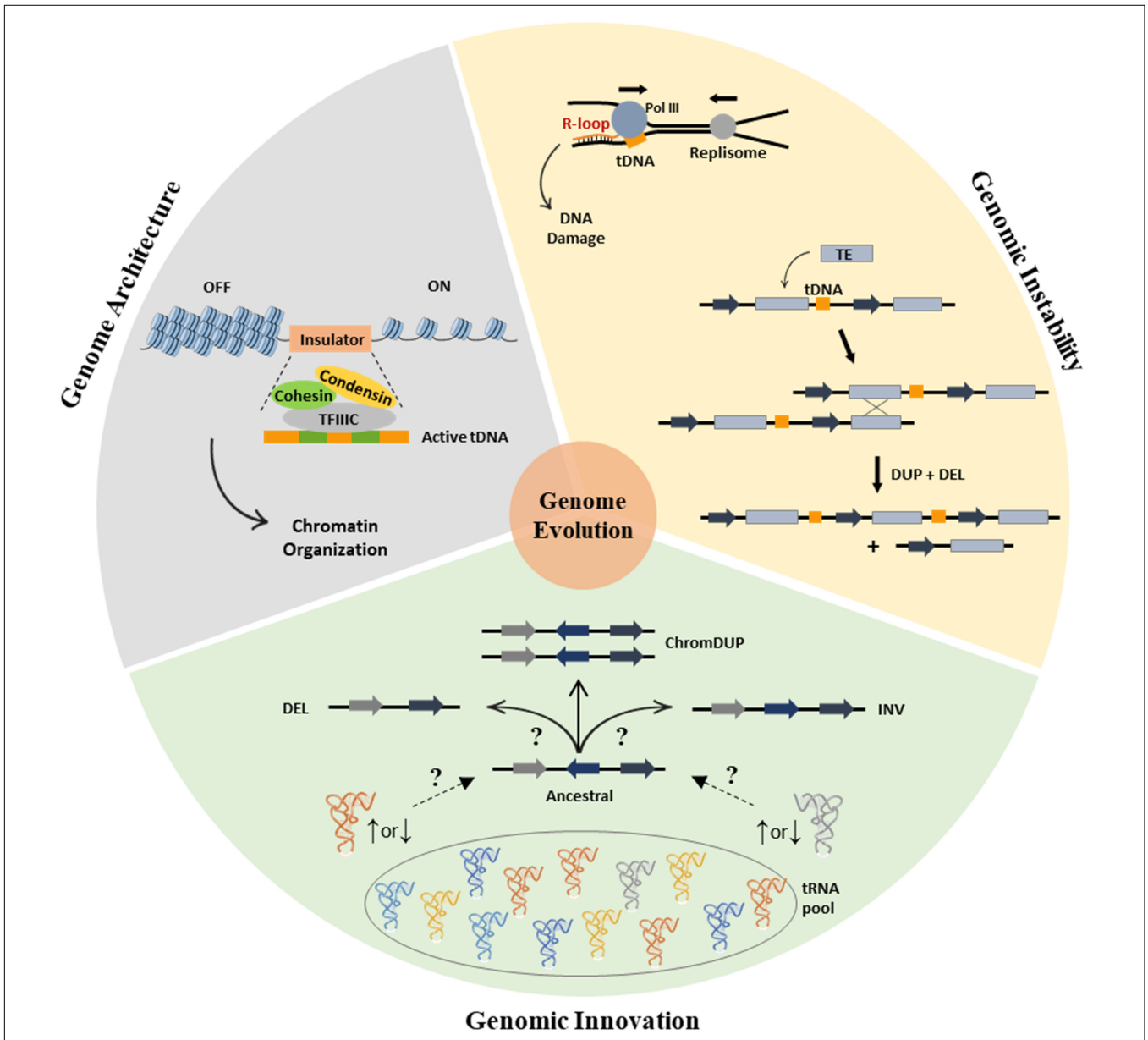


FIGURE 1 | The involvement of tRNAs in genome organization and evolution. (Gray section) Role of tDNAs in genome architecture. TFIII recognizes the tDNA promoters (green boxes) enabling the assembly of the complete Pol III transcription apparatus and the recruitment of cohesin and condensin. Recruitment of the latter is crucial as it blocks the spread of heterochromatin (silenced state, OFF) into the active euchromatin (ON). (Yellow section) Contribution of tDNAs for genomic instability. tRNA genes are known sites for R-loop formation, which can be precursors of genomic instability, particularly when the directions of genome replication and of tRNA transcription collide (top). TEs integrate into the genome preferentially upstream of tDNAs. TEs are prone sites for ectopic recombination, like the one depicted, where recombination between sister chromatids results in deletions and duplications of the sequences located within the TEs (bottom). (Green section) The elusive connection between alterations in the tRNA pool and the genome instability. Alterations of tRNA's abundance level destabilize the proteome, leading to adaptive genome instability and mutations through poorly understood mechanisms (represented by "?").

to CCU, without affecting cellular fitness, highlighting how the plasticity of the tRNA pool can overcome translational challenges in changing environments.

Maintaining the proteome's good health is of extreme importance, but several bacterial and fungal species are able to decrease translation fidelity during stress to functionally diversify the proteome, a phenomenon called adaptive translation (Pan, 2013). Although alterations in the identity of a sense codon are a rare phenomenon, several budding yeasts reassigned the CUG codon to serine (Santos and Tuite, 1995) and to alanine (Muhlhausen et al., 2016; Riley et al., 2016). The CUG reassignments occurred independently during evolution and involved different tRNA genes that convergently mutated anticodons to CAG (Krassowski et al., 2018). *Candida albicans* is the most studied example of adaptive translation, where the identity of the CUG codon was altered to serine but residual leucine identity was still maintained. This results in an ambiguous CUG codon that is translated 97% of the times as serine and 3% as leucine, in standard growth conditions (Gomes et al., 2007). This is accomplished by a single Ser-tRNA$_{CAG}$ with identity elements for both seryl- and leucyl-tRNA synthetases (Suzuki et al., 1997). However, alteration of the levels of the CUG-decoding tRNA is surprisingly adaptative. Bezerra et al. (2013) engineered a set of *C. albicans* strains with different combinations of tDNA$_{CAG}$ copy number, where one, two, or both copies of the endogenous Ser-tRNA$_{CAG}$ genes were deleted and one or two copies of the *S. cerevisiae* Leu-tRNA$_{CAG}$ genes were inserted, thus shifting the ratio of leucine/serine incorporated in the proteome. Strains tolerated increasing Leu incorporation and displayed unexpected phenotypic variability, with highly variable colony and cell morphologies, and increased tolerance to fluconazole and itraconazole. Interestingly, altering the copy number of the CUG decoding tDNAs leads to the rapid accumulation of unique single-nucleotide polymorphisms (SNPs) and LOH events. Strains with higher deregulation of the tRNA pool, and therefore higher levels of Leu incorporation at CUG sites, showed higher number of SNPs, indicating the potential mutagenic effect of tRNA codon misreading. Of note was the fact that strains with the most extreme alterations in the tRNA pool (*i.e.*, with highest level of Leu incorporation) presented a near-complete LOH on chromosome V. A set of genes related with stress response, antifungal drug resistance, filamentous growth, and pathogenesis is located in this chromosome, showing that these alterations are not random and have an adaptative role (Bezerra et al., 2013). One could hypothesize that the observed genomic alterations triggered by tRNA misexpression are associated with the peculiar features of the *C. albicans* biology and its highly plastic genome (Selmecki et al., 2010). However, a similar phenotype was also uncovered in *S. cerevisiae*. Kalapis et al. (2015) experimentally evolved a yeast strain engineered with a mutant Ser-tRNA$_{CAG}$ that misincorporates serine at CUG codons. Although this insertion was highly detrimental for fitness, cells were able to adapt to their new condition after 250 generations to their new condition. Genome sequencing showed that tolerance and adaptation to translational stress were achieved by large genomic rearrangements. These repeatedly involved a partial deletion of 127 kb at chromosome V, enriched with genes involved in deubiquitination processes, and a duplication of 540 kb in chromosome IV, enriched with genes involved in glucose uptake. Together, these allowed cells to adapt to imbalances in the tRNA pool that culminate in CUG mistranslation by accelerated protein turnover and a high rate of glucose internalization (Kalapis et al., 2015). In other words, alterations in the tRNA pool and mRNA decoding accuracy destabilize the proteome in a dynamic way that reciprocates to the genome and produce important adaptive genome instabilities.

CONCLUSION AND FUTURE PERSPECTIVES

We highlighted the non-canonical function of tRNAs and tDNAs as drivers of genome evolution (**Figure 1**) and summarized how tDNA can play a role in the three-dimensional and functional organization of the genome and potentiate genome rearrangements events. Additionally, the discovery that alterations in the yeast tRNA pool generate genome instability associated with phenotypic variation of high adaptation potential adds a new dimension to the study of tRNA-driven genome evolution. Precisely how these mechanisms operate remains to be determined, but future work should elucidate how the tRNA pool provides evolutionary plasticity in environmental changing conditions. It is of high biological importance to understand the complex relationship between the tRNA pool and the genome, since the produced genomic instabilities may be relevant to human diseases, including cancer where extensive tRNA pool alterations and aneuploidies have been observed.

AUTHOR CONTRIBUTIONS

AG and AB: conceptualization, literature study, writing, and editing of the manuscript. MS: editing and revision of the manuscript. All authors listed have made a substantial, direct and intellectual contributions to the work, and approved it for publication.

FUNDING

This work was supported by FEDER (Fundo Europeu de Desenvolvimento Regional) funds through the COMPETE 2020, Operational Program for Competitiveness and Internationalization (POCI), and by Portuguese national funds via Fundação para a Ciência e a Tecnologia, I.P. (FCT) under the projects PTDC/BIA-MIB/31238/2017 and PTDC/BIA-MIC/31849/2017. The iBiMED research unit was supported by FCT funds under UIDP/04501/2020. AG was supported directly by an FCT grant (SFRH/BD/121358/2016). AB was supported by national funds (OE), through FCT, I.P., in the scope of the framework contract foreseen in numbers 4, 5, and 6 of article 23, of Decree-Law 57/2016 of August 29, changed by Law 57/2017 of July 19.

REFERENCES

Arimbasseri, A. G., Rijal, K., and Maraia, R. J. (2014). Comparative overview of RNA polymerase II and III transcription cycles, with focus on RNA polymerase III termination and reinitiation. *Transcription* 5:e27639. doi: 10.4161/trns.27369

Bagadia, M., Singh, A., and Singh Sandhu, K. (2016). Three dimensional organization of genome might have guided the dynamics of gene order evolution in eukaryotes. *Genome Biol. Evol.* 8, 946–954. doi: 10.1093/gbe/evw050

Bensasson, D. (2011). Evidence for a high mutation rate at rapidly evolving yeast centromeres. *BMC Evol. Biol.* 11:211. doi: 10.1186/1471-2148-11-211

Bezerra, A. R., Simoes, J., Lee, W., Rung, J., Weil, T., Gut, I. G., et al. (2013). Reversion of a fungal genetic code alteration links proteome instability with genomic and phenotypic diversification. *Proc. Natl. Acad. Sci. U.S.A.* 110, 11079–11084. doi: 10.1073/pnas.1302094110

Biswas, M., Maqani, N., Rai, R., Kumaran, S. P., Iyer, K. R., Sendinc, E., et al. (2009). Limiting the extent of the RDN1 heterochromatin domain by a silencing barrier and Sir2 protein levels in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 29, 2889–2898. doi: 10.1128/MCB.00728-08

Bloom-Ackermann, Z., Navon, S., Gingold, H., Towers, R., Pilpel, Y., and Dahan, O. (2014). A comprehensive tRNA deletion library unravels the genetic architecture of the tRNA pool. *PLoS Genet.* 10:e1004084. doi: 10.1371/journal.pgen.1004084

Bradshaw, V. A., and McEntee, K. (1989). DNA damage activates transcription and transposition of yeast Ty retrotransposons. *Mol. Gen. Genet.* 218, 465–474. doi: 10.1007/BF00332411

Carr, M., Bensasson, D., and Bergman, C. M. (2012). Evolutionary genomics of transposable elements in *Saccharomyces cerevisiae*. *PLoS One* 7:e50978. doi: 10.1371/journal.pone.0050978

Chan, P. P., and Lowe, T. M. (2016). GtRNAdb 2.0: an expanded database of transfer RNA genes identified in complete and draft genomes. *Nucleic Acids Res.* 44, D184–D189. doi: 10.1093/nar/gkv1309

Chan, Y. A., Aristizabal, M. J., Lu, P. Y., Luo, Z., Hamza, A., Kobor, M. S., et al. (2014). Genome-wide profiling of yeast DNA:RNA hybrid prone sites with DRIP-chip. *PLoS Genet.* 10:e1004288. doi: 10.1371/journal.pgen.1004288

Cole, H. A., Howard, B. H., and Clark, D. J. (2012). Genome-wide mapping of nucleosomes in yeast using paired-end sequencing. *Methods Enzymol.* 513, 145–168. doi: 10.1016/B978-0-12-391938-0.00006-9

Czaja, W., Bensasson, D., Ahn, H. W., Garfinkel, D. J., and Bergman, C. M. (2020). Evolution of Ty1 copy number control in yeast by horizontal transfer and recombination. *PLoS Genet.* 16:e1008632. doi: 10.1371/journal.pgen.1008632

D'Ambrosio, C., Schmidt, C. K., Katou, Y., Kelly, G., Itoh, T., Shirahige, K., et al. (2008). Identification of cis-acting sites for condensin loading onto budding yeast chromosomes. *Genes Dev.* 22, 2215–2227. doi: 10.1101/gad.1675708

De Cecco, M., Criscione, S. W., Peterson, A. L., Neretti, N., Sedivy, J. M., and Kreiling, J. A. (2013). Transposable elements become active and mobile in the genomes of aging mammalian somatic tissues. *Aging* 5, 867–883. doi: 10.18632/aging.100621

Dennis, S., Sheth, U., Feldman, J. L., English, K. A., and Priess, J. R. (2012). C. elegans germ cells show temperature and age-dependent expression of Cer1, a Gypsy/Ty3-related retrotransposon. *PLoS Pathog.* 8:e1002591. doi: 10.1371/journal.ppat.1002591

Donze, D., Adams, C. R., Rine, J., and Kamakaka, R. T. (1999). The boundaries of the silenced HMR domain in *Saccharomyces cerevisiae*. *Genes Dev.* 13, 698–708. doi: 10.1101/gad.13.6.698

Donze, D., and Kamakaka, R. T. (2001). RNA polymerase III and RNA polymerase II promoter complexes are heterochromatin barriers in *Saccharomyces cerevisiae*. *EMBO J.* 20, 520–531. doi: 10.1093/emboj/20.3.520

dos Reis, M., Savva, R., and Wernisch, L. (2004). Solving the riddle of codon usage preferences: a test for translational selection. *Nucleic Acids Res.* 32, 5036–5044. doi: 10.1093/nar/gkh834

Duan, Z., Andronescu, M., Schutz, K., McIlwain, S., Kim, Y. J., Lee, C., et al. (2010). A three-dimensional model of the yeast genome. *Nature* 465, 363–367. doi: 10.1038/nature08973

Dunham, M. J., Badrane, H., Ferea, T., Adams, J., Brown, P. O., Rosenzweig, F., et al. (2002). Characteristic genome rearrangements in experimental evolution of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U.S.A.* 99, 16144–16149. doi: 10.1073/pnas.242624799

El Hage, A., Webb, S., Kerr, A., and Tollervey, D. (2014). Genome-wide distribution of RNA-DNA hybrids identifies RNase H targets in tRNA genes, retrotransposons and mitochondria. *PLoS Genet.* 10:e1004716. doi: 10.1371/journal.pgen.1004716

Fischer, G., James, S. A., Roberts, I. N., Oliver, S. G., and Louis, E. J. (2000). Chromosomal evolution in *Saccharomyces*. *Nature* 405, 451–454. doi: 10.1038/35013058

Gasch, A. P., Spellman, P. T., Kao, C. M., Carmel-Harel, O., Eisen, M. B., Storz, G., et al. (2000). Genomic expression programs in the response of yeast cells to environmental changes. *Mol. Biol. Cell* 11, 4241–4257. doi: 10.1091/mbc.11.12.4241

Gomes, A. C., Miranda, I., Silva, R. M., Moura, G. R., Thomas, B., Akoulitchev, A., et al. (2007). A genetic code alteration generates a proteome of high diversity in the human pathogen *Candida albicans*. *Genome Biol.* 8:R206. doi: 10.1186/gb-2007-8-10-r206

Gordon, J. L., Byrne, K. P., and Wolfe, K. H. (2009). Additions, losses, and rearrangements on the evolutionary route from a reconstructed ancestor to the modern *Saccharomyces cerevisiae* genome. *PLoS Genet.* 5:e1000485. doi: 10.1371/journal.pgen.1000485

Gutin, J., Joseph-Strauss, D., Sadeh, A., Shalom, E., and Friedman, N. (2019). Genetic screen of the yeast environmental stress response dynamics uncovers distinct regulatory phases. *Mol. Syst. Biol.* 15:e8939. doi: 10.15252/msb.20198939

Haeusler, R. A., Pratt-Hyatt, M., Good, P. D., Gipson, T. A., and Engelke, D. R. (2008). Clustering of yeast tRNA genes is mediated by specific association of condensin with tRNA gene transcription complexes. *Genes Dev.* 22, 2204–2214. doi: 10.1101/gad.1675908

Hamdani, O., Dhillon, N., Hsieh, T. S., Fujita, T., Ocampo, J., Kirkland, J. G., et al. (2019). tRNA genes affect chromosome structure and function via local effects. *Mol. Cell. Biol.* 39, e00432–18. doi: 10.1128/MCB.00432-18

Hani, J., and Feldmann, H. (1998). tRNA genes and retroelements in the yeast genome. *Nucleic Acids Res.* 26, 689–696. doi: 10.1093/nar/26.3.689

Hegazy, Y. A., Fernando, C. M., and Tran, E. J. (2020). The balancing act of R-loop biology: the good, the bad, and the ugly. *J. Biol. Chem.* 295, 905–913. doi: 10.1074/jbc.REV119.011353

Ivessa, A. S., Lenzmeier, B. A., Bessler, J. B., Goudsouzian, L. K., Schnakenberg, S. L., and Zakian, V. A. (2003). The *Saccharomyces cerevisiae* helicase Rrm3p facilitates replication past nonhistone protein-DNA complexes. *Mol. Cell* 12, 1525–1536. doi: 10.1016/s1097-2765(03)00456-8

Iwasaki, O., Tanaka, A., Tanizawa, H., Grewal, S. I., and Noma, K. (2010). Centromeric localization of dispersed Pol III genes in fission yeast. *Mol. Biol. Cell* 21, 254–265. doi: 10.1091/mbc.E09-09-0790

Kalapis, D., Bezerra, A. R., Farkas, Z., Horvath, P., Bodi, Z., Daraba, A., et al. (2015). Evolution of robustness to protein mistranslation by accelerated protein turnover. *PLoS Biol.* 13:e1002291. doi: 10.1371/journal.pbio.1002291

Kellis, M., Patterson, N., Endrizzi, M., Birren, B., and Lander, E. S. (2003). Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature* 423, 241–254. doi: 10.1038/nature01644

Krassowski, T., Coughlan, A. Y., Shen, X. X., Zhou, X., Kominek, J., Opulente, D. A., et al. (2018). Evolutionary instability of CUG-Leu in the genetic code of budding yeasts. *Nat. Commun.* 9:1887. doi: 10.1038/s41467-018-04374-7

LaBella, A. L., Opulente, D. A., Steenwyk, J. L., Hittinger, C. T., and Rokas, A. (2019). Variation and selection on codon usage bias across an entire subphylum. *PLoS Genet.* 15:e1008304. doi: 10.1371/journal.pgen.1008304

Li, W., Prazak, L., Chatterjee, N., Gruninger, S., Krug, L., Theodorou, D., et al. (2013). Activation of transposable elements during aging and neuronal decline in *Drosophila*. *Nat. Neurosci.* 16, 529–531. doi: 10.1038/nn.3368

Maxwell, P. H., Burhans, W. C., and Curcio, M. J. (2011). Retrotransposition is associated with genome instability during chronological aging. *Proc. Natl. Acad. Sci. U.S.A.* 108, 20376–20381. doi: 10.1073/pnas.1100271108

Mieczkowski, P. A., Lemoine, F. J., and Petes, T. D. (2006). Recombination between retrotransposons as a source of chromosome rearrangements in the yeast *Saccharomyces cerevisiae*. *DNA Repair* 5, 1010–1020. doi: 10.1016/j.dnarep.2006.05.027

Muhlhausen, S., Findeisen, P., Plessmann, U., Urlaub, H., and Kollmar, M. (2016). A novel nuclear genetic code alteration in yeasts and the evolution of codon

reassignment in eukaryotes. *Genome. Res.* 26, 945–955. doi: 10.1101/gr.200931.115

Noma, K., Cam, H. P., Maraia, R. J., and Grewal, S. I. (2006). A role for TFIIIC transcription factor complex in genome organization. *Cell* 125, 859–872. doi: 10.1016/j.cell.2006.04.028

Osmundson, J. S., Kumar, J., Yeung, R., and Smith, D. J. (2017). Pif1-family helicases cooperatively suppress widespread replication-fork arrest at tRNA genes. *Nat. Struct. Mol. Biol.* 24, 162–170. doi: 10.1038/nsmb.3342

Pan, T. (2013). Adaptive translation as a mechanism of stress response and adaptation. *Annu. Rev. Genet.* 47, 121–137. doi: 10.1146/annurev-genet-111212-133522

Percudani, R., Pavesi, A., and Ottonello, S. (1997). Transfer RNA gene redundancy and translational selection in *Saccharomyces cerevisiae*. *J. Mol. Biol.* 268, 322–330. doi: 10.1006/jmbi.1997.0942

Phizicky, E. M., and Hopper, A. K. (2010). tRNA biology charges to the front. *Genes Dev.* 24, 1832–1860. doi: 10.1101/gad.1956510

Pryce, D. W., Ramayah, S., Jaendling, A., and McFarlane, R. J. (2009). Recombination at DNA replication fork barriers is not universal and is differentially regulated by Swi1. *Proc. Natl. Acad. Sci. U.S.A.* 106, 4770–4775. doi: 10.1073/pnas.0807739106

Raab, J. R., Chiu, J., Zhu, J., Katzman, S., Kurukuti, S., Wade, P. A., et al. (2012). Human tRNA genes function as chromatin insulators. *EMBO J.* 31, 330–350. doi: 10.1038/emboj.2011.406

Raina, M., and Ibba, M. (2014). tRNAs as regulators of biological processes. *Front. Genet.* 5:171. doi: 10.3389/fgene.2014.00171

Riley, R., Haridas, S., Wolfe, K. H., Lopes, M. R., Hittinger, C. T., Goker, M., et al. (2016). Comparative genomics of biotechnologically important yeasts. *Proc. Natl. Acad. Sci. U.S.A.* 113, 9882–9887. doi: 10.1073/pnas.1603941113

Saha, A., Mitchell, J. A., Nishida, Y., Hildreth, J. E., Ariberre, J. A., Gilbert, W. V., et al. (2015). A trans-dominant form of Gag restricts Ty1 retrotransposition and mediates copy number control. *J. Virol.* 89, 3922–3938. doi: 10.1128/JVI.03060-14

Santos, M. A., and Tuite, M. F. (1995). The CUG codon is decoded in vivo as serine and not leucine in *Candida albicans*. *Nucleic Acids Res.* 23, 1481–1486. doi: 10.1093/nar/23.9.1481

Santos-Pereira, J. M., and Aguilera, A. (2015). R loops: new modulators of genome dynamics and function. *Nat. Rev. Genet.* 16, 583–597. doi: 10.1038/nrg3961

Schramm, L., and Hernandez, N. (2002). Recruitment of RNA polymerase III to its target promoters. *Genes Dev.* 16, 2593–2620. doi: 10.1101/gad.1018902

Scott, K. C., Merrett, S. L., and Willard, H. F. (2006). A heterochromatin barrier partitions the fission yeast centromere into discrete chromatin domains. *Curr. Biol.* 16, 119–129. doi: 10.1016/j.cub.2005.11.065

Selmecki, A., Forche, A., and Berman, J. (2010). Genomic plasticity of the human fungal pathogen *Candida albicans*. *Eukaryot Cell* 9, 991–1008. doi: 10.1128/EC.00060-10

Simms, T. A., Dugas, S. L., Gremillion, J. C., Ibos, M. E., Dandurand, M. N., Toliver, T. T., et al. (2008). TFIIIC binding sites function as both heterochromatin barriers and chromatin insulators in *Saccharomyces cerevisiae*. *Eukaryot Cell* 7, 2078–2086. doi: 10.1128/EC.00128-08

Simms, T. A., Miller, E. C., Buisson, N. P., Jambunathan, N., and Donze, D. (2004). The *Saccharomyces cerevisiae* TRT2 tRNAThr gene upstream of STE6 is a barrier to repression in MATalpha cells and exerts a potential tRNA position effect in MATa cells. *Nucleic Acids Res.* 32, 5206–5213. doi: 10.1093/nar/gkh858

Su, Z., Wilson, B., Kumar, P., and Dutta, A. (2020). Noncanonical roles of tRNAs: tRNA fragments and beyond. *Annu. Rev. Genet.* 54, 47–69. doi: 10.1146/annurev-genet-022620-101840

Suzuki, T., Ueda, T., and Watanabe, K. (1997). The 'polysemous' codon–a codon with multiple amino acid assignment caused by dual specificity of tRNA identity. *EMBO J.* 16, 1122–1134. doi: 10.1093/emboj/16.5.1122

Thompson, M., Haeusler, R. A., Good, P. D., and Engelke, D. R. (2003). Nucleolar clustering of dispersed tRNA genes. *Science* 302, 1399–1401. doi: 10.1126/science.1089814

Todeschini, A. L., Morillon, A., Springer, M., and Lesage, P. (2005). Severe adenine starvation activates Ty1 transcription and retrotransposition in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 25, 7459–7472. doi: 10.1128/MCB.25.17.7459-7472.2005

Torrent, M., Chalancon, G., de Groot, N. S., Wuster, A., and Madan Babu, M. (2018). Cells alter their tRNA abundance to selectively regulate protein synthesis during stress conditions. *Sci. Signal.* 11:eaat6409. doi: 10.1126/scisignal.aat6409

Tran, P. L. T., Pohl, T. J., Chen, C. F., Chan, A., Pott, S., and Zakian, V. A. (2017). PIF1 family DNA helicases suppress R-loop mediated genome instability at tRNA genes. *Nat. Commun.* 8:15025. doi: 10.1038/ncomms15025

Wahba, L., Costantino, L., Tan, F. J., Zimmer, A., and Koshland, D. (2016). S1-DRIP-seq identifies high expression and polyA tracts as major contributors to R-loop formation. *Genes Dev.* 30, 1327–1338. doi: 10.1101/gad.280834.116

Warner, J. R. (1999). The economics of ribosome biosynthesis in yeast. *Trends Biochem. Sci.* 24, 437–440. doi: 10.1016/s0968-0004(99)01460-7

Yeung, R., and Smith, D. J. (2020). Determinants of replication-fork pausing at tRNA genes in *Saccharomyces cerevisiae*. *Genetics* 214, 825–838. doi: 10.1534/genetics.120.303092

Yona, A. H., Bloom-Ackermann, Z., Frumkin, I., Hanson-Smith, V., Charpak-Amikam, Y., Feng, Q., et al. (2013). tRNA genes rapidly change in evolution to meet novel translational demands. *Elife* 2:e01339. doi: 10.7554/eLife.01339

Yuan, G. C., Liu, Y. J., Dion, M. F., Slack, M. D., Wu, L. F., Altschuler, S. J., et al. (2005). Genome-scale identification of nucleosome positions in *S. cerevisiae*. *Science* 309, 626–630. doi: 10.1126/science.1112178

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 21 May 2021
doi: 10.3389/fmicb.2021.635070

Prediction of Novel Bacterial Small RNAs From RIL-Seq RNA–RNA Interaction Data

*Amir Bar, Liron Argaman, Yael Altuvia and Hanah Margalit**

Department of Microbiology and Molecular Genetics, Institute for Medical Research Israel-Canada, Faculty of Medicine, The Hebrew University of Jerusalem, Jerusalem, Israel

The genomic revolution and subsequent advances in large-scale genomic and transcriptomic technologies highlighted hidden genomic treasures. Among them stand out non-coding small RNAs (sRNAs), shown to play important roles in post-transcriptional regulation of gene expression in both pro- and eukaryotes. Bacterial sRNA-encoding genes were initially identified in intergenic regions, but recent evidence suggest that they can be encoded within other, well-defined, genomic elements. This notion was strongly supported by data generated by RIL-seq, a RNA-seq-based methodology we recently developed for deciphering chaperon-dependent sRNA-target networks in bacteria. Applying RIL-seq to Hfq-bound RNAs in *Escherichia coli*, we found that ∼64% of the detected RNA pairs involved known sRNAs, suggesting that yet unknown sRNAs may be included in the ∼36% remaining pairs. To determine the latter, we first tested and refined a set of quantitative features derived from RIL-seq data, which distinguish between Hfq-dependent sRNAs and "other RNAs". We then incorporated these features in a machine learning-based algorithm that predicts novel sRNAs from RIL-seq data, and identified high-scoring candidates encoded in various genomic regions, mostly intergenic regions and 3′ untranslated regions, but also 5′ untranslated regions and coding sequences. Several candidates were further tested and verified by northern blot analysis as Hfq-dependent sRNAs. Our study reinforces the emerging concept that sRNAs are encoded within various genomic elements, and provides a computational framework for the detection of additional sRNAs in Hfq RIL-seq data of *E. coli* grown under different conditions and of other bacteria manifesting Hfq-mediated sRNA-target interactions.

Keywords: sRNA (small RNA), RIL-seq, prediction, *E. coli* – *Escherichia coli*, post-transcriptional regulation, Hfq

OPEN ACCESS

Edited by:
Jiqiang Ling,
University of Maryland, United States

Reviewed by:
Erik Holmqvist,
Uppsala University, Sweden
Carin Vanderpool,
University of Illinois at Urbana-Champaign, United States
Lydia Contreras,
University of Texas at Austin, United States

***Correspondence:**
Hanah Margalit
hanahm@ekmd.huji.ac.il

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *29 November 2020*
Accepted: *06 April 2021*
Published: *21 May 2021*

Citation:
Bar A, Argaman L, Altuvia Y and Margalit H (2021) Prediction of Novel Bacterial Small RNAs From RIL-Seq RNA–RNA Interaction Data. Front. Microbiol. 12:635070. doi: 10.3389/fmicb.2021.635070

INTRODUCTION

Trans-acting small RNAs (sRNAs) have emerged as a major class of post-transcriptional gene expression regulators in bacteria. These are short RNA molecules, 50–400 nucleotides long, which regulate their targets in trans, usually by incomplete base pairing with their mRNAs, affecting translation and/or mRNA stability (Wagner and Romby, 2015; Hör et al., 2020). sRNAs were discovered in many bacteria and were shown to play regulatory roles in diverse cellular processes, and in particular in the response to various stress conditions. Often, these RNA regulators are associated with chaperon proteins, such as Hfq (Vogel and Luisi, 2011) or ProQ (Melamed et al., 2020).

In many Gram-negative bacteria the protein chaperon Hfq mediates many of the sRNA-target interactions and stabilizes the sRNAs (Vogel and Luisi, 2011; De Lay et al., 2013; Updegrove et al., 2016; Santiago-Frangos and Woodson, 2018). Yet, there are sRNAs in *Escherichia coli* for which it was suggested that their RNA-binding activity is Hfq-independent (Mihailovic et al., 2018). In the present study, we focus on Hfq-dependent sRNAs in *E. coli*.

While the initial discovery of the first sRNA in *E. coli*, Spot 42, is dated to 1973 (Ikemura and Dahlberg, 1973a,b) and a few other sRNAs were discovered serendipitously along the years [e.g., MicF (Mizuno et al., 1984), DsrA (Sledjeski and Gottesman, 1995), OxyS (Altuvia et al., 1997)], their big burst occurred following the genomic revolution in the mid-1990s. The completion of the genome sequencing of *E. coli* inspired several systematic computational-experimental expeditions, attempting to identify additional sRNA-encoding genes based on the genome information. As all previously known sRNAs were encoded by genes located between two protein coding genes, the initial screens were focused at intergenic regions and identified novel sRNA-encoding genes only in those regions (Argaman et al., 2001; Rivas et al., 2001; Wassarman et al., 2001; Chen et al., 2002). Yet, subsequent experimental screens of RNAs bound to Hfq, carried out in several bacterial species, revealed putative Hfq-bound sRNAs encoded in various genomic regions, including coding sequences (CDS) and 5′ and 3′ untranslated regions (UTR) (Zhang et al., 2003; Chao et al., 2012; Bilusic et al., 2014; Tree et al., 2014; Huber et al., 2020). These sRNAs may be either independently transcribed, or processed from mRNAs by endoribonucleases (Miyakoshi et al., 2015b). When processed from the mRNA of their hosting gene they often regulate genes involved in the same pathways as the parent gene and may generate efficient regulatory circuits [e.g., CpxQ and *cpxP*, and GadF and *gadE* (Chao and Vogel, 2016; Grabowicz et al., 2016; Melamed et al., 2016)].

The discovery of novel Hfq-bound sRNAs that are encoded within a variety of genomic elements was enhanced by RIL-seq (RNA Interaction by Ligation and sequencing), a high-throughput methodology we recently developed for mapping direct RNA–RNA interactions mediated by Hfq (Melamed et al., 2016, 2018). The idea behind RIL-seq is that a sRNA and a target RNA co-bound to Hfq could be ligated and then identified by sequencing as chimeric fragments. The major steps of RIL-seq involve *in vivo* protein-RNA crosslinking, co-immunoprecipitation of Hfq and bound RNAs, RNA ligation and paired-end RNA sequencing. Interacting pairs are identified by mapping the ends of sequenced fragments to the genome and identifying chimeric fragments in which the two ends are mapped to two different genomic locations. Only chimeric fragments whose abundance exceeds random expectation are kept and considered as representing RNA interacting pairs (statistically significant chimeras, hereinafter, S-chimeras). Application of RIL-seq to *E. coli* grown to exponential phase, stationary phase and exponential phase under iron limitation revealed ∼2800 RNA–RNA interactions, ∼64% of which involved well-established sRNAs and the rest involved RNAs derived from various genomic entities (Melamed et al., 2016). Interestingly, in most of the chimeric fragments including known sRNAs, the sRNA was the second RNA in the chimera (at the 3′ part of the chimeric RNA). This regarded both class I and class II sRNAs (Schu et al., 2015). The positioning of the sRNAs as second in the chimeras is consistent with the known binding mode of many sRNAs within Hfq, where the uridine-rich 3′ terminus of the sRNA (hereinafter, U-tract) is bound by Hfq (Otaka et al., 2011; Sauer and Weichenrieder, 2011; Dimastrogiovanni et al., 2014; Schu et al., 2015). RIL-seq involves, prior to the ligation of Hfq-bound RNAs, a step where RNA regions that are not protected by Hfq or by base pairing are trimmed by riboendonucleases and treated with polynucleotide kinase, generating 5′P end of the sRNA. This 5′P end is accessible to the ligase, resulting in chimeras where the sRNA is the second RNA. In fact, this finding provided further support to the suggested binding mode of sRNAs on Hfq (Dimastrogiovanni et al., 2014), by the identification of a common motif in the second RNAs of RIL-seq chimeras, comprising a GC-rich sequence followed by a U tail (Melamed et al., 2016), compatible with a transcription terminator. In addition, Holmqvist et al. (2016) identified a similar motif in mRNA 3′ UTR sequences bound by Hfq in *Salmonella*. The observation that sRNAs are often second in their respective chimeric fragments has raised the intriguing conjecture that the second RNAs in chimeric fragments that do not contain known sRNAs may be novel sRNAs. Furthermore, many of the RNAs found at the 3′ part of the chimeric fragments (second RNAs of the chimeras) were derived from intergenic regions or from 3′ UTRs, underpinning their potential as novel sRNAs. Indeed, some of these second RNAs, such as those embedded in the 3′ UTR of *cutC* and in the 3′ UTR of *cpxP* were identified in independent studies as sRNAs (Guo et al., 2014; Chao and Vogel, 2016).

In total, RIL-seq data comprised ∼1000 RNA–RNA pairs that did not include a known sRNA (Melamed et al., 2016), suggesting that they may include yet unknown sRNAs. To identify novel sRNAs systematically, we characterized the RNAs in all RIL-seq chimeras by various features inferred from the data and from their sequences. The distributions of several of these features, such as the number of unique interactions that a RNA is involved in, were found to differ statistically significantly between known sRNAs and "other RNAs", reaffirming them as informative features. Here, we describe and discuss the set of informative features of sRNAs as well as a predictive algorithm utilizing them, provide a list of potential novel sRNAs and report the experimental verification of novel sRNAs encoded in intergenic regions within operons, in 5′ and 3′ UTRs and within the coding sequence. Our computational and experimental results support the expanding concept that there is a reservoir of sRNAs encoded within a variety of genomic entities and expressed under various conditions (Adams and Storz, 2020; Adams et al., 2021). The computational framework that we provide for analysis of Hfq RIL-seq data can be used to identify novel sRNAs in RIL-seq data generated for *E. coli* grown under additional cellular conditions and in RIL-seq data generated for other bacteria manifesting Hfq-mediated sRNA-target interactions. It may also inspire the application of similar algorithms for analysis of large-scale data generated by equivalent protocols in other contexts.

MATERIALS AND METHODS

Computational Analysis

Data

We used three data sets of chimeric fragments corresponding to S-chimeras, obtained in RIL-seq experiments applied to bacteria grown to exponential phase, to stationary phase and to exponential phase under iron limitation (Melamed et al., 2016). The data set of exponential growth phase was obtained from six biological replicates of the experiment, while the data sets of the stationary phase and growth under iron limitation were obtained from three biological replicates in each condition. Each RNA in the data was annotated as either "known sRNA" or "other RNA" (**Supplementary Table 1**). We included in the set of known sRNAs all RNAs that were annotated as sRNAs prior or in parallel to RIL-seq publication (Melamed et al., 2016). The latter regard CpxQ (Chao and Vogel, 2016), SroC (Miyakoshi et al., 2015a) and 3′ETS-leuZ (Lalaouna et al., 2015). Any RNA that is not a known sRNA was annotated as "other RNA". The total numbers of known sRNAs and "other RNAs" in each group within each of the three data sets is summarized in **Table 1**.

Selecting Features Distinguishing sRNAs From "Other RNAs"

We describe each RNA by features mainly extracted from RIL-seq data (**Supplementary Table 2**) and compare their distributions between the groups of sRNAs and "other RNAs" by Mann–Whitney U test (with Bonferroni correction for multiple hypotheses testing). For features that differ statistically significantly between the two groups we compute the Pearson correlation coefficient between every pair of features, cluster the features based on their correlation coefficients, and select one of the features in a cluster as representative. As the data corresponding to a RNA in one RIL-seq experiment (e.g., exponential phase) may differ from the data corresponding to this RNA in another RIL-seq experiment (e.g., stationary phase), all analyses were carried out separately for each data set. We verified that the selected features were found to statistically significantly differ between the group of known sRNAs and the group of "other RNAs" in all data sets. These selected features were used in the successive analyses.

Predicting the Probability of a RNA to Be a sRNA

The development of the predictive scheme was carried out separately for each data set. Each RNA in the data was described by a vector of the selected features. The data set was split in a ratio of 2:1 into a training set and a test set, respectively, where in each set the ratio of known sRNAs to "other RNAs" was maintained (**Table 1**). We applied logistic regression (python sklearn module) to the training set. The logistic regression provides weights to the different features (β_i) and an intercept (β_0), such that $l = \beta_0 + \sum_{i=1}^{n} \beta_i \cdot x_i$, where n is the number of selected features and x_i is the value of feature i. The probability of a RNA to be a sRNA is then computed as $1/(1 + \exp(-l))$. We tested the obtained logistic regression model by applying it to RNAs in the test set. In practice, we conducted 10,000 iterations of this procedure, and recorded the probabilities a RNA obtained when it was included in the test set of an iteration. The final predicted probability of each RNA was computed as the mean of the predicted probabilities across all the iterations in which it was included in the test set. The logistic regression was trained using the default parameters of the sklearn linear_model LogisticRegression class, i.e., using L2 regularization.

TABLE 1 | Number of sRNAs and "other RNAs" in the various data sets.

Condition/growth phase	Number of known sRNAs	Number of "other RNAs"
Exponential phase	26	751
Stationary phase	29	1201
Exponential phase under iron limitation	29	1248

Computation of Feature Contribution

It is common to examine the weights in order to learn on the relative contributions of the various features to the computed probability. However, as the values of the different features span different numeric scales, comparison of the weights *per se* is not informative. Instead, we can transform the feature values into z-scores, and compare the products of weight and the feature standard deviation:

$$\begin{aligned}\beta_0 + \sum_{i=1}^{n} \beta_i \cdot x_i &= \beta_0 + \sum_{i=1}^{n} \beta_i \cdot (x_i + m_i - m_i) \\ &= \left(\beta_0 + \sum_{i=1}^{n} \beta_i \cdot m_i\right) + \sum_{i=1}^{n} \beta_i \cdot (x_i - m_i) \\ &= \left(\beta_0 + \sum_{i=1}^{n} \beta_i \cdot m_i\right) + \sum_{i=1}^{n} \beta_i \cdot \frac{s_i}{s_i} \cdot (x_i - m_i) \\ &= \left(\beta_0 + \sum_{i=1}^{n} \beta_i \cdot m_i\right) + \sum_{i=1}^{n} \beta_i \cdot s_i \cdot zscore(x_i)\end{aligned}$$

where m_i and s_i are the mean and the standard deviation, respectively, of the RNA's i^{th} feature values. The equation shows that transforming the data to z-scores is associated by an appropriate change of the intercept by the weighted sum of the mean feature values, and the weights of the features are represented by the products of the original weight and standard deviation of each feature, which are comparable. In practice, we applied this transformation to the average coefficients from the 10,000 logistic regression iterations we conducted per growth condition.

Principal Component Analysis (PCA)

Feature vectors of the RNAs were initially scaled with python's sklearn preprocessing module using the *robust_scale* function. Then, we applied the PCA transformation for the first two dimensions using sklearn decomposition PCA class.

5′ and 3′ Boundaries of sRNA Transcripts

5′ and 3′ transcript boundaries of the recently published sRNAs relied on the original papers reporting them (**Table 2A**). 5′ and

TABLE 2 | Novel sRNAs.

A. Novel sRNAs predicted based on RIL-seq results and recently reported in published papers

Novel sRNA	Hosting gene or operon	Genomic region	Genomic position	Number of unique targets	Prediction score (probability of being a sRNA)[a]			Comment/ References
					Exponential	Stationary	Iron limitation	
FlgO	*flgL*	3′ UTR	1,140,986 → 1,141,063	4	0.073	0.087	0.152	Hör et al. (2020)[b]
FliX	*FliC*	3′ UTR	2,001,912 ← 2,002,106	17	0.859	0.272	0.672	Hör et al. (2020)[b]
GadF	*gadE*	3′ UTR	3,658,992 → 3,659,082	24	–	0.684	–	Melamed et al. (2016)
MalH	*malG*	3′ UTR	4,242,531 ← 4,242,629 or 4,242,531 ← 4,242,633[c]	38	0.416	0.468	0.596	Iosub et al. (2020)
MotR	*motA*	5′ UTR	1,977,208 ← 1,977,300	19	0.562	0.158	0.444	Hör et al. (2020)
NarS-L	*narK*	3′ UTR	1,279,286 → 1,279,520	8	0.376	–	0.380	Wang et al. (2020)
NarS-S			1,279,337 → 1,279,520					
PspH	*pspG*	3′ UTR	4,263,139 → 4,263,249	1	0.003	0.002	0.001	Melamed et al. (2016)
RaiZ	*raiA*	3′ UTR	2,737,381 → 2,737,542	29	0.337	0.482	0.296	Smirnov et al. (2017)
RaiZ-S			2,737,417 → 2,737,542					
RbsZ	*rbsB-rbsK*	Intergenic in operon	3,937,045 → 3,937,278	4	0.091	0.035	0.105	Melamed et al. (2020)
SdhX (RybD)	*sucD*	3′ UTR	765,050 → 765,150	33	0.638	0.836	0.735	De Mets et al. (2019); Miyakoshi et al. (2019)
UhpU	*uhpT*	3′ UTR	3,845,730 ← 3,845,995	111	0.751	0.347	0.786	Hör et al. (2020)[b]

B. Novel sRNA candidates predicted based on RIL-seq results and verified by northern blot analysis in the current study

AceK-int	*aceK*	CDS	4,218,879 → 4,218,963	15	–	0.592	0.084	Recently verified also by Adams et al. (2021)
AllZ	*allR*	3′ UTR	533,629 → 533,863	5	0.042	0.295	0.046	
BhsB	*bhs*A	3′ UTR	1,169,303 → 1,169,402	2	–	0.099	–	
FadZ	*fadA*	3′ UTR	4,027,232 ← unknown	12	0.031	0.202	0.081	
KilS	*kilR*	5′ UTR	1,418,405 ← 1,418,502	3	0.033	0.154	0.032	
XylZ	*xylA-xylB*	Intergenic in operon	3,729,386 ← 3,729,545	10	0.064	0.346	0.068	
ZbiJ	*ybiJ*	3′ UTR	837,435 ← 837,531	21	0.223	0.374	0.506	Recently verified also by Han and Lory (2021), who called it "asYbiE"

[a] *Dashed cells mean the RNA was not included in the data of the corresponding experiment.*
[b] *Referring to unpublished results from Storz's lab.*
[c] *Mapping of the 5′ end is according to Iosub et al. (2020). According to our RNA-seq results the 5′ end is at 4,242,609.*

3′ transcript boundaries of the new sRNAs predicted here were determined based on the read coverage in corresponding RNA-seq libraries (see below). For *fadA* 3′ UTR we were not able to determine the 5′ end (marked unknown in **Table 2B**) and estimated its 5′ end position based on the size of the band observed in the Hfq-dependent northern blot (see below).

Identification of Transcription Start Sites and RNase E Cleavage Sites Near Predicted sRNAs

In order to appreciate if the transcripts of the novel sRNAs were generated by independent transcription or by cleavage of the hosting mRNA, we used published data of large-scale screens of transcription start sites (TSSs) (Thomason et al., 2015; Ju et al., 2019) and RNase E cleavage sites (Clarke et al., 2014), and searched for TSSs and cleavage sites located between the determined 3′ end and up to 50 nucleotides upstream the 5′ end.

Identification of Putative sRNAs in Hfq-CLASH Data

To verify whether the putative sRNAs we report are supported by other data sets, we compared their estimated coordinates to chimeric fragments included in the Hfq-CLASH data set (Iosub et al., 2020). We considered a RNA as found in Hfq-CLASH chimera if its estimated coordinates were at most 50-nt apart from the coordinates reported in Iosub et al. (2020).

Experimental Testing

Strains and Growth Conditions

For the verification of novel sRNA expression, cultures of *Escherichia coli* MG1655 and its isogenic strain MG1655 *hfq::Kn* were grown over-night in LB medium and then diluted 1:100 in fresh LB medium and grown while shaking at 37°C. Samples of culture were collected throughout growth, and centrifuged at 4°C. The pelleted cells were resuspended in 50 μl of TE buffer (10 mM Tris HCl pH 8.0, 1 mM EDTA pH 8.0), mixed with lysozyme to a final concentration of 0.9 mg/ml and fast frozen in liquid nitrogen. The samples were then subjected to two cycles of thawing at 37°C and freezing in liquid nitrogen.

Northern Analysis

Total RNA was extracted from harvested cells using TRI-reagent (Sigma). 30 μg of total RNA were separated in 7 M urea/6% polyacrylamide gels in 44.5 mM Tris-base, 44.5 mM boric acid and 2 mM EDTA pH 8.0, and transferred to Zeta-Probe membrane (Bio-Rad) by electroblotting. The membranes were hybridized with specific [^{32}P] end labeled DNA probes. For each tested sRNA, the northern blot was repeated at least twice, with a different replicate of total RNA. The probe sequences are listed in **Supplementary Table 3**.

RNA-Seq

We used compatible RNA-seq data available in the lab, which were generated as following: Three single colonies of MG1655 cells carrying a pJV300 plasmid (Urban and Vogel, 2007) were grown over night at 37°C in LB medium supplied with Ampicillin (100 μg/ml). The cultures were diluted 1:100 in fresh medium and grown while shaking at 37°C for 6 h. Cells were collected and RNA was extracted as described above. RNA-seq libraries were constructed according to the RNAtag-seq protocol (Shishkin et al., 2015), with few modifications described in Melamed et al. (2018). The libraries were paired-end sequenced using Illumina NextSeq 500 machine, with read length of 45 and 40 bp for first and second read, respectively. Raw reads were split into their original three replicate libraries using an in-house script. Cutadpat was applied to remove adapter sequences, low quality ends and sequences shorter than 25 nucleotides (Martin, 2011). We applied bwa aln followed by bwa sampe (Li, 2013) to align the reads to the genome. We applied stringent mapping allowing only two mismatches. The total number of reads in the libraries of the three replicates 1, 2, 3 was 10674656, 16274638, 10368982 reads, respectively. In all three libraries 99% of the reads passed the processing filter and 89% of the processed reads were successfully mapped. Library 2, which had the highest number of reads, was used to define the novel sRNA boundaries.

RESULTS

Examination of the chimeric fragments corresponding to S-chimeras in RIL-seq data hinted at several properties that may aid in the classification of RNAs represented in these chimeras as either sRNA or target RNA (Melamed et al., 2016). sRNAs were included in many chimeric fragments, were found to interact with multiple targets and were preferentially identified as the second RNA in the chimeric fragments. In contrast, RNAs found in interaction with a known sRNA were usually found to be involved in a small number of chimeric fragments, were found to interact with only a few partners (mainly, the sRNA), and were frequently identified as the first RNA in the chimeric fragments.

While previously these properties were intuitively considered for supporting or rejecting a sRNA candidate (Melamed et al., 2016), our aim here is to quantify them and carry out a systematic analysis, selecting informative features that will be incorporated in a sRNA predictive scheme. The features that we propose to examine for each RNA are of two types: (*i*) features derived from the chimeric fragments the RNA is involved in (first layer), and (*ii*) features of the RNA interactors (second layer). The incorporation of both layers of features in the analysis is inspired by the acknowledgment that up to date the number of identified sRNA–sRNA interactions is very small and far below the number of identified sRNA interactions with mRNAs. Recognizing first layer features that support the RNA as a sRNA along with second layer features that do not support the interactors as sRNAs should provide stronger support for a sRNA candidate than expected from its first layer features alone. Combining the two layers of traits is expected to enhance the discriminative power of the model and to increase the reliability of predicted sRNAs.

Feature Selection

All analyses described hereinafter were carried out for each RIL-seq data set separately (exponential phase, stationary phase, exponential phase under iron limitation). We describe in the text the results for the stationary phase data and in the Supplementary Material the results for the two other data sets. When relating

to the chimeric fragments, we refer to chimeric fragments corresponding to S-chimeras identified in RIL-seq results.

Some of the traits characterizing a RNA can be quantitatively described in several ways (hereinafter, features). For example, let X be a RNA that was identified as interacting with k RNAs $Y_1, ..., Y_i, ..., Y_k$ and is involved in $n_1, ..., n_i, ..., n_k$ chimeric fragments corresponding to each RNA, respectively, making up a total of N chimeric fragments. The trait 'number of chimeric fragments the RNA is involved in' can be described as N, or as the mean of n_i, or as the median of n_i. We examined 18 features in total, several of which regard different representations of the same trait (**Supplementary Table 2**). We assigned each RNA the values of the features. For each feature we compared the distributions of its values between the group of known sRNAs and "other RNAs" by two-tailed Mann–Whitney U test with Bonferroni correction for multiple hypotheses testing (**Supplementary Material** and **Supplementary Table 2**). We then clustered all features that differed statistically significantly between the group of known sRNAs and group of "other RNAs" (**Supplementary Material** and **Supplementary Figure 1**), and selected from each cluster of features one representative feature (usually the one with simplest intuitive interpretation) to be used in successive analyses. In addition to features solely based on RIL-seq data, we also included the length of the U-tract of the RNA, as we previously observed that sRNAs have longer U-tracts at their terminators compared to "other RNAs" (Melamed et al., 2016). The U-tract length is also considered a first-layer feature, as it is a feature of the RNA itself.

Six features were selected, four are first layer features and two are second layer features (**Figure 1**). It is of note that these six features were consistently selected in the analyses of all three data sets. For each RNA these features are: (A) *Total number of chimeric fragments*: The total number of chimeric fragments that included the RNA. This value was normalized by the total number of chimeric fragments in the data set. (B) *Number of unique interactions*: Number of unique interactions the RNA was involved in (k). This value was normalized by the total number of unique interactions in the data set. (C) *Second-In-Chimera (SIC) score*: A score representing the fraction of chimeric fragments in which the RNA was the second RNA of the chimera, while taking into account the number of unique interactions this RNA is involved in. We defined this score as $S - \frac{1}{k}$, where S is the fraction of chimeric fragments in which the RNA was the second RNA of the chimera and k is as defined above. Intuitively, for RNAs with many interactions the score is approximately S, while the score of RNAs with a small number of interactions is penalized to prevent high SIC scores that are based on one or only a few interactions. Note that due to this correction SIC may also get negative values. (D) *U-tract length*: For each RNA we assigned the length of the longest U-tract that could be identified in a region spanning 50 nucleotides around the segments of this RNA in the chimeric fragments of the RIL-seq data. (E) *Median number of interactions of interactors*: Each interactor of a RNA is also annotated by feature B. We take the median of these values across all interactors of the RNA. (F) *Median SIC score of interactors*: Each interactor has a SIC score, as defined above. We take the median of these values across all interactors of the RNA. Note that since the variances of features A, B, and E were extremely large, their values were transformed to Log_{10} scale for further analysis.

The distributions of these feature values differed statistically significantly (after Bonferroni correction for multiple hypotheses testing) between the groups of known sRNAs and "other RNAs" (**Figure 2** and **Supplementary Figures 2**, **3**). It is evident from **Figure 2** that compared to the "other RNAs", the known sRNAs were found to be involved in more chimeric fragments and in more unique interactions; the fraction of the chimeric fragments in which they appear as second RNA is higher; and they have longer U-tracts. As for the interactors of sRNAs, the fraction of interactions in which they are second RNA in the chimera and the number of unique interactions they are involved in are lower compared to interactors of "other RNAs". The statistically significant differences between the distributions of the features in the two RNA groups (p values between 10^{-23} and 10^{-11}) suggested that they can be used for classifying sRNAs and for the determination of novel, yet unknown sRNAs, which may be hidden in RIL-seq data.

Each RNA in RIL-seq data was represented by a vector of the above six features. Analysis of the vectors by principal component analysis (PCA) further demonstrated the separation of known sRNAs from "other RNAs" by the features and the contribution of the various features to this separation (**Figures 3A,B** and **Supplementary Figures 4A,B, 5A,B**). Furthermore, this analysis showed additional RNAs in close proximity to the previously known sRNAs, suggesting these might be novel sRNA candidates. Intriguingly, several novel sRNAs derived from 3′ UTRs, which were recently verified experimentally (**Table 2A**), are clustered together with known sRNAs in the PCA plots.

Prediction of Novel sRNAs

To systematically and comprehensively identify novel sRNA candidates, we applied logistic regression, using these six features as characteristics of each RNA in the data. We applied 10,000 iterations of the logistic regression, where in each iteration we randomly split the data into a training set and a test set. Each training set included 2/3 of the known sRNAs and 2/3 of the "other RNAs" and the test set included the rest of the data. At each iteration we trained a logistic regression model on the training set, resulting in a linear combination of the features, which provides the probability of a RNA in the data to be a sRNA (see the section "Materials and Methods"). The model was then used to compute these probabilities for RNAs in the test set. At the end of the process, each RNA had a list of N probabilities, where N is the number of test sets that included this RNA. The final sRNA probability of a specific RNA was the average of these probabilities, considered hereinafter as the sRNA score of the RNA. Determining different probability thresholds above which a RNA is determined as a sRNA, we obtained a receiver operating characteristic (ROC) curve and a precision–recall (PR) curve for each iteration and for the average results (**Figures 3C,D** and **Supplementary Figures 4C,D, 5C,D**), showing the consistency and high predictive power provided by the logistic regression.

Due to the low frequency of sRNAs in the data and the inclusion of sRNAs not yet discovered in the training set, we expect the model to output uncalibrated prediction probabilities.

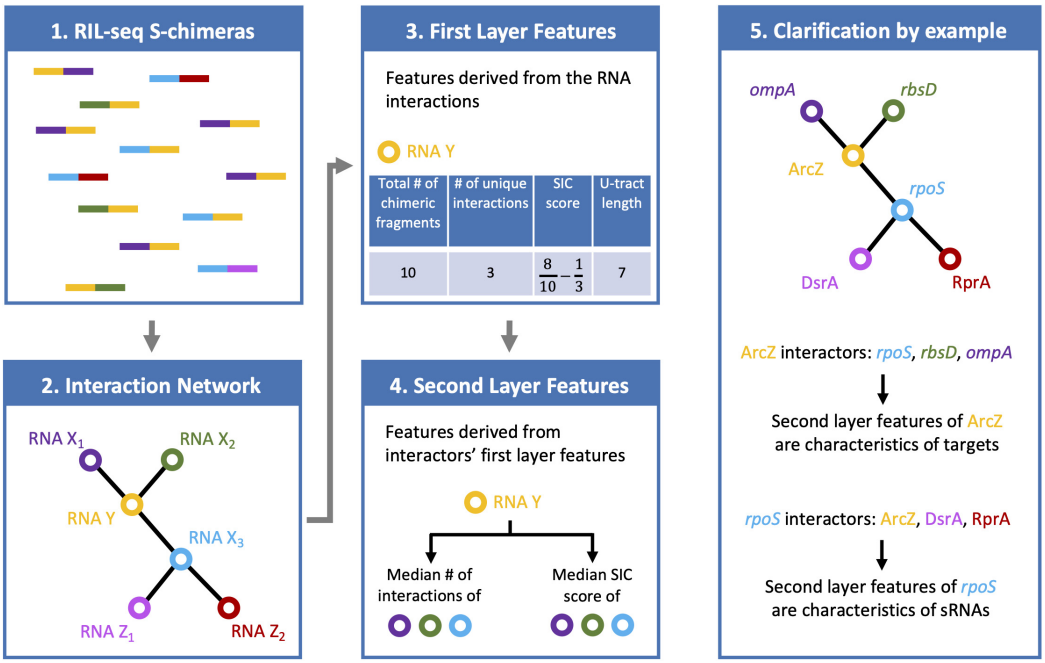


FIGURE 1 | Features characterizing a RNA in RIL-seq data. RIL-seq final data set includes chimeric fragments corresponding to S-chimeras (first box), determining RNA interacting pairs, which can be described as a RNA–RNA interaction network (second box). Each node in the network represents a RNA. Two nodes are connected by an edge if they were determined as interacting. For every RNA, first layer features were computed based on the chimeric fragments it was involved in (third box; U-tract length is not shown on the chimeric fragments). Finally, for each RNA, second layer features were computed based on the first layer features of its interacting partners (forth box). If the RNA is a sRNA, its first layer features are expected to be characteristic of a sRNA, namely, many chimeric fragments, many interactions, high fractions of chimeras in which it is second RNA (high SIC, Second-In-Chimera, score), long U-tract. The second layer features of the sRNA, which involve the first-layer features of its interactors, are expected to be characteristic of targets, namely, low number of interactions and low SIC scores (fifth box).

We therefore did not determine a probability threshold above which a RNA is predicted as a sRNA, but ranked the RNAs by their sRNA scores, scanned the ranked RNAs from top down and searched for RNAs ranked above or in the vicinity of known sRNAs (**Figure 3E** and **Supplementary Figures 4E, 5E**). As the logistic regression was performed for each RIL-seq data set separately, the ranking of a specific RNA can change between conditions. This stems from the fact that both the feature vectors of RNAs and the annotated sRNAs that are included in a data set are condition specific. This implies a RNA can be predicted as a sRNA under one condition, but not necessarily under another condition, consistent with the acknowledged condition-specific expression of sRNAs (Wagner and Romby, 2015). In fact, we find an association between the change in expression levels of the sRNAs between conditions and the differences in their sRNA score between the corresponding conditions (**Supplementary Figure 6**). This implies a relationship between the expression level of the sRNA and its sRNA score per condition. For example, SroC, known to be expressed in stationary phase (Miyakoshi et al., 2015a), got a sRNA score of 0.5 in stationary phase data, but scores of 0.004 and 0.003 in the data sets of exponential phase and exponential phase under iron limitation, respectively. Encouragingly, many of the known sRNAs have obtained high ranking scores in at least one data set (**Supplementary Table 1**). Thus, the implementation of the selected features of the RNAs in a machine learning approach, such as the logistic regression, enables the distinction of sRNAs from "other RNAs", and therefore may enable the discovery of novel sRNAs. Notably, we identified most of the recently discovered sRNAs (that were not annotated as such in our data) among the top ranking RNAs (**Table 2A**), as well as additional novel sRNAs, some of which (**Table 2B**) we tested experimentally, as detailed below.

To verify that our results do not depend on the number of iterations or the selected ratio of 2:1 between the sizes of the training and test sets, we conducted the analyses for different numbers of iterations and different ratios of training to test set sizes. These analyses confirmed that the results are independent of these parameters (**Supplementary Material** and **Supplementary Figures 7–9**).

Contribution of Individual Features to the Classification

The logistic regression assigns weights to the features, which are used for the computation of the probability of a RNA in the data to be a sRNA (see the section "Materials and Methods," **Table 3**, and **Supplementary Table 4**). As shown in **Table 3** and **Supplementary Table 4**, the various features differ in their contributions to the predicted probability. First, since the first-layer features directly assess a RNA as a sRNA and the second-layer features are expected to contribute to the prediction by rejecting its interactors as sRNAs (**Figure 1**), it is affirmative

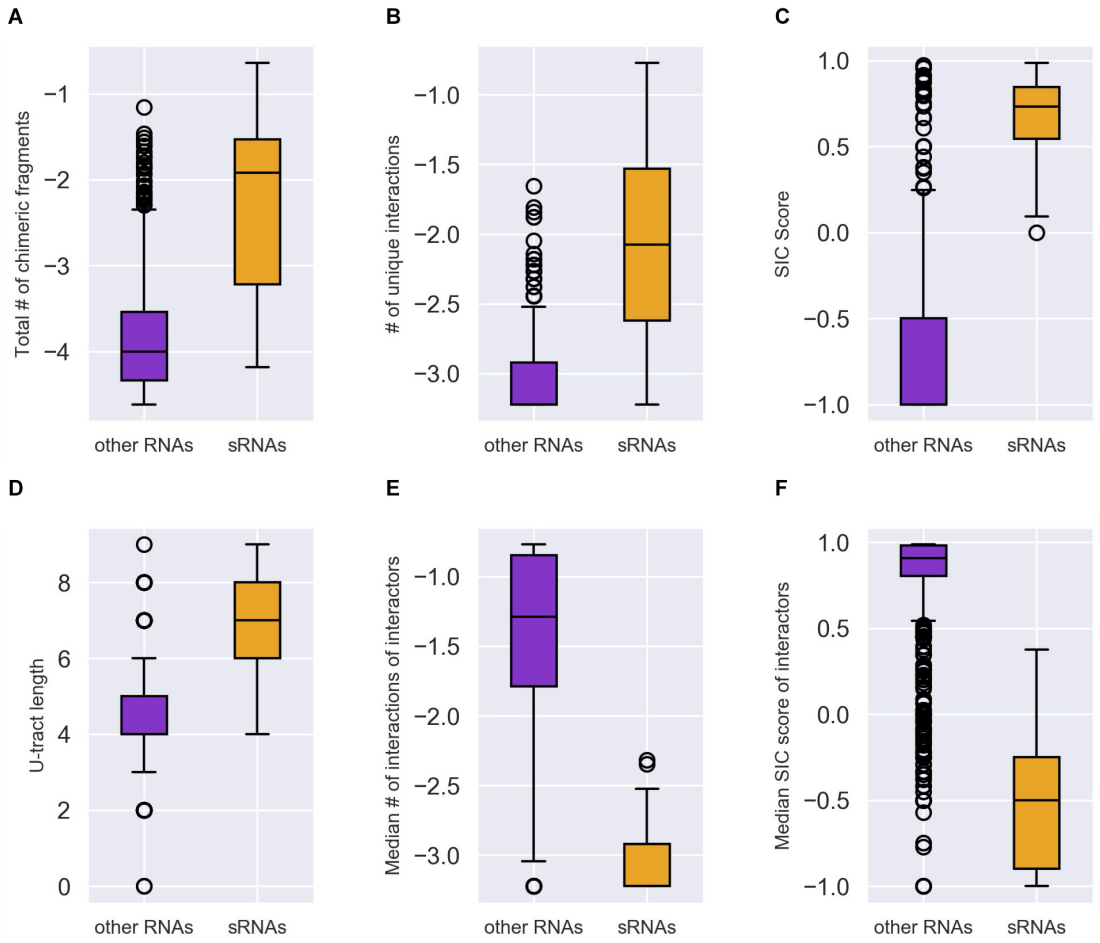


FIGURE 2 | Distinction between sRNAs and "other RNAs" by the characteristic features. Each RNA was characterized by the following features: **(A)** Total number of the chimeric fragments the RNA is involved in. This value was normalized by the total number of chimeric fragments in the data set and then transformed to Log_{10} scale. **(B)** Number of unique interactions the RNA is involved in. This value was normalized by the total number of interactions in the data set and then transformed to Log_{10} scale. **(C)** SIC score (SIC, for Second-In-Chimera), namely the percentage of chimeric fragments in which the RNA was second in the chimera, penalized by the number of unique interactions the RNA is involved in. **(D)** The U-tract length of the RNA. **(E)** The median number of interacting partners of the RNA interactors (normalized as in B and expressed by Log_{10}). **(F)** The median SIC score of the RNA interactors. The distributions of each feature values in the group of known sRNAs (orange) and "other RNAs" (purple) are described by boxplots. The differences between the two distributions **(A–F)** were found to be statistically significant by two-tailed Mann–Whitney U test (*p* values between 10^{-23} and 10^{-11} after Bonferroni correction for multiple hypotheses testing). The distributions in panels **(A–F)** are based on the data of stationary phase RIL-seq experiment (**Supplementary Table 1**). Results for the exponential phase data sets are shown in **Supplementary Figures 2**, **3**.

that the weights reflecting the contributions of the second layer features are in opposite signs to the contributions of the first layer features (**Table 3**, **Figure 4**, **Supplementary Table 4**, and **Supplementary Figure 10**). Secondly, as explained in the Section "Materials and Methods," we can assess the relative contributions of the various features to the computed probability by examining the products of the weight and standard deviation of the feature values (**Figure 4** and **Supplementary Figure 10**). It seems that the major contributors to the final sRNA score involve both first- and second-layer features. The features that are high contributors in all data sets are 'the total number of chimeric fragments' and the 'median number of interactions the interactors are involved in', while the feature that consistently has the least contribution is 'number of unique interactions'. The contributions of the SIC (Second-In-Chimera), median SIC of interactors and the U-tract length seem to be more data set-dependent.

The second layer features were expected to prevent misclassification of a RNA targeted by multiple sRNAs ("target hub") as a sRNA. To assess this, we examined the sRNA scores of "target hubs", defined as RNAs interacting with at least four different sRNAs in at least one condition. Indeed, out of 18 "target hubs", 16 got low sRNA scores (**Supplementary Table 5**). When "target hubs" present high values of first layer features, such as a long U-tract, the second layer features may not be sufficient to prevent their misclassification. Indeed, the two "target hubs" *lpp* and *ompF* have a long U-tract of eight nucleotides each, are involved in 9 and 15 unique interactions, respectively, and have many chimeric fragments, together

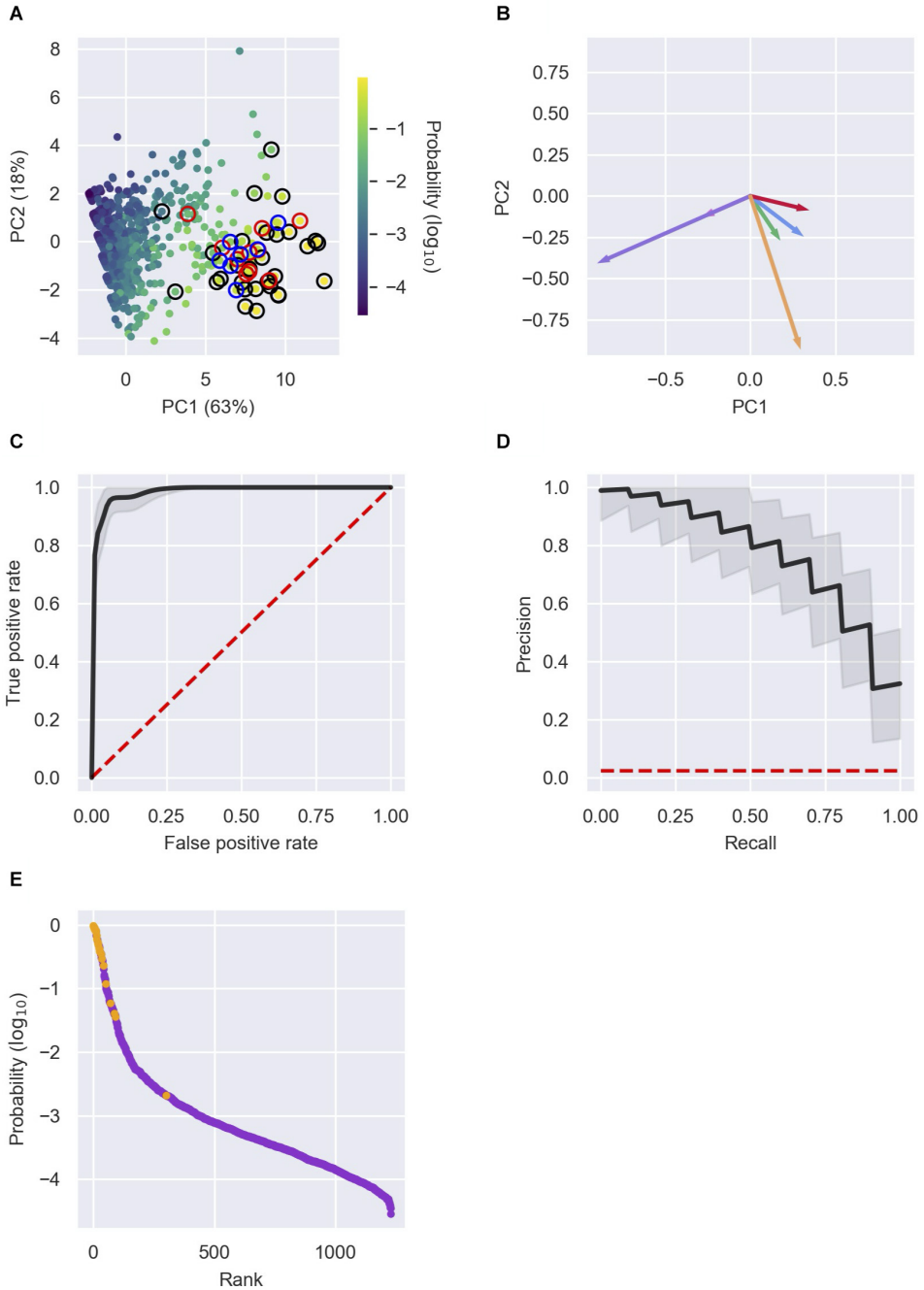


FIGURE 3 | Detection of novel sRNAs. **(A)** Principal component analysis (PCA) of RNAs characterized by the six features. The RNAs (dots) are plotted in two dimensions, using their projections onto the first two principal components. Each RNA in the data is colored by its sRNA probability, as assigned by the logistic regression analysis. Colored circles surrounding the dots represent: a well-established sRNA marked in **Supplementary Table 1** by 1 (black), a recently discovered sRNA listed in **Table 2A** (red) or a newly discovered sRNA listed in **Table 2B** (blue). **(B)** Contribution of the features to PC1 and PC2. The vectors represent the coefficients of the features in each PC: Total number of chimeric fragments (green), number of unique interactions (blue), SIC score (red), U-tract length (orange), median number of interactions of interactors (pink), median SIC score of interactors (purple). **(C,D)** Receiver operating characteristic (ROC) curve **(C)** and precision–recall (PR) curve **(D)** showing the high predictive power of the logistic regression model. Shown in black are the curves obtained from the mean probabilities of 10,000 iterations of the logistic regression, and the curves of individual iterations in the range of one standard deviation around the curve of mean probabilities. The curves are compared to the expected curve of a random classifier (red dashed line). The area under curve (AUC) of the ROC curve is 0.98 ± 0.01. **(E)** Known sRNAs and "other RNAs" (colored orange and purple, respectively) were ranked by their computed sRNA scores. Highly ranked RNAs, yet unknown as sRNAs, are predicted as putative novel sRNAs. Presented results are for the data set of stationary phase RIL-seq experiment. Results for the exponential phase data sets are shown in **Supplementary Figures 4**, **5**.

TABLE 3 | Weights of the logistic regression model for stationary phase data[a].

Total number of chimeric fragments	Number of unique interactions	SIC (Second-In-Chimera) score	U-tract length	Median number of interactions of interactors	Median SIC score of interactors	Intercept
0.956	0.354	0.683	0.584	−1.345	−0.955	−5.837

[a] *The table shows the mean intercept and weights of the 10,000 logistic regression iterations.*

enforcing their seemingly misclassification as a sRNAs, although with relatively low sRNA scores (**Supplementary Table 1**). Interestingly, a recent study identified a premature transcription termination site downstream to the transcription start site of *ompF*, suggesting that, in addition to being targeted by sRNAs in its 5′ UTR, a yet unknown small RNA overlapping *ompF* 5′ UTR might be generated (Adams et al., 2021). In general, sRNAs that function mainly as sponges of other sRNAs are not expected to be predicted by our algorithm as they usually have very few interactions (**Supplementary Table 5**). Yet, in a few cases the combination of various features in the prediction has allowed their identification by the computational scheme. For example, we found in RIL-seq stationary phase data that 70% of the chimeric fragments including the sRNA GcvB involve SroC, a recently discovered sRNA encoded in the 3′ UTR of *gltI*, a target of GcvB (Miyakoshi et al., 2015a). SroC sponges GcvB under stationary phase, relieving the repression of its targets. While SroC is involved mainly in the interaction with GcvB, our computational scheme awards it a relatively high sRNA probability in the stationary phase data, which is obtained by the combined contributions of all features (**Supplementary Tables 1, 5**).

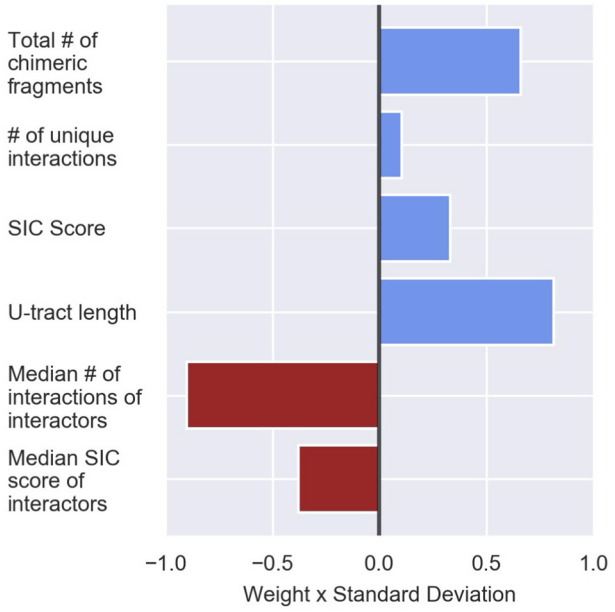


FIGURE 4 | Contribution of the various features to the logistic regression predictions. Presented are the logistic regression weights after z-score transformation of the feature values (see the section "Materials and Methods"). The presented weights, which are the original weights (**Table 3**) multiplied by the standard deviation of the feature value, are comparable. The weight value represents its contribution to the probability the logistic regression model provides, and the sign signifies the direction in which the weight affects this probability (i.e., positive values increase the sRNA probability and negative values reduce the sRNA probability). The results are based on the data set of stationary phase RIL-seq experiment. Results for the exponential phase data sets are shown in **Supplementary Figure 10**.

Experimental Verification of sRNA Candidates

Our computational scheme reported newly predicted sRNAs, encoded within various genomic elements (**Supplementary Table 1** summary tab). Many are encoded in 3′ UTRs, but there were also sRNA candidates encoded in 5′ UTRs and in coding sequences. We tested experimentally eleven candidates that got relatively high sRNA scores, but not necessarily those that ranked the highest above known sRNAs. These were selected to span the whole range of sRNA scores above known sRNAs and included seven candidates encoded in the 3′ UTR of protein-coding genes (*allR*, *bhsA*, *fadA*, *glpX*, *malG*, *ybiJ* and *ykgH*), two in intergenic regions (*rbsB-rbsK, xylA-xylB*), one in 5′ UTR (*kilR*), and one in a coding sequence (*aceK*). To validate the expression of these sRNAs we probed them by northern blotting, which provides information on both the expression pattern of the RNA and on its approximate size. Total RNA was extracted from wild type K-12 and Δ*hfq* strains grown to different growth phases and the expression of the sRNA candidates was tested. As *malG* 3′ UTR and *rbsB-rbsK* IGR were in the meanwhile reported by other groups as sRNAs [MalH (Iosub et al., 2020) and RbsZ (Melamed et al., 2020), respectively], we included them in **Table 2A** and report their northern blot results in **Supplementary Figure 11**. Seven of the remaining nine putative sRNAs were verified experimentally by northern blotting, where expression was evident in wild type but not in the Δ*hfq* strain (**Figure 5** and **Table 2B**). While this paper was under revision, the expression of AceK-int was confirmed also in another publication (Adams et al., 2021). The expression of sRNAs encoded at the 3′ UTRs of *glpX* and *ykgH* could not be verified by the northern blot experiments using two different probes for each of the candidates (**Supplementary Table 3**). However, the accumulation of RNA-seq reads at the 3′ UTR of *ykgH* and, with a less distinct pattern, at the 3′ UTR of *glpX* hint that transcripts originating from these loci do exist independently of the hosting gene (**Supplementary Figure 12**). All the verified sRNA candidates obtained high sRNA scores in the analysis of RIL-seq stationary phase data, and indeed they all accumulated during the stationary growth phase (**Figure 5** and **Table 2B**). Using RNA-seq data of stationary phase cells studied in our laboratory, we obtained estimates of the sizes of most sRNA candidates, and these sizes were confirmed by the northern blots (**Figure 5**).

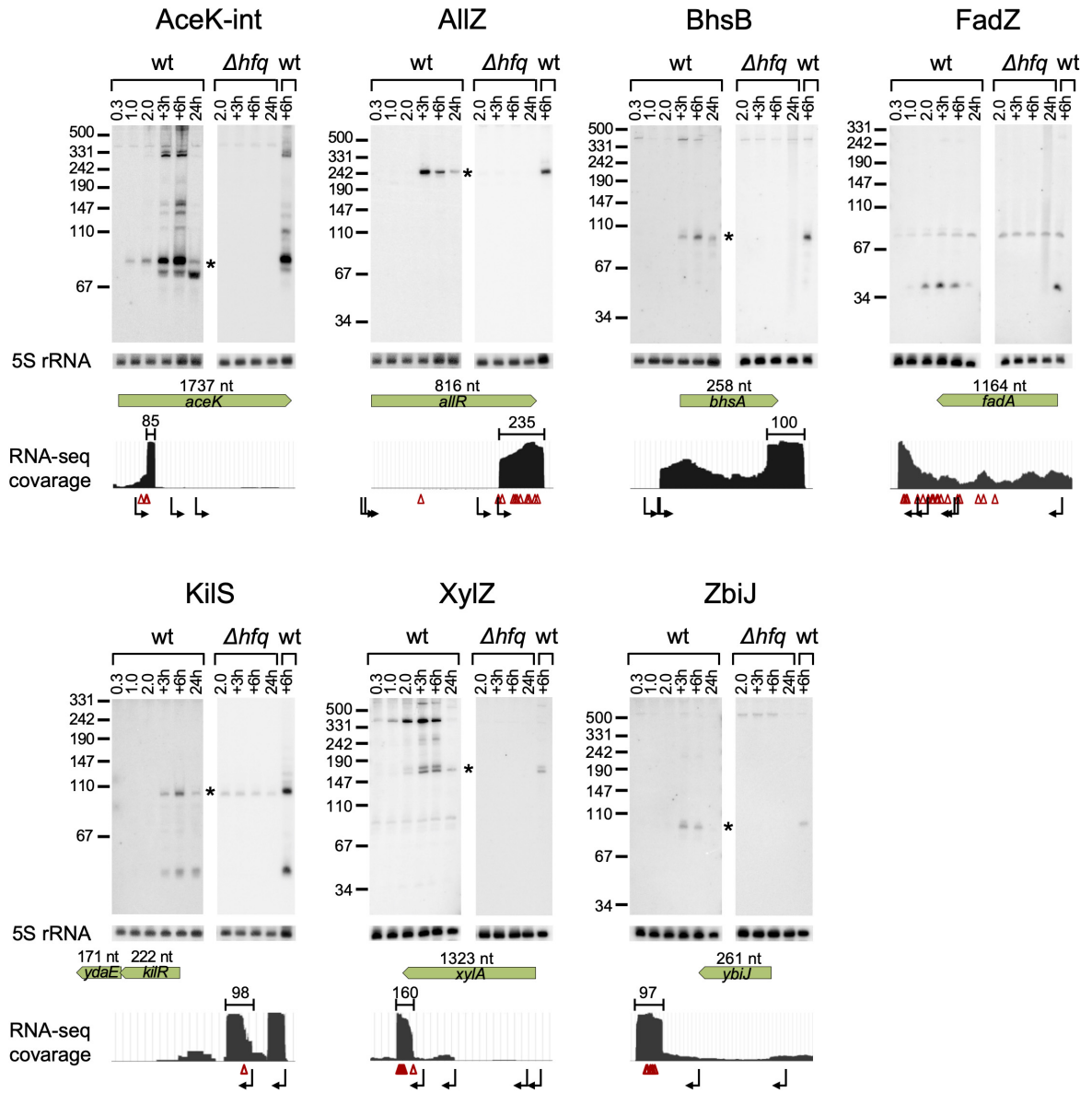


FIGURE 5 | Verification of the novel sRNAs by northern analysis. Total RNA was extracted from wt *E. coli* and Δ*hfq* cultures throughout growth. Samples of the wt culture were taken at an OD_{600} of 0.3, 1.0, and 2.0, 3 h and 6 h after the culture reached an OD_{600} of 2.0 (+3 h and +6 h, respectively) and after 24 h of growth (24 h). Samples of the Δ*hfq* were taken at an OD_{600} of 2.0, 3 h and 6 h after the culture reached an OD_{600} of 2.0 and after 24 h of growth. 30 μg total RNA were subjected to northern analysis using specific probes. The membrane used for the probing of AceK-int was re-used for the probing of AllZ, after the AceK-int probe radio-labeling has faded. 5S rRNA was probed as a loading control. For each sRNA, a coverage plot of RNA-seq library made of total RNA from a stationary phase (6 h growth) culture is shown. The green arrows indicate the coding sequence (CDS) region and gene orientation, with the CDS size above the arrow in nucleotides (nt). The approximated size of each sRNA is indicated above the read coverage plot (nt). Starlet indicates the band fitting in size to the RNA-seq data. Transcription start sites, based on data of Thomason et al. (2015) and Ju et al. (2019), and RNase E cleavage sites, based on data of Clarke et al. (2014) are shown below the read coverage plots along the transcript by bent black arrows and red triangles, respectively. Transcription start and cleavage sites in the vicinity of the suspected sRNA are recorded also in **Supplementary Table 6**.

To get clues whether the novel sRNAs were transcribed independently from an internal promoter or were processed from the hosting mRNA by an endoribonuclease, we examined global TSS data (Thomason et al., 2015; Ju et al., 2019) and large-scale cleavage data of RNase E (Clarke et al., 2014). These analyses indicated that AllZ and KilS can be transcribed from independent promoters, while the other novel sRNAs seem to be processed by an endoribonuclease from longer transcripts. We identified cleavage sites at the 5′ end position of AceK-int and two nucleotides upstream to the approximated 5′ end of XylZ (**Figure 5** and **Supplementary Table 6**). The generation of ZbiJ and BhsB cannot be explained by the previously mapped

TSSs or RNase E cleavage sites, as none were mapped near their approximated 5′ ends.

DISCUSSION

Systematic detection of sRNAs in large-scale RNA-seq data is highly valuable. As the fraction of genomic elements producing sRNAs out of all genes expressed in a cell is very small and estimated to be around 2% [∼100 sRNAs out of ∼4500 transcribed genes (Keseler et al., 2017)], the probability of detecting a genomic region encoding a sRNA at random is very small. In contrast, a prediction of "not a sRNA" for a genomic element has a high chance to be correct. Therefore, if our interest was in classification *per se*, it would be worthwhile to declare each RNA as non-sRNA, promising high chance of success. However, our challenge has been to find these needles in the haystack of all genes, and indeed we demonstrated that using informative features extracted from RIL-seq data and from the RNA sequences, it is feasible to distinguish the sRNAs from other genes. Using these features and the predictive scheme they are incorporated in, we predict additional novel sRNAs and demonstrate experimentally their expression as Hfq-dependent sRNAs.

There is an inherent difficulty in analyzing data that include ambiguous annotations for some genes, where some genomic elements classified as "other RNAs" are actually sRNAs that have not yet been detected. This causes the precision of the prediction to be underestimated. Indeed, if we re-label the recently discovered sRNAs and the seven additional experimentally verified sRNAs in the data (**Table 2**) as sRNAs and re-compute the precision rates we obtain better results (**Figure 6**). Interestingly, training the logistic regression model on the re-labeled data does not provide substantial improvement in the precision–recall results (**Figure 6**). The ambiguity of the initial labeling has also guided our strategy for determining new putative sRNAs. Thus, we chose to scan the RNAs ranked by their sRNA scores from top down, and classify RNAs ranked above known sRNAs as putative sRNAs that were wrongly labeled as "other RNAs." Using this strategy, we predicted nine novel sRNAs that obtained sRNA scores of 0.1–0.59, seven of which were verified experimentally. As stated above, as the chance probability for a genomic element to encode a sRNA is about 0.02, a sRNA score of 0.1 is also high above random expectation.

It is interesting, yet not surprising, that a RNA can be ranked differently in the different data sets, since it gets different sRNA scores depending on the feature values and weights in each data set. As most of the feature values are derived from RIL-seq data, which may change for a particular gene from experiment to experiment, it is conceivable that its computed sRNA score may change (**Supplementary Figure 6**). For example, a sRNA that is weakly expressed under one of the conditions may be involved in fewer chimeric fragments under this condition, and the weak contribution of the feature "total number of chimeric fragments" may lead to a final low probability by the predictor. In no way this means that the RNA is a sRNA under one condition and not under another.

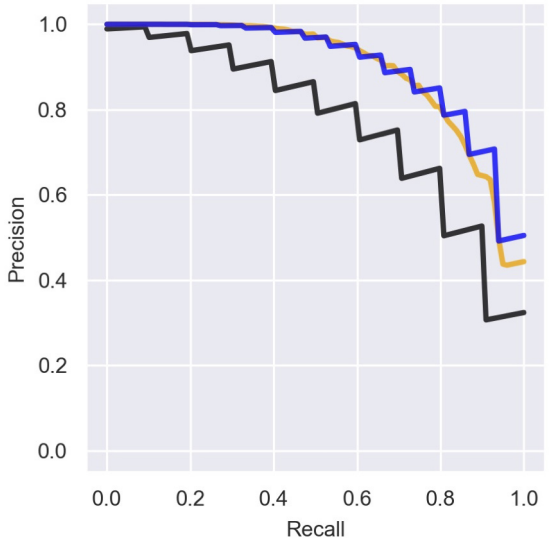


FIGURE 6 | Precision of the predictions with and without labeling of novel sRNAs. Presented is a comparison of the mean precision–recall (PR) curve over the logistic regression iterations for different sets of known sRNAs. We consider three cases: (1) Precision curve as in **Figure 3** (black line), when only sRNAs marked by 1 in **Supplementary Table 1** are considered as known sRNAs in both training and test sets. (2) The training is done with known sRNAs as in (1) but for the assessment we label all the sRNAs in **Table 2** as known sRNAs (orange line). (3) RNAs from **Table 2** are labeled as sRNAs for the training and for the assessment (blue line).

It simply means that the data of this RNA under a certain condition was not sufficient to allow its identification as a sRNA. Hence, we consider a genomic element as encoding a putative sRNA if it was ranked high and among known sRNAs in at least one data set.

The computed weights are also data set-dependent, and we examined whether their relative contributions are consistent or differ among the data sets. Comparing the original weights (**Supplementary Table 4**) and the products of weight and standard deviation (**Figure 4** and **Supplementary Figure 10**), we observed that, as expected, the directions of the contributions of the various features are consistent in all data sets, as well as the features that are main contributors. In all data sets the total number of chimeric fragments had a substantial positive effect, while the median number of the interactions of interactors had a large negative effect. The large contribution of this latter feature emphasizes that the recognition of the interactors as targets rather than sRNAs is highly important for the success of the predictions. Interestingly, the contribution of the U-tract length changes between the various conditions. This might be due to differences in the compositions of chimeric fragments among the various data sets, which may result in sRNAs with short U-tract in a particular data set, affecting its weight. The slight differences in the weights among the data sets suggest that it will be preferable to develop a predictive scheme per data set by repeating the learning process. Yet, the features we present can be easily extracted from the RIL-seq data and the execution of the

TABLE 4 | Common pathways involving the host genes of novel sRNAs and their targets.

Host gene	Host gene function	sRNA	sRNA target gene[a]	Target gene function	Suggested common pathway	References	Additional RIL-seq targets[a]
aceK	Regulator of the branch point between the TCA cycle and the glyoxylate cycle	AceK-int	*gatY*	Tagatose-1,6-bisphosphate aldolase 2 subunit; galactitol metabolism	Carbohydrate metabolic process	LaPorte and Koshland (1982); Richet and Raibaud (1989); Nobelmann and Lengeler (1996)	*clpB.rrsG*.IGR; *glpQ*; *fur*; *ryjB*; *ryjB.sgcQ*.IGR; *ycaK*.3UTR; *ydgA*; *yfjJ*; *yqeG*
			gatR	*gat* operon repressor; galactitol metabolism			
			malT	*mal* operon activator; maltose catabolism and transport			
allR	Transcriptional repressor of genes involved in anaerobic utilization of allantonin as a nitrogen source	AllZ	*grcA*	Stress-induced alternate pyruvate formate-lyase subunit; important in anaerobic maintenance of redox balance	Anaerobic metabolism	Cusa et al. (1999); Rintoul et al. (2002); Kramer et al. (2010)	*ftsA.ftsZ*.IGR; *ftsZ*; *bssR*
bhsA	Outer membrane stress protein; induced by H_2O_2 and increases cell resistance to H_2O_2 induced stress	BhsB	*ompC*	Outer membrane protein; was shown to facilitate and regulate the diffusion of H_2O_2 through the outer membrane in *Salmonella*	Cell response to oxidative stress	Pomposiello et al. (2001); Zheng et al. (2001); van der Heijden et al. (2016); Iwadate and Kato (2017)	
			ytfK	Stringent response activator; induced by paraquat, involved in H_2O_2 tolerance			
fadA	3-ketoacyl-CoA thiolase, involved in fatty acid degradation via β-oxidation and generation of acetyl-CoA	FadZ	*dctA*	C4 dicarboxylate/orotate:H^+ symporter; importer of metabolites that can serve as substrates in the TCA cycle	TCA cycle	Kay and Kornberg (1969, 1971); Darlison et al. (1984); Buck et al. (1986)	*clpS*; *ompC*; *ompF*; *yhsB*
			kgtP	a-Ketoglutarate:H^+ symporter			
			sucA	Component of the 2-oxoglutarate dehydrogenase multienzyme complex			
kilR	Killing protein; inhibits cell division by binding FtsZ	KilS	*yncL*	Inner membrane protein of unknown function	KilR targets FtsZ and YncL, both localize to the inner membrane	Overath and Raufuss (1967); Kay and Kornberg (1969, 1971); Lo et al. (1972)	

[a] *Only RIL-seq targets detected in at least one individual library in addition to the unified libraries were regarded in this analysis.*

logistic regression is straightforward, making our approach feasible for detecting novel sRNAs in *E. coli* grown under other conditions and in other bacteria to which RIL-seq is applied. It is of note that applying similar computational approaches to data sets of sRNA-target pairs detected by different methods may result in different informative features. For example, in a comparable data set of chimeric fragments including sRNAs and targets, recently determined by the CLASH methodology applied to Hfq in *E. coli* (Iosub et al., 2020), the sRNAs were not found to be preferentially the second RNAs in their respective chimeras. Encouragingly, all the novel sRNAs reported here were included in the CLASH chimeras, and half of them were located mostly second in their chimeras.

While most of the previously known sRNAs are properly classified, we do encounter and expect misclassifications emerging mainly from three major RNA classes. The first class comprises "target hubs" that interact with multiple sRNAs (**Supplementary Table 5**). While in most cases the second-layer features prevent their misclassification as sRNAs, some RNAs exhibiting very strong first layer features might be misclassified. Note however that the exact classification of sRNAs and targets is not always obvious and some of these allegedly misclassified sRNAs may turn out to be true sRNAs (e.g., the above described

ompF 5′ UTR). The second class comprises sRNAs with very few targets. This group includes sRNAs that are not highly expressed in the conditions studied here and, thus, are lowly ranked at these conditions, but they are likely to be detected under the relevant condition. In addition, this class includes highly specialized sRNAs with specific targets, mostly considered as sponges. While some sponges, with extremely high first-layer properties, such as SroC, are classified as sRNAs, others are not (**Supplementary Table 5**). As this special class of sRNAs is not expected to be predicted by an algorithm like the one presented here, which is trained on information drawn mostly from the RNA interactome, loading the training set with single target sRNAs is not recommended. The third class of misclassifications can be traced back to ambiguous annotation of the RIL-seq data itself, and in particular to reads overlapping different genomic annotations (e.g., CDS and 3′ UTR of the same gene). Putative sRNA for which the reads are split between two annotations, are more likely to be missed. Furthermore, one annotation, e.g., CDS, can be misclassified as sRNA at the expense of the second annotation, e.g., the respective 3′ UTR-derived sRNA (e.g., *uhpT* and *sucD*). Re-examination of the proximity of the chimeric fragment coordinates of the CDS-derived candidates to 5′ UTR or 3′ UTR can resolve some of these misclassifications.

Finding that the targets of the newly revealed sRNAs have functions that are associated with the function of the hosting mRNA would support the functionality of the novel sRNAs as regulatory molecules. It would also suggest that the sRNA and hosting gene affect the same pathways at different regulation levels, and, in case they share targets they may generate regulatory circuits combining multiple regulation levels. However, as RIL-seq data do not provide information whether the sRNA enhances or represses the target expression, it would not be possible at this stage to draw mechanistic conclusions on such possible circuits. Yet, we found for several of the novel sRNAs that the hosting genes and their targets are involved in common pathways (**Table 4**). For example, BhsB is derived from the 3′ end of the *bhsA* mRNA, encoding a small outer membrane protein that is involved in various stress responses. Oxidative stress induced by hydrogen peroxide or paraquat activates *bhsA* transcription (Pomposiello et al., 2001; Zheng et al., 2001). Also, BhsA was shown to increase cell resistance to hydrogen peroxide (Zhang et al., 2007). The RIL-seq data indicate that BhsB interacts with two targets, *ytfK* and *ompC,* and both were shown to be involved in the cellular response to oxidative stress. *ytfK*, induced by paraquat (Pomposiello et al., 2001), was shown to be involved in hydrogen peroxide tolerance (Iwadate and Kato, 2017). OmpC, an outer membrane protein, was shown in *Salmonella* to facilitate and regulate the diffusion of hydrogen peroxide through the outer membrane (van der Heijden et al., 2016). Thus, RIL-seq results suggest a shared pathway for the hosting gene and the sRNA derived from its transcript, further supporting the functionality of the 3′ UTR-derived BhsB as a sRNA.

In summary, using our methodology followed by experimental verification we reaffirmed that there is a rich repertoire of sRNAs encoded within various genomic elements and generated under different conditions. The use of our systematic approach has allowed us to identify putative sRNAs that would not have been considered otherwise, such as AllZ, KilS and BhsB. Each of them is involved in only a few interactions and a few hundred chimeric fragments, and they would not have been suspected as sRNAs by examining their individual features alone. Yet, the overall combination of their features has allowed their detection. Especially, their interactors had very few interactions with "other RNAs", rejecting the interactors as sRNAs and supporting AllZ, KilS and BhsB as sRNAs. We believe that taking into account both the first- and second-layer features empowers our predictions. Hence, taking into account the information extracted directly from RIL-seq data while accounting for the RNA–RNA interaction network inferred from RIL-seq results is highly rewarding. Our approach may be generalized to other RNA-seq-based methodologies, where the results may imply a network structure or hierarchy of the genes. Combining features based on the direct sequencing results with features based on a higher-order structure of the data may prove beneficial to the inference of novel biological insights in other contexts.

DATA AVAILABILITY STATEMENT

The data used in this article can be found in https://www.ebi.ac.uk/arrayexpress/E-MTAB-9834 and in https://www.ebi.ac.uk/arrayexpress/E-MTAB-3910.

AUTHOR CONTRIBUTIONS

HM and YA initiated and supervised the study. LA contributed to the experimental investigation. AB contributed to the software development and computational analysis. YA, LA, AB, and HM contributed to the writing – original draft and review and editing. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the European Research Council Advanced Grants #322920 and #833598. AB was supported by the Data Sciences Program of the Planning and Budgeting Committee of the Israel Council for Higher Education.

ACKNOWLEDGMENTS

We are thankful to Gal Elidan for very helpful advice.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.635070/full#supplementary-material

REFERENCES

Adams, P. P., Baniulyte, G., Esnault, C., Chegireddy, K., Singh, N., Monge, M., et al. (2021). Regulatory roles of *Escherichia coli* 5′ UTR and ORF-internal RNAs detected by 3′ end mapping. *eLife* 10:e62438.

Adams, P. P., and Storz, G. (2020). Prevalence of small base-pairing RNAs derived from diverse genomic loci. *Biochim. Biophys. Acta Gene Regul. Mech.* 1863:194524. doi: 10.1016/j.bbagrm.2020.194524

Altuvia, S., Weinstein-Fischer, D., Zhang, A., Postow, L., and Storz, G. (1997). A small, stable RNA induced by oxidative stress: role as a pleiotropic regulator and antimutator. *Cell* 90, 43–53. doi: 10.1016/s0092-8674(00)80312-8

Argaman, L., Hershberg, R., Vogel, J., Bejerano, G., Wagner, E. G., Margalit, H., et al. (2001). Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. *Curr. Biol.* 11, 941–950. doi: 10.1016/s0960-9822(01)00270-6

Bilusic, I., Popitsch, N., Rescheneder, P., Schroeder, R., and Lybecker, M. (2014). Revisiting the coding potential of the E. *coli* genome through Hfq co-immunoprecipitation. *RNA Biol.* 11, 641–654. doi: 10.4161/rna.29299

Buck, D., Spencer, M., and Guest, J. (1986). Cloning and expression of the succinyl-CoA synthetase genes of *Escherichia coli* K12. *J. Gen. Microbiol.* 132:10.

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

Chao, Y., and Vogel, J. (2016). A 3′ UTR-derived small RNA provides the regulatory noncoding arm of the inner membrane stress response. *Mol. Cell* 61, 352–363. doi: 10.1016/j.molcel.2015.12.023

Chen, S., Lesnik, E. A., Hall, T. A., Sampath, R., Griffey, R. H., Ecker, D. J., et al. (2002). A bioinformatics based approach to discover small RNA genes in the *Escherichia coli* genome. *Biosystems* 65, 157–177. doi: 10.1016/s0303-2647(02)00013-8

Clarke, J. E., Kime, L., Romero, A. D., and McDowall, K. J. (2014). Direct entry by RNase E is a major pathway for the degradation and processing of RNA in *Escherichia coli*. *Nucleic Acids Res.* 42, 11733–11751. doi: 10.1093/nar/gku808

Cusa, E., Obradors, N., Baldomà, L., Badía, J., and Aguilar, J. (1999). Genetic analysis of a chromosomal region containing genes required for assimilation of allantoin nitrogen and linked glyoxylate metabolism in *Escherichia coli*. *J. Bacteriol.* 181, 7479–7484. doi: 10.1128/jb.181.24.7479-7484.1999

Darlison, M. G., Spencer, M. E., and Guest, J. R. (1984). Nucleotide sequence of the sucA gene encoding the 2-oxoglutarate dehydrogenase of *Escherichia coli* K12. *Eur. J. Biochem.* 141, 351–359. doi: 10.1111/j.1432-1033.1984.tb08199.x

De Lay, N., Schu, D. J., and Gottesman, S. (2013). Bacterial small RNA-based negative regulation: Hfq and its accomplices. *J. Biol. Chem.* 288, 7996–8003. doi: 10.1074/jbc.r112.441386

De Mets, F., Van Melderen, L., and Gottesman, S. (2019). Regulation of acetate metabolism and coordination with the TCA cycle via a processed small RNA. *Proc. Natl. Acad. Sci. U.S.A.* 116, 1043–1052. doi: 10.1073/pnas.1815288116

Dimastrogiovanni, D., Fröhlich, K. S., Bandyra, K. J., Bruce, H. A., Hohensee, S., Vogel, J., et al. (2014). Recognition of the small regulatory RNA RydC by the bacterial Hfq protein. *eLife* 3:e05375.

Grabowicz, M., Koren, D., and Silhavy, T. J. (2016). The CpxQ sRNA negatively regulates Skp to prevent mistargeting of beta-barrel outer membrane proteins into the cytoplasmic membrane. *mBio* 7:e00312-16.

Guo, M. S., Updegrove, T. B., Gogol, E. B., Shabalina, S. A., Gross, C. A., and Storz, G. (2014). MicL, a new sigmaE-dependent sRNA, combats envelope stress by repressing synthesis of Lpp, the major outer membrane lipoprotein. *Genes Dev.* 28, 1620–1634. doi: 10.1101/gad.243485.114

Holmqvist, E., Wright, P. R., Li, L., Bischler, T., Barquist, L., Reinhardt, R., 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

Hör, J., Matera, G., Vogel, J., Gottesman, S., and Storz, G. (2020). Trans-acting small RNAs and their effects on gene expression in *Escherichia coli* and *Salmonella enterica*. *EcoSal Plus* 9.

Huber, M., Fröhlich, K. S., Radmer, J., and Papenfort, K. (2020). Switching fatty acid metabolism by an RNA-controlled feed forward loop. *Proc. Natl. Acad. Sci. U.S.A.* 117, 8044–8054. doi: 10.1073/pnas.1920753117

Ikemura, T., and Dahlberg, J. E. (1973a). Small ribonucleic acids of *Escherichia coli*. I. Characterization by polyacrylamide gel electrophoresis and fingerprint analysis. *J. Biol. Chem.* 248, 5024–5032.

Ikemura, T., and Dahlberg, J. E. (1973b). Small ribonucleic acids of *Escherichia coli*. II. Noncoordinate accumulation during stringent control. *J. Biol. Chem.* 248, 5033–5041. doi: 10.1016/s0021-9258(19)43667-3

Iosub, I. A., van Nues, R. W., McKellar, S. W., Nieken, K. J., Marchioretto, M., Sy, B., et al. (2020). Hfq CLASH uncovers sRNA-target interaction networks linked to nutrient availability adaptation. *eLife* 9:e54655.

Iwadate, Y., and Kato, J. I. (2017). Involvement of the ytfK gene from the PhoB regulon in stationary-phase H2O2 stress tolerance in *Escherichia coli*. *Microbiology* 163, 1912–1923. doi: 10.1099/mic.0.000534

Ju, X., Li, D., and Liu, S. (2019). Full-length RNA profiling reveals pervasive bidirectional transcription terminators in bacteria. *Nat. Microbiol.* 4, 1907–1918. doi: 10.1038/s41564-019-0500-z

Kay, W. W., and Kornberg, H. L. (1969). Genetic control of the uptake of C(4)-dicarboxylic acids by *Escherichia coli*. *FEBS Lett.* 3, 93–96. doi: 10.1016/0014-5793(69)80105-5

Kay, W. W., and Kornberg, H. L. (1971). The uptake of C4—Dicarboxylic acids by *Escherichia coli*. *Eur. J. Biochem.* 18, 274–281. doi: 10.1111/j.1432-1033.1971.tb01240.x

Keseler, I. M., Mackie, A., Santos-Zavaleta, A., Billington, R., Bonavides-Martínez, C., Caspi, R., et al. (2017). The EcoCyc database: reflecting new knowledge about *Escherichia coli* K-12. *Nucleic Acids Res.* 45, D543–D550.

Kramer, G., Sprenger, R. R., Nessen, M. A., Roseboom, W., Speijer, D., de Jong, L., et al. (2010). Proteome-wide alterations in *Escherichia coli* translation rates upon *Anaerobiosis*. *Mol. Cell. Proteom.* 9, 2508–2516. doi: 10.1074/mcp.m110.001826

Lalaouna, D., Carrier, M. C., Semsey, S., Brouard, J. S., Wang, J., Wade, J. T., et al. (2015). A 3′ external transcribed spacer in a tRNA transcript acts as a sponge for small RNAs to prevent transcriptional noise. *Mol. Cell* 58, 393–405. doi: 10.1016/j.molcel.2015.03.013

LaPorte, D. C., and Koshland, D. E. Jr. (1982). A protein with kinase and phosphatase activities involved in regulation of tricarboxylic acid cycle. *Nature* 300, 458–460. doi: 10.1038/300458a0

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* [preprint] arXiv:1303.3997,

Lo, T. C. Y., Rayman, M. K., and Sanwal, B. D. (1972). Transport of Succinate in *Escherichia coli*. *J. Biol. Chem.* 247, 6323–6331. doi: 10.1016/s0021-9258(19)44800-x

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 17:10. doi: 10.14806/ej.17.1.200

Melamed, S., Adams, P. P., Zhang, A., Zhang, H., and Storz, G. (2020). RNA-RNA interactomes of ProQ and Hfq reveal overlapping and competing roles. *Mol. Cell* 77, 411–425. doi: 10.1016/j.molcel.2019.10.022

Melamed, S., Faigenbaum-Romm, R., Peer, A., Reiss, N., Shechter, O., Bar, A., et al. (2018). Mapping the small RNA interactome in bacteria using RIL-seq. *Nat. Protoc.* 13, 1–33. doi: 10.1038/nprot.2017.115

Melamed, S., Peer, A., Faigenbaum-Romm, R., Gatt, Y. E., Reiss, N., Bar, A., 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

Mihailovic, M. K., Vazquez-Anderson, J., Li, Y., Fry, V., Vimalathas, P., Herrera, D., et al. (2018). High-throughput in vivo mapping of RNA accessible interfaces to identify functional sRNA binding sites. *Nat. Commun.* 9:4084.

Miyakoshi, M., Chao, Y., and Vogel, J. (2015a). Cross talk between ABC transporter mRNAs via a target mRNA-derived sponge of the GcvB small RNA. *EMBO J.* 34, 1478–1492. doi: 10.15252/embj.201490546

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

Miyakoshi, M., Matera, G., Maki, K., Sone, Y., and Vogel, J. (2019). Functional expansion of a TCA cycle operon mRNA by a 3′ end-derived small RNA. *Nucleic Acids Res.* 47, 2075–2088. doi: 10.1093/nar/gky1243

Mizuno, T., Chou, M. Y., and Inouye, M. (1984). A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA). *Proc. Natl. Acad. Sci. U.S.A.* 81, 1966–1970. doi: 10.1073/pnas.81.7.1966

Nobelmann, B., and Lengeler, J. W. (1996). Molecular analysis of the gat genes from *Escherichia coli* and of their roles in galactitol transport and metabolism. *J. Bacteriol.* 178, 6790–6795. doi: 10.1128/jb.178.23.6790-6795.1996

Otaka, H., Ishikawa, H., Morita, T., and Aiba, H. (2011). PolyU tail of rho-independent terminator of bacterial small RNAs is essential for Hfq action. *Proc. Natl. Acad. Sci. U.S.A.* 108, 13059–13064. doi: 10.1073/pnas.1107050108

Overath, P., and Raufuss, E. M. (1967). The induction of the enzymes of fatty acid degradation in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* 29, 28–33.

Pomposiello, P. J., Bennik, M. H., and Demple, B. (2001). Genome-wide transcriptional profiling of the *Escherichia coli* responses to superoxide stress and sodium salicylate. *J. Bacteriol.* 183, 3890–3902. doi: 10.1128/jb.183.13.3890-3902.2001

Richet, E., and Raibaud, O. (1989). MalT, the regulatory protein of the *Escherichia coli* maltose system, is an ATP-dependent transcriptional activator. *EMBO J.* 8, 981–987. doi: 10.1002/j.1460-2075.1989.tb03461.x

Rintoul, M. R., Cusa, E., Baldomà, L., Badia, J., Reitzer, L., and Aguilar, J. (2002). Regulation of the *Escherichia coli* allantoin regulon: coordinated function of the repressor AllR and the activator AllS. *J. Mol. Biol.* 324, 599–610. doi: 10.1016/s0022-2836(02)01134-8

Rivas, E., Klein, R. J., Jones, T. A., and Eddy, S. R. (2001). Computational identification of noncoding RNAs in E. *coli by comparative genomics*. *Curr. Biol.* 11, 1369–1373. doi: 10.1016/s0960-9822(01)00401-8

Santiago-Frangos, A., and Woodson, S. A. (2018). Hfq chaperone brings speed dating to bacterial sRNA. *Wiley Interdiscip. Rev. RNA* 9:e1475. doi: 10.1002/wrna.1475

Sauer, E., and Weichenrieder, O. (2011). Structural basis for RNA 3′-end recognition by Hfq. *Proc. Natl. Acad. Sci. U.S.A.* 108, 13065–13070. doi: 10.1073/pnas.1103420108

Schu, D. J., Zhang, A., Gottesman, S., and Storz, G. (2015). Alternative Hfq-sRNA interaction modes dictate alternative mRNA recognition. *EMBO J.* 34, 2557–2573. doi: 10.15252/embj.201591569

Shishkin, A. A., Giannoukos, G., Kucukural, A., Ciulla, D., Busby, M., Surka, C., et al. (2015). Simultaneous generation of many RNA-seq libraries in a single reaction. *Nat. Methods* 12, 323–325. doi: 10.1038/nmeth.3313

Sledjeski, D., and Gottesman, S. (1995). A small RNA acts as an antisilencer of the H-NS-silenced rcsA gene of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 92, 2003–2007. doi: 10.1073/pnas.92.6.2003

Smirnov, A., Wang, C., Drewry, L. L., and Vogel, J. (2017). Molecular mechanism of mRNA repression in trans by a ProQ-dependent small RNA. *EMBO J.* 36, 1029–1045. doi: 10.15252/embj.201696127

Thomason, M. K., Bischler, T., Eisenbart, S. K., Förstner, K. U., Zhang, A., Herbig, A., et al. (2015). Global transcriptional start site mapping using differential RNA sequencing reveals novel antisense RNAs in *Escherichia coli*. *J. Bacteriol.* 197, 18–28. doi: 10.1128/jb.02096-14

Tree, J. J., Granneman, S., McAteer, S. P., Tollervey, D., and Gally, D. L. (2014). Identification of bacteriophage-encoded anti-sRNAs in pathogenic *Escherichia coli*. *Mol. Cell.* 55, 199–213. doi: 10.1016/j.molcel.2014.05.006

Updegrove, T. B., Zhang, A., and Storz, G. (2016). Hfq: the flexible RNA matchmaker. *Curr. Opin. Microbiol.* 30, 133–138. doi: 10.1016/j.mib.2016.02.003

Urban, J. H., and Vogel, J. (2007). Translational control and target recognition by *Escherichia coli* small RNAs in vivo. *Nucleic Acids Res.* 35, 1018–1037. doi: 10.1093/nar/gkl1040

van der Heijden, J., Reynolds, L. A., Deng, W., Mills, A., Scholz, R., Imami, K., et al. (2016). *Salmonella* rapidly regulates membrane permeability to survive oxidative stress. *mBio* 7:e01238-16.

Vogel, J., and Luisi, B. F. (2011). Hfq and its constellation of RNA. *Nat. Rev. Microbiol.* 9, 578–589. doi: 10.1038/nrmicro2615

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

Wang, C., Chao, Y., Matera, G., Gao, Q., and Vogel, J. (2020). The conserved 3′ UTR-derived small RNA NarS mediates mRNA crossregulation during nitrate respiration. *Nucleic Acids Res.* 48, 2126–2143. doi: 10.1093/nar/gkz1168

Wassarman, K. M., Repoila, F., Rosenow, C., Storz, G., and Gottesman, S. (2001). Identification of novel small RNAs using comparative genomics and microarrays. *Genes Dev.* 15, 1637–1651. doi: 10.1101/gad.901001

Zhang, A., Wassarman, K. M., Rosenow, C., Tjaden, B. C., Storz, G., and Gottesman, S. (2003). Global analysis of small RNA and mRNA targets of Hfq. *Mol. Microbiol.* 50, 1111–1124. doi: 10.1046/j.1365-2958.2003.03734.x

Zhang, X. S., García-Contreras, R., and Wood, T. K. (2007). YcfR (BhsA) influences *Escherichia coli* biofilm formation through stress response and surface hydrophobicity. *J. Bacteriol.* 189, 3051–3062. doi: 10.1128/jb.01832-06

Zheng, M., Wang, X., Templeton, L. J., Smulski, D. R., LaRossa, R. A., and Storz, G. (2001). DNA microarray-mediated transcriptional profiling of the *Escherichia coli* response to hydrogen peroxide. *J. Bacteriol.* 183, 4562–4570. doi: 10.1128/jb.183.15.4562-4570.2001

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 21 July 2021
doi: 10.3389/fmicb.2021.649858

Interspecies RNA Interactome of Pathogen and Host in a Heritable Defensive Strategy

*Marcela Legüe[1], Blanca Aguila[1,2] and Andrea Calixto[1]**

[1] Centro Interdisciplinario de Neurociencia de Valparaíso, Instituto de Neurociencia, Facultad de Ciencias, Universidad de Valparaíso, Valparaíso, Chile, [2] Programa de Doctorado en Microbiología, Universidad de Chile, Santiago, Chile

Communication with bacteria deeply impacts the life history traits of their hosts. Through specific molecules and metabolites, bacteria can promote short- and long-term phenotypic and behavioral changes in the nematode *Caenorhabditis elegans*. The chronic exposure of *C. elegans* to pathogens promotes the adaptive behavior in the host's progeny called pathogen-induced diapause formation (PIDF). PIDF is a pathogen avoidance strategy induced in the second generation of animals infected and can be recalled transgenerationally. This behavior requires the RNA interference machinery and specific nematode and bacteria small RNAs (sRNAs). In this work, we assume that RNAs from both species co-exist and can interact with each other. Under this principle, we explore the potential interspecies RNA interactions during PIDF-triggering conditions, using transcriptomic data from the holobiont. We study two transcriptomics datasets: first, the dual sRNA expression of *Pseudomonas aeruginosa* PAO1 and *C. elegans* in a transgenerational paradigm for six generations and second, the simultaneous expression of sRNAs and mRNA in intergenerational PIDF. We focus on those bacterial sRNAs that are systematically overexpressed in the intestines of animals compared with sRNAs expressed in host-naïve bacteria. We selected diverse *in silico* methods that represent putative mechanisms of RNA-mediated interspecies interaction. These interactions are as follows: heterologous perfect and incomplete pairing between bacterial RNA and host mRNA; sRNAs of similar sequence expressed in both species that could mimic each other; and known or predicted eukaryotic motifs present in bacterial transcripts. We conclude that a broad spectrum of tools can be applied for the identification of potential sRNA and mRNA targets of the interspecies RNA interaction that can be subsequently tested experimentally.

Keywords: small RNAs, *C. elegans*, *P. aeruginosa* PAO1, interspecies communication, host behavioral defenses, RNA–RNA interaction, dual-RNA-seq transcriptomics, pathogen-induced diapause

OPEN ACCESS

Edited by:
Jiqiang Ling,
University of Maryland, United States

Reviewed by:
Balamurugan Krishnaswamy,
Alagappa University, India
Alejandro Vasquez-Rifo,
University of Massachusetts Medical School, United States

****Correspondence:***
Andrea Calixto
andrea.calixto@uv.cl;
tigonas@gmail.com

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *05 January 2021*
Accepted: *17 June 2021*
Published: *21 July 2021*

Citation:
Legüe M, Aguila B and Calixto A (2021) Interspecies RNA Interactome of Pathogen and Host in a Heritable Defensive Strategy. Front. Microbiol. 12:649858. doi: 10.3389/fmicb.2021.649858

INTRODUCTION

Bacteria and animals are adapted to living together, and their interaction impacts the physiology of both entities (Rosenberg and Zilber-Rosenberg, 2013; Moënne-Loccoz et al., 2015). Many small molecules and metabolites are known to be mediators of this communication (Cleary et al., 2017). The emerging relevance of RNA molecules in the bacteria–host interplay is based on the discovery of specific bacterial RNAs that directly induce phenotypic changes in hosts.

For instance, *Escherichia coli* non-coding RNAs (ncRNAs) OxyS and DsrA affect chemotaxis and longevity in *Caenorhabditis elegans* by downregulating *che-2* (Liu et al., 2012). *Salmonella enterica* microRNA-like RNA fragment Sal-1 targets host-inducible nitric oxide synthase (Zhao et al., 2017); *Pseudomonas aeruginosa* PA14 methionine-transfer RNAs (tRNA) fragment from outer membrane vesicles (OMVs) induces IL-8 secretion in human epithelial cells (Koeppen et al., 2016); ncRNA p11 induces intergenerational learned avoidance in *C. elegans* by targeting *maco-1* (Kaletsky et al., 2020); and *P. aeruginosa* PAO1 RsmY triggers transgenerational diapause (Legüe et al., 2021).

The aforementioned works support the idea that RNAs from both species can be transferred between animal tissues and be co-expressed spatially and temporally. This transcriptomic layer of regulation of the bacteria–host *holobiont* (Celluzzi and Masotti, 2016; Westermann et al., 2016; Rosenberg and Zilber-Rosenberg, 2018) is called *holo-transcriptome* (Palominos et al., 2017; Legüe and Calixto, 2019) and implies a tight communication directly at the RNA expression level. Indirect evidence of RNA exchange came from the fact that RNA is selectively secreted in bacterial extracellular vesicles and exosomes, which could be transferred between organisms (Ghosal et al., 2015; Sjöström et al., 2015; Celluzzi and Masotti, 2016; Koeppen et al., 2016; Liu et al., 2016; Westermann et al., 2016; Lefebvre and Lécuyer, 2017). Despite the advances in understanding the small RNA (sRNA)-mediated bacteria–host interaction, the broad spectrum of mechanisms by which RNAs from both species could interact has not been fully explored.

The search for inter-kingdom RNA interactions represents a conundrum. On one hand, the regulation of intraspecies RNA–RNA interaction is complex and diverse. On the other hand, mechanisms in bacteria and eukaryotes differ, making the process of finding commonalities between them a challenge. Until now, the search for interspecies RNA communication has relied on an oversimplification of the possibilities, using RNA sequence similarity as the only parameter (Nguyen et al., 2018). Here, we aim to systematize the possible mechanisms of interspecies RNA interaction and evaluate *in silico* the applicability of different bioinformatics tools for each proposed mechanism. To accomplish this, we take advantage of a behavioral paradigm in *C. elegans* called pathogen-induced diapause formation (PIDF) (Palominos et al., 2017; Gabaldón et al., 2020), which involves sRNAs from both bacteria and hosts. PIDF is a strategy of defense by which animals enter diapause to effectively avoid feeding on pathogens. For PIDF to take place, RsmY from *P. aeruginosa* (Legüe et al., 2021) and *mir-243* from *C. elegans* (Gabaldón et al., 2020) are required. In this paradigm, animals are in contact with pathogens for two generations and are re-introduced to pathogens after two generations in non-pathogenic bacteria. The datasets we use include intergenerational and transgenerational sRNA and polyA+ RNA-seq transcriptomics. This allows the comparison of generations of animals from the same cohort feeding on pathogens and non-pathogens. Additionally, this design permits the discrimination between dynamic changes in sRNA expression from the constitutive expression that are mostly due to intestinal life. Among RNA species that emerged as players of interspecies interaction are the tRNAs, which we show are implicated in the process of PIDF.

In this work, we explore interspecies RNA–RNA interactions based on the mechanisms described for bacteria and Eukarya independently. For each putative RNA interaction mechanism, we test and select the most appropriate existing bioinformatics tool. Finally, from the predicted relevant players, we show that ELPC-3, a tRNA elongation factor, plays a role in PIDF. These analyses open new insights and interesting lines of research in interspecies communication.

MATERIALS AND METHODS

Datasets Used

For downstream interaction prediction, we use the following datasets: (1) *C. elegans* sRNA transcriptomics deposited in the NCBI under BioProject no. PRJNA659467 (Gabaldón et al., 2020); (2) *P. aeruginosa* PAO1 sRNA transcriptomics deposited in the NCBI under BioProject no. PRJNA708299 (Legüe et al., 2021); (3) polyA+ transcriptomics to obtain downregulated mRNAs and lncRNAs during bacteria–nematode interaction, from **Supplementary Table 2** (Gabaldón et al., 2020) available at https://mbio.asm.org/content/mbio/11/5/e01950-20/DC2/embed/inline-supplementary-material-2.xlsx; and (4) normalized sRNA expression from *P. aeruginosa* PAO1 and *C. elegans* (Legüe et al., 2021).

Small RNA libraries were based on size selection of fragments shorter than 200 nt. Based on these criteria, we use the broad term of sRNA for transcripts shorter than 200 nt, and we refer to specific biotypes when relevant.

The data obtained from the *C. elegans* polyA+ RNA transcription was pre-processed using Trimmomatic v. 0.36 (Bolger et al., 2014). Reads with a quality (Phred score) less than 35 and a length less than 36 were removed. The reads were aligned using TopHat (Trapnell et al., 2009) and quantified with HTSeq (Anders et al., 2015). Differentially expressed genes (DEGs) were determined using EdgeR (Robinson et al., 2010) and DeSeq (Anders and Huber, 2010). DEG cutoff was set with a p-adjusted value < 0.01.

The sRNA data from *P. aeruginosa* and *C. elegans* sRNAs (Gabaldón et al., 2020; Legüe et al., 2021) was previously processed as follows unless explained otherwise:

Small RNA Data Pre-processing and Quality Control

Quality visualization was made with FastQC[1]. Trimming was performed with Cutadapt (Martin, 2011), using Diagenode-recommended parameters for CATS Library Preparation Kits available at https://www.diagenode.com/en/documents/diagenode-trimming-tools-for-cats-rnaseq.

Mapping

For each sample, reads were mapped to the *E. coli* OP50 genome assembly ASM435501v1 or *P. aeruginosa* PAO1, assembly ASM676v1, available at NCBI[2] as appropriate, using Bowtie2

[1] http://www.bioinformatics.babraham.ac.uk/projects/fastqc

[2] ftp://ftp.ncbi.nlm.nih.gov/genomes/all/

version 2.2.6 (Langmead and Salzberg, 2012) with one allowed mismatch and seed set to 17 base pairs. As a result, a bam file was produced for each sample.

Detection of Previously Unannotated Transcripts

Units of expression were defined as transcriptional peaks (TP) following the methodology described in Gabaldón et al. (2020). Briefly, the TPs were generated by merging all the bam files and selecting the coordinates of the expression peaks (more than 10 per base). For the subsequent analysis, the transcripts with moderate or high expression (10 or more reads per nucleotide) were kept. Finally, a comparison of the TP obtained with annotations reported in databases was made and classified according to their genomic context.

Matching Genomic Sequences Between Species

To identify nearly perfect complementary transcripts between *C. elegans* and *P. aeruginosa*, we first needed to distinguish those reads belonging to each species. To this end, we considered that *bona fide* bacterial transcripts need to be expressed in naïve *P. aeruginosa*. We aligned reads from the sRNA transcriptome of naïve bacteria against the *C. elegans* genome with Bowtie2 (Langmead and Salzberg, 2012), allowing one mismatch and setting the seed to 17 bp. Next, to evaluate the expression of these transcripts in intestinal bacteria, we mapped reads from the holobiont bacteria–nematode against *C. elegans* (WBcel235) and *P. aeruginosa* PAO1 genome with the same parameters. Finally, we counted the naïve bacterial reads that mapped *C. elegans* features with HTSeq-count. Reads from the bacteria–nematode holobiont that were indistinguishable were counted against the features that previously matched with *P. aeruginosa*-naïve reads.

Matching Transcriptional Expression in the Holobiont

Transcriptional peaks from *P. aeruginosa* PAO1 in dataset from Legüe et al. (2021) were compared with *C. elegans* TPs in dataset from Gabaldón et al. (2020) by Reciprocal Best Hit (RBH) analysis using BLAST+ (Altschul et al., 1990; Camacho et al., 2009). The rationale is that if two genes in different genomes are the best hit of each other in the other genome, they are orthologs candidates (Moreno-Hagelsieb and Latimer, 2008). We performed a reciprocal blast using nucleotides and with a discontiguous megablast task, which is optimized for interspecies comparison. Our dataset contained short sequences ranging from 18 to 200 nt. Therefore, to avoid spurious matches, we set the *e*-value < 0.01 for considering a hit.

Bacterial RNAs Systematically Overexpressed During Bacteria–Host Interaction

Pseudomonas aeruginosa sRNAs overexpressed in the intestines of the F1, F2, F5, and F6 nematode generations compared to naïve bacteria were operationally defined as *induced sRNAs*. We defined overexpression as a log2 fold change 1.5 and $p_{adj} < 0.01$. *P. aeruginosa* sRNAs whose expression level did not change during exposure to the animal intestine compared to the naïve condition were called *constitutive*. We selected those with log2 fold change less than 0.5 and p_{adj} value > 0.85. As input for selecting *induced and constitutive sRNAs* in *P. aeruginosa*, we used the differential expression data reported in Legüe et al. (2021). This sRNA annotation was based on TPs that allowed the identification of RNA fragments based on their genomic coordinates (Gabaldón et al., 2020; Legüe et al., 2021).

Bacterial microRNA-Like sRNAs Targeting Host mRNAs

We use the rationale that bacterial transcripts with a size similar to microRNAs (miRNAs) could act over host mRNAs using miRNA-induced post-transcriptional inhibition. We defined putative bacterial miRNA-like genes as those transcripts or fragments with size between 17 and 28 nt, filtered from Legüe et al. (2021). We selected as potential targets of miRNA-like RNAs the host-downregulated mRNAs reported in Gabaldón et al. (2020). The putative interspecies interactions between miRNA-sized-induced sRNAs from bacteria and host-downregulated mRNAs were assessed with IntaRNA 2.0 (Mann et al., 2017). We set the temperature parameter to 20°C and seed length as default.

Eukaryotic Regulatory Motifs Contained in Bacterial Transcripts

The functional eukaryotic motifs we focused on are listed in **Supplementary File 2**. For the identification of these motifs in the sequences of bacterial sRNA transcripts, the online tool RegRNA 2.0 (Chang et al., 2013) was used. This tool identifies regulatory RNA motifs and elements in mRNA sequences by integrating information on regulatory elements from data sources and analytical approaches (Chang et al., 2013). The *induced* and *constitutive* sRNAs mentioned above were used as input sequences. The Fisher's exact test was applied to discriminate the statistical significance of motifs in *induced* and *constitutive* sRNAs.

Enrichment Analysis

Enrichment analysis for tissue, phenotype, and gene ontology for *C. elegans* genes that are putative targets of the induced sRNAs from bacteria was performed using the enrichment tool from www.wormbase.org database (Angeles-Albores et al., 2016, 2018).

Cross-Kingdom RNA Network Construction

For constructing this network, we assumed that sRNAs coming from bacteria and nematodes are co-localized and co-expressed. The network was constructed using as nodes those *P. aeruginosa* RNAs that emerged as *induced*, *constitutive*, *microRNA-sized*, and *common TPs*, from the predictions of this work. Additionally, *C. elegans* transcripts that qualified

as targets in any of the aforementioned predictions were also used as nodes. Attributes considered for nodes were differential expression, their organism, and their biotype, and for bacteria, whether the sRNAs were *induced* or *constitutive* and microRNA-sized. Edge types considered the inhibitory or activating potential of interaction and their putative interaction mechanism. Edges were based on all the cross-kingdom interactions predicted in this paper. Statistical analysis and graph visualization were performed in Gephi 0.9.2. To understand the general structure of the network, statistical analysis was executed with the built-in tools in Gephi 0.9.2, to calculate the following metrics: average degree, graph density, modularity, Eigenvector centrality, average path length, and average cluster coefficient. For graph visualization, a Force Atlas layout was set. Adjusted repulsion strength was adjusted to 500, maintaining the other parameters as default. Node sizes were ranked by betweenness centrality, and color partitions were made by biotype.

Data Availability

Raw data were deposited in the NCBI under BioProject no. PRJNA659467. All the scripts used in these analyses are available at Bitbucket: https://bitbucket.org/srnainterspeciesinteraction/workspace/projects/ISIS_RNA.

C. elegans and Bacterial Growth

Wild-type and mutant *C. elegans* strains were grown at 20°C as previously described (Brenner, 1974). All nematode strains were grown on *E. coli* OP50 before pathogen exposure. *P. aeruginosa* PAO1 (ATCC 15692) was used for infection protocols. Bacteria were grown overnight on Luria-Bertani (LB) plates at 37°C from glycerol stocks. The next morning, a large amount of the bacterial lawn is inoculated in LB broth and grown for 6 h at 250 rpm and at 37°C. Three milliliters of the resulting bacterial culture is seeded onto 90-mm NGM plates and allowed to dry for 36 h before worms are placed on them.

C. elegans Strains

We used the following strains of the *Caenorhabditis* Genetics Center (CGC): wild type (N2), VC1937 (*elpc-2*), and VC463 (*rsp-2*).

C. elegans Growth in Pathogenic Bacteria

Five L4 (P0) wild-type worms or mutants previously maintained in *E. coli* OP50 were picked and transferred to a 90-mm-diameter plate seeded with 3 ml of *P. aeruginosa* PAO1 or *E. coli* OP50 control bacteria. In all cases, the bacterial lawn covered the plate. After 8 days, the total number of worms and dauer larvae were quantified. The number of bacteria seeded allowed animals to be well fed for the length of the experiment. If worms starved, the experiment was discarded. Each assay was performed in three independent experiments (technical replicates) generating a biological replica. A total of three biological replicates were considered for each analysis.

Quantification of Population and Dauer Larvae

Dauer Formation on Pathogens

Entire worm populations on each plate were collected in 1 ml of M9. This initial stock was diluted 1:10 in M9. To count the total population of worms under a Nikon SMZ745 stereomicroscope, 10 μl of this 1:10 dilution was used. To quantify the number of dauers in each population, the initial stock was diluted 1:10 in a 1% SDS solution and maintained in constant agitation for 20 min (Cassada and Russell, 1975). To count the number of total animals and dauers, 10 μl of this last dilution was placed in a glass slide under the stereomicroscope. Each condition was scored three times (triplicates of each technical replica), and dauers were plotted as a percentage of the total populations of animals.

RESULTS

Matching sRNA Expressed Sequences From Bacteria With the Genome of *C. elegans*

The sustained interaction between *P. aeruginosa* PAO1 and *C. elegans* triggers PIDF, a protective response to infection in host progenies (Palominos et al., 2017; **Figure 1A**). PIDF requires the RNA interference machinery of the host and is maintained transgenerationally. We hypothesize that cross-kingdom RNA interactions in the holobiont underlie this behavior. From many possible mechanisms of RNA–RNA interaction, we first examined the common sequences that could act as a *cis*-encoded regulatory element in bacteria (Waters and Storz, 2009; Storz et al., 2011) or as small interfering RNAs in the nematode host. In this scenario, interspecies sequences would interact with nearly perfect complementarity. To prove this, we compared public genomic data and transcriptomic data generated in a previous work (Legüe et al., 2021) from bacteria grown on standard LB broth (naïve bacteria) and for six generations in the transgenerational paradigm shown in **Figure 1A**. We searched for sRNAs from naïve *P. aeruginosa* PAO1 that matched regions of the *C. elegans* genome (**Figure 1B**). We aligned the *P. aeruginosa* PAO1 reads against the *C. elegans* genome WBcel235 with Bowtie2 (Langmead and Salzberg, 2012). Accordingly, from the cross-mapped reads between the two species, we selected those with annotations in the nematode (**Figure 1C**). We found that naïve *P. aeruginosa* reads map to 88 *C. elegans* annotations (**Supplementary Table 1**). Half of these features matched genes that can be expressed either as coding or non-coding isoforms (**Supplementary Figure 1A**). Matching non-coding sRNAs include ncRNA, Piwi-interacting RNAs (piRNA), tRNA, and one pre-miRNA. The *C. elegans* gene with the most abundant matching reads from naïve *P. aeruginosa* is the ribosomal RNA *rrn-3.1* (**Supplementary Figure 1B**). Because the sequence identity potentially allows the interaction

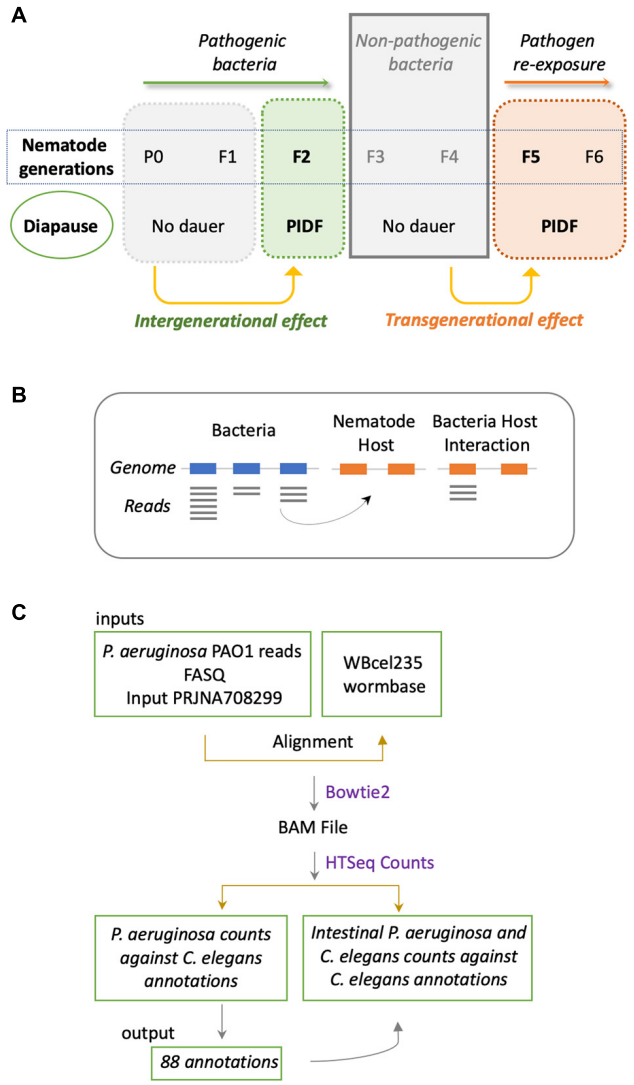


FIGURE 1 | (A) Schematic representation of the inter- and transgenerational paradigm of pathogen-induced diapause formation (PIDF). Animals chronically feeding on pathogenic bacteria such as *P. aeruginosa* PAO1 form dauers in the second generation of exposure to the bacteria. Re-exposure to pathogens after two generations on non-pathogenic bacteria triggers the immediate formation of the dauer larvae. **(B)** Representation of the putative mechanism of interaction based on matching sRNA sequences from bacteria with genomic sequences of *C. elegans*. **(C)** Methodological flow to address the mechanism proposed in panel **(B)**.

between the two transcripts, we speculate that if this mechanism occurs, it should reflect in expression changes upon intestinal interaction between the two species. Therefore, we evaluated the expression of matching *P. aeruginosa*–*C. elegans* sequences (genome level) in the condition of bacteria colonizing the nematode intestine (transcription level) taking data from dual RNA-sequencing experiments mentioned before. We focused on matching transcripts highly expressed in naïve bacteria (more than 10 TMM, **Table 1**) and explored their expression change in the first and second intestinal generations: While *rrn-3.1*

TABLE 1 | Non-coding *C. elegans* genes matching bacterial reads with high expression in host-naïve and intestinal conditions.

***C. elegans* annotation**	**Biotype**	**Bacterial Read Count (TPMs)**		
		Host-Naïve	**F1**	**F2**
rrn-3.1	rRNA	74997.5	164489.2	119781.3
21ur-6043	piRNA	144471.0	1958.7	3130.2
mir-235	pre_miRNA	46295.1	327.6	175.9
K07C10.5	ncRNA	292653.5	164.3	659.5
M163.t1	tRNA	13745.4	160.8	88.8
C06E4.17	ncRNA	50543.3	135.2	988.4
C02F4.11		2221.5	129.7	83.8
21ur-3672	piRNA	19341.1	95.8	0.0
Y43F8A.8	ncRNA	7011.9	84.1	158.2
C43H6.11		7908.7	59.0	98.6
W05B2.13		2937.3	40.0	152.8
K05G3.t1	tRNA	5368.3	34.5	0.0
W03G11.8	ncRNA	4838.9	34.1	17.3

matching sequences in *P. aeruginosa* increased dramatically in read number in the intestinal conditions, all other genes that were highly expressed in naïve bacteria dropped their expression. Based on that, we conclude that genomic matches are not insightful enough to assert physiological relevance.

Matching Transcriptional Expression in the Holobiont

The cross-mapping approach (**Figures 1B,C**) has two main conceptual gaps: (1) the lack of information on how reads are distributed along feature coordinates; and (2) the uncertainty of whether reads are simultaneously co-expressed in both organisms to make possible their interaction. To address these shortcomings, we used synchronous dual RNA-sequencing data generated by us of bacteria and nematodes under PIDF-inducing conditions, both intergenerational (Gabaldón et al., 2020) and transgenerational (Legüe et al., 2021). In these datasets, the annotation of transcripts was based on TPs, which consider pervasive transcription (Lybecker et al., 2014) and add genomic context information (Gabaldón et al., 2020). We classify the TPs as matching annotations, nested, or overlapped in genomic features or novel fragments situated in intergenic regions. The relevance of this approach is that it uncovers functional sRNA fragments previously unannotated or intragenic sense-encoded transcripts (Huang et al., 2010). We hypothesize that co-expressed sequences in both species could mimic each other or form heterologous RNA–RNA duplexes (**Figure 2A**), changing the RNA homeostasis by affecting regulatory loops or RNA sponge occupation (Azam and Vanderpool, 2015). To define interspecies common sequences, we looked for perfectly matching sequences between expressed sequences, by RBH analysis with the BLAST+ tool (Camacho et al., 2009). **Figure 2B** shows the workflow, where we used the TPs expressed in *P. aeruginosa* (454) and *C. elegans* (7,964) during any generation of their interaction (Legüe et al., 2021). Four transcripts mutually hit each other on the same strand and one in the opposite

TABLE 2 | Reciprocal best hits of TPs expressed in *P. aeruginosa* PAO1 and *C. elegans* simultaneously.

P. aeruginosa		*C. elegans*			*P. aeruginosa–C. elegans* interaction		
Gene name	**Protein name**	**Gene name**	**Protein name**	**Tissue**	**% Identity**	**mismatches**	***E*-value**
rhlE	ATP-dependent RNA helicase	*laf-1*	RNA strand annealing activity; RNA-dependent ATPase	P granule/germline	80	8	6.54E-05
Lpd	Dihydrolipoamide dehydrogenases	*dld-1*	Dihydrolipoyl dehydrogenase activity and flavin adenine dinucleotide binding activity	Germline precursor cell	78	9	0.003
SucC	Succinyl-CoA syntetase	*suca-1*	Succinate-CoA ligase ADP-forming subunit beta	Anterior gonad arm	82	9	1.49E-07
GlgP	Glycogen phosphorylase	*pygl-1*	Glycogen phosphorylase and pyridoxal phosphate binding activity	Germline	81	9	8.11E-07
hscA	Heat shock protein	*hsp-1(sj/hsp-2(as)*	Heat shock protein/pseudogene	Germline and intestine	80	9	6.35E-06

s, sense; as, antisense.

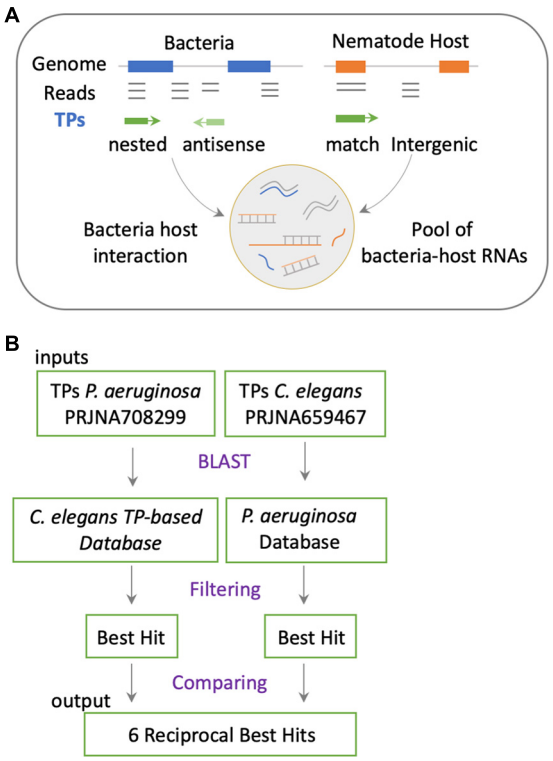


FIGURE 2 | (A) Representation of the putative mechanism of interaction based on transcripts expressed in holobiont and that have the potential of sequence matching. This approach considers actual transcriptomic data of both species in a synchronous paradigm. **(B)** Methodological flow to address the mechanism proposed in panel **(A)**.

strand (**Table 2**). These hits are nested in proteins with similarly described functions, as seen in **Table 2**, suggesting that they could be functional orthologs.

A Core of Bacterial RNAs Is Systematically Overexpressed During Bacteria–Host Interaction

The colonization of the *C. elegans* intestine by *P. aeruginosa* for two continued generations is a requisite for PIDF to take place and highlights the bidirectional molecular interaction between microbe and host (Palominos et al., 2017). This interaction also leaves a memory that allows the defensive strategy to be established immediately upon reencounter (F5 in paradigm of **Figure 1A**). We reasoned that pathogen sRNAs with a role in triggering PIDF and the subsequent transgenerational memory are overexpressed in the intestines of generations F1, F2, F5, and F6 compared to host-naïve bacteria (**Figure 3A**), herein *induced RNAs* (log2FC $>$ 0.5 and $p_{adj} < 0.01$). Forty-four genes were upregulated in intestinal bacteria in the four generations compared to the naïve condition (**Supplementary Table 2**). According to our rationale, these bacterial genes would generate molecular interactions with host genes that are conducive to behavioral changes. On the other hand, we hypothesized that genes constitutively expressed in both intestinal and naïve

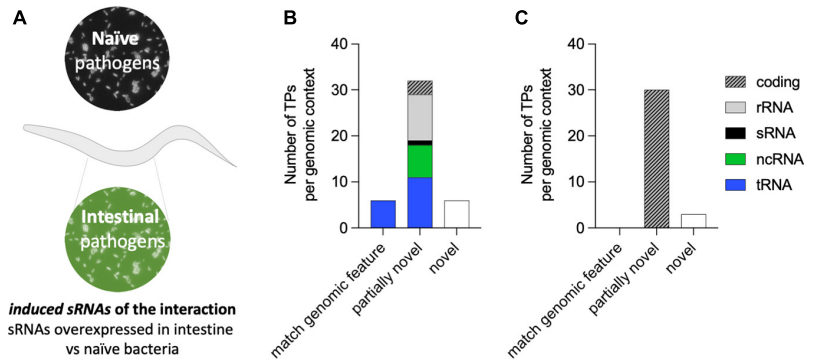


FIGURE 3 | (A) Diagram of the concept of bacterial-induced small RNAs (sRNAs). Induced sRNAs are those bacterial transcripts showing upregulation in all intestinal conditions (F1, F2, F5, and F6) compared to their expression in naïve bacteria. **(B,C)** RNA biotypes and genomic contexts of the induced bacterial sRNA **(B)** and constitutively expressed sRNAs **(C)** from *P. aeruginosa*. The genomic context category of matching genomic feature refers to transcripts that fit the size and genomic coordinate of an existing feature, and the category partially novel refers to transcripts that are either nested or overlapping a previously annotated genomic feature. The category novel refers to a transcript matching an intergenic region of the genome.

conditions (log2FC < 0.5 and $p_{adj} > 0.85$) are unresponsive to the interaction with the nematode, herein *constitutive sRNAs*. Thirty-three genes were constitutively expressed in naïve and intestinal conditions (**Supplementary Table 3**). Interestingly, induced RNAs were predominantly nested in non-coding transcripts such as tRNAs and rRNAs (**Figure 3B**), while constitutive sRNAs were mostly nested in coding transcripts (**Figure 3C**). At the moment, we do not know the implication of this finding, but it may highlight the importance of transcription in non-coding regions in the adaptation to new environments. Throughout the following text, the *induced RNAs* and their comparison with constitutively expressed RNAs serve as a source for predictive interaction analysis.

Exploring the miRNA-Like Mechanisms of Bacterial sRNAs Targeting Host mRNA

Bacterial sRNAs can regulate host mRNAs using the RNAi machinery in a similar way to an endogenous microRNA (Cardin and Borchert, 2017; Zhao et al., 2017; **Figure 4A**). This mechanism is analogous to post-transcriptional inhibition by bacterial *trans*-encoded base-pairing RNAs, which act via limited base pairing (Waters and Storz, 2009; Storz et al., 2011). We aimed to find feasible miRNA-like interactions between the *induced* sRNAs from *P. aeruginosa* and *C. elegans* mRNAs. To that end, we first selected from all TPs from bacteria expressed upon interaction with the nematode (Legüe et al., 2021) those with a length between 17 and 28 nucleotides, hereby *microRNA-sized* (**Figure 4B**). From the 456 bacterial genes expressed in either F1, F2, F5, or F6 nematode generations (**Supplementary Table 4**), we found 44 *microRNA-sized* transcripts (**Figure 4C**). Among those, five were *induced sRNAs*, all of which were either nested or antisense to tRNAs (**Figures 4D,E**). Conversely, two of the constitutively expressed transcripts was *microRNA-sized*, one of them nested in a protein-coding gene and one novel (**Supplementary Table 5**).

Based on the expectation that the regulation of a target by a miRNA is repressive, we used as putative targets 313 downregulated polyA+ *C. elegans*-coding genes upon PIDF (Gabaldón et al., 2020) and their isoforms. We predicted miRNA-like interactions using the IntaRNA 2.0 tool (Mann et al., 2017), using a cutoff of −8 (Trotta, 2014) for the minimum free energy (MFE) of interaction. Two hundred thirty-six genes and 1,096 isoforms surpass this cutoff (**Supplementary File 1**). The complete interaction data and selected datasets are shown in **Supplementary File 1**. We speculate that the more times a polyA+ is predicted to be a target of a bacterial RNA, the probability of being a *bona fide* target increases. One hundred fifty-five mRNAs were targeted by the five *microRNA-sized*-induced sRNAs (**Figure 4F** and **Supplementary Table 6**). The 155 targets had three putative sites for the binding of each *microRNA-sized* sequence. The candidate target genes (155) were subject to tissue, phenotype, and gene ontology enrichment analysis (Angeles-Albores et al., 2016, 2018). These genes show intestinal expression in *C. elegans* with a phenotypical enrichment in intestinal uptake (**Figure 4G**). We analyzed the interactions for all *microRNA-sized* sRNAs expressed in any of the intestinal generations (**Supplementary Tables 4**, **7**). **Supplementary Figure 2** shows that the five *microRNA-sized* sRNAs and the expressed sRNAs in any intestinal generation share most predicted targets. This suggests that the expression and not necessarily the overexpression of a bacterial sRNA can impact the global pool of mRNAs in the host.

Eukaryotic Regulatory Motifs Contained in Bacterial-Induced sRNAs

Numerous functional RNA motifs in eukaryotes control gene expression. These motifs are defined by specific sequences or secondary structures that bind or are targeted by transcriptional or translational regulators, such as transcription factors (Morris, 2011), splicing enhancers or inhibitors, RNA editing sites, and other functional sequences (Mattick and Makunin, 2006). To the extent of our knowledge, the presence of the aforementioned motifs in bacterial RNAs has not been explored in an interaction

paradigm such as PIDF. Therefore, we aimed to identify bacterial-induced sRNAs motifs that could potentially bind eukaryotic gene expression regulators. We tested a wide variety of motif types potentially present in *induced RNAs* by using RegRNA 2.0 (**Figure 5**; Chang et al., 2013). As a control, we performed the same analysis with the constitutive sRNAs whose expression is not modified during interaction.

We found that induced and constitutive RNAs did not differ in the total number of eukaryotic-regulatory motifs (**Supplementary File 2**), both displaying a high number of them. The common motifs shared by induced and constitutive sRNAs (total number of 146 and 77, respectively) were varied. By far, the most frequent motifs were *splicing regulators* (114/146 and 68/77). In both *induced* and *constitutive* sRNAs, the exon splicing enhancers (ESE) predominate with 52 and 62%, respectively, followed by intron splicing enhancer (ISE) (16.9 and 19.4%, respectively) motifs. Interestingly, the most frequent ESE motifs were members of the serine-rich protein family, such as *srp40* (27/114–23.7% of splicing regulatory motifs). *srp40* and *srp55* are homologous to *C. elegans* SR protein RSP-2, which is involved in larval development. This suggests that commonly expressed bacterial sRNAs should carry different splicing regulatory motifs. However, in our paradigm, their presence is not increased with the interaction with hosts. However, the type of motifs was different between groups, being the *functional regulatory sequences* and *RNA editing sites* significantly overrepresented in the *induced sRNAs* (Fisher's exact test, $p = 0.038$ and $p = 0.04$, respectively). The *functional regulatory sequences* corresponded mainly to tRNA motifs (6/7) and p16, also termed RgsA, involved in quorum sensing and regulated by GacS/GacA system (Brencic et al., 2009). This could suggest a role of tRNAs and their modifications in interspecies interaction.

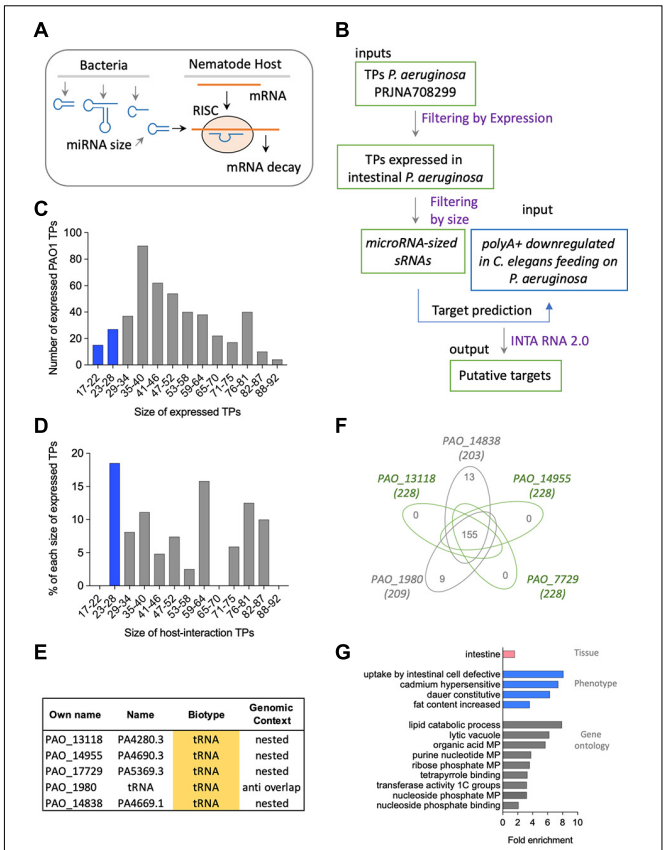


FIGURE 4 | (A) Representation of the putative mechanism of interaction based on *miRNA-like* mechanisms of bacterial sRNAs targeting host mRNA. **(B)** Methodological flow to address the mechanism proposed in panel **(A)**. **(C)** Size and number of expressed transcriptional peaks (TPs) in *P. aeruginosa* PAO1. **(D)** Size and number of induced RNA TPs in *P. aeruginosa* PAO1. **(E)** Biotype and genomic context of *microRNA-sized*-induced TPs in *P. aeruginosa* PAO1. **(F)** Venn diagram of mRNAs targeted by the five miRNA-like-induced sRNAs. **(G)** Tissue (pink), phenotype (blue), and gene ontology (gray) enrichment of mRNAs targeted by the five miRNA-like-induced sRNAs.

Integrative Analysis of the Cross-Kingdom RNA Interactome Network

To gain insight on co-occurrence and co-regulation between the bacterial sRNAs and host targets, we integrated all the mechanisms of interspecies interaction proposed in this work and constructed a global RNA interaction network of bacteria and host putative interactors (**Supplementary File 3**). In order to classify the main interrelations, we use as input those *P. aeruginosa* RNAs that emerged as *induced*, *constitutive*, *microRNA-sized*, and *common TPs*, from the predictions of this work. Additionally, *C. elegans* transcripts that qualified as targets in any of the aforementioned predictions were used as nodes. Attributes considered for nodes were as follows: differential expression, organism type, biotype, and in bacteria, whether the sRNAs are *induced* or *constitutive* and *microRNA-sized* or not. The network consisted of 456 nodes and 1,569 edges, with an average weighted degree of 5.7 and 6.8, respectively, and a power law degree distribution. Network metrics are reported in **Supplementary File 3**. Nodes that correspond to *C. elegans* were 78.4, and 21.6% to *P. aeruginosa*. The most represented biotype in the network was host mRNA regulatory motifs and bacterial microRNA-sized sRNAs, followed by tRNAs and tRNA fragments of both species. We discriminate influential nodes based on the betweenness centrality and closeness centrality metrics (**Figure 6**). The biotypes with the higher betweenness centrality were tRNAs and tRNA-related motifs, from both species. Nodes with the higher betweenness centrality in bacteria were enriched in quorum sensing-related processes, such as *RhlE*, *RgsA*, and *CrcZ*. We also found an alternative splicing motif of fibronectin (EDA exon) as relevant. The modularity analysis renders four main modules. The cluster with the highest number of nodes corresponded to microRNAs' predicted mechanism. The integration of RNA interactions in this network allowed us to uncover potentially relevant players of PIDF and pathogenesis in our paradigm, some of which can be further tested by mutant and phenotypic analysis.

In vivo Phenotypic Analysis

Through the integration of diverse *in silico* analysis, we found that the overrepresentation of tRNA motifs and RNA

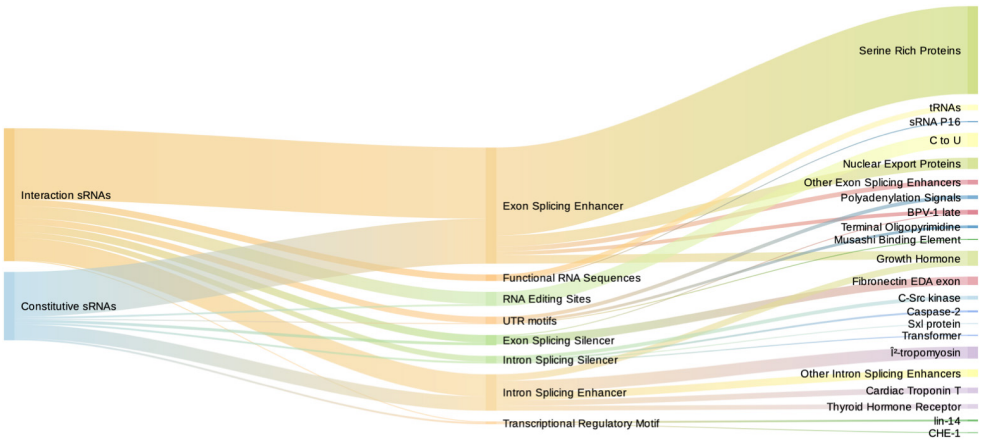


FIGURE 5 | Sankey diagram depicting the number of motifs of induced sRNAs (red) and constitutive sRNAs (blue) in the first column, motif type in the second, and specific targets in the third column. Intron and exon splicing regulatory motifs are similarly represented among induced and constitutive sRNAs. Functional RNA sequences (comprising tRNAs and sRNA P16) are exclusive from induced sRNAs.

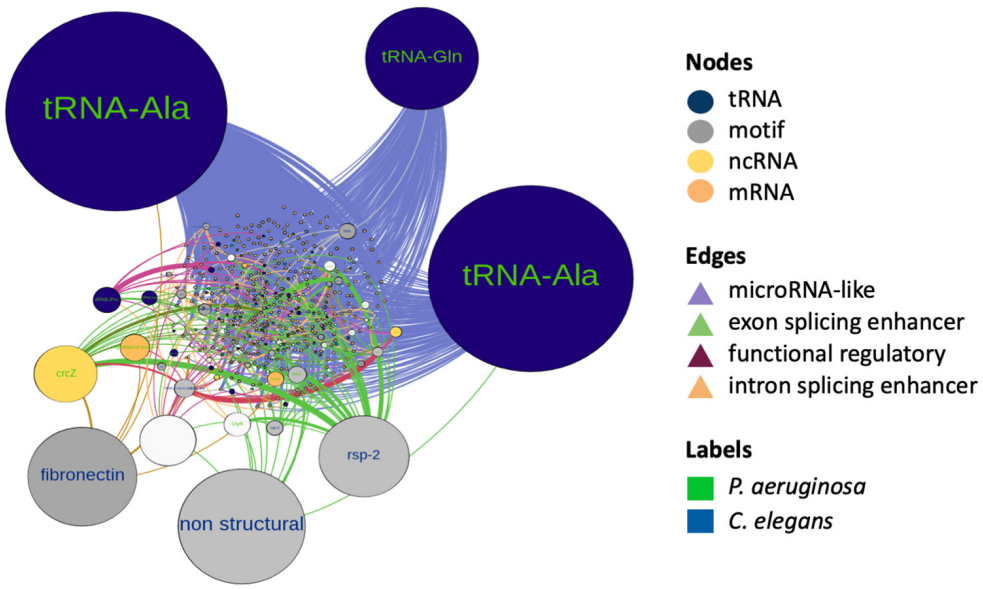


FIGURE 6 | RNA interactome network of cross-kingdom predictions between bacterial non-coding sRNAs and host sRNAs and long non-coding RNAs and coding RNA candidate targets. Node colors represent biotypes and are differentiated by organisms. Nodes are scaled by size according to their betweenness centrality. The edges represent predicted interactions of this work and are colored according to the putative mechanism.

editing sites suggested that base modification in tRNAs could be implicated in interspecies interaction. Previous works have shown that tRNA base modifications determine *C. elegans* behavior (Fernandes De Abreu et al., 2020). We tested the possibility that modified tRNAs are good candidates for testing interspecies elicitation of PIDF. We asked whether worms that lack the ability to modify tRNAs show abnormal PIDF. We chose mutant nematodes defective of *elpc-2* and *elpc-3*, an elongation factor required for tRNA base modification, to study their potential role in PIDF. We first quantified the growth of *elpc-3* and *rsp-2* animals in *P. aeruginosa* PAO1 and compared it with wild-type animals. **Figure 7A** shows that growth of mutants is similar to the wild-type strain, indicating that the lack of *elpc-3* and *rsp-2* does not impair growth nor the wild-type innate response required for development under mild pathogens. We then quantified dauer formation after two generations feeding on *P. aeruginosa* PAO1. Interestingly, both mutants were able to form large numbers of dauers in pathogenic foods (**Figure 7B**), showing that neither is defective in PIDF. Furthermore, *elpc-3* animals formed significantly more dauers under pathogens, suggesting that tRNAs are halting the dauer defensive mechanism.

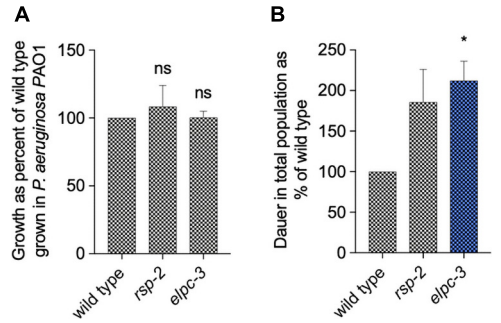


FIGURE 7 | Growth **(A)** and diapause formation **(B)** of wild-type, *elpc-2*, and *rsp-2 C. elegans* mutants on *P. aeruginosa* PAO1 for two generations. *$P < 0.05$.

DISCUSSION

We investigated putative mechanisms by which RNAs from *P. aeruginosa* could target host RNAs and proteins to promote a heritable response and memory of the infection. We use the PIDF transgenerational paradigm of defense against pathogens where animals enter diapause when exposed to harmful bacteria. The analysis is based on RNA transcripts since sRNAs from *C. elegans* and bacteria have been described as necessary for PIDF (Gabaldón et al., 2020; Legüe et al., 2021). We approached this question by applying *in silico* tools that operationalize possible mechanistic frameworks of interaction between RNAs from two species. These analyses generate potential candidate triggers and targets of the interspecies communication to be tested *in vivo* for their role in specific life history traits. For example, here we show that effectors of tRNA modifications are relevant for wild-type PIDF.

From Context-Specific RNA Expression to *in silico* Tools to Address RNA-Based Interspecies Communication

We selected the PIDF behavioral paradigm that results from the RNA interplay between *P. aeruginosa* and *C. elegans* for two generations. PIDF is triggered rapidly after animals are re-exposed to pathogens after having been withdrawn for two generations. We previously produced transcriptomic profiles of animals and their intestinal bacteria for two and six generations, respectively (Gabaldón et al., 2020; Legüe et al., 2021). We focus on bacterial sRNAs always upregulated in the intestines of the nematodes (induced sRNAs) compared to naïve conditions and contrast them with sRNAs constitutively expressed across generations. We complemented the conventional sRNA annotation with the information of the peaks of expression (TPs, Gabaldón et al., 2020), enabling the scrutiny of novel sRNAs and RNA fragments with potential biological relevance. The dual RNA-sequencing analysis allows us to study the simultaneous gene expression changes during the pathogen–host interaction.

The interaction analysis we perform is built on the following assumptions: First, RNA from both species is highly mobile, capable of being transferred between organisms by RNA transporters or membrane vesicles. These RNA species constitute the holo-transcriptome where RNAs expressed by either species are susceptible to interacting with each other. The RNA expression is context-dependent and tissue-specific. While we have taken context-specific transcriptomics data in highly synchronized populations, the extraction was done from whole animals. A form to solve this shortcoming would be to use data from single-cell-type transcriptomics in future investigations. Alternatively, individual interactions predicted by our analysis could be tested in specific tissues and under specific conditions. Second, we assume that RNAs induced under encounter with the *C. elegans* intestine are those most relevant for the interaction with the host because it acknowledges the connection of functionally relevant transcripts with PIDF. Third, we consider the pervasiveness of transcription (Leitão et al., 2020) and use a method that estimates as equally likely the expression of transcripts that have been previously annotated and those unannotated or expressed from intergenic regions.

The first mechanism evaluated was the interspecies post-transcriptional inhibition triggered by perfectly or incompletely paired sequences such as siRNA, miRNA-like, and acting sRNAs (asRNA). We also implemented approaches previously reported such as genomic alignment (Celluzzi and Masotti, 2016; Kaletsky et al., 2020). For siRNA exploration, we conducted the mapping of reads with Bowtie2, which is a validated and standard tool for global alignment. We chose this tool based on its use (Celluzzi and Masotti, 2016) even though the global alignment is not the best choice for dissimilar sequences such as those from different species. Therefore, we looked for paired transcripts with local alignment using BLAST (Camacho et al., 2009) and restricted it to expressed fragments, which allowed us to find less but more precise matches. It is relevant to highlight that the BLAST alignment does not constitute evidence for homology by itself but of sequence similarity. For selecting biologically relevant sequences, it is imperative to filter these results by *e*-value. We selected those with an *e*-value lower than 0.01, keeping in mind that small fragments have more chance of aligning randomly. This threshold is more restrictive than the one used by other authors (Kaletsky et al., 2020). Furthermore, even though interspecies matching sequences meet the statistical criteria of significance, they have a bit-score lower than 50, so we do not call them homologous despite their similarity. We additionally use the reciprocal best-hit analysis (Ward and Moreno-Hagelsieb, 2014), which consists in keeping the alignments in which both sequences are the best match of each other. This method was first used for finding coding sequences or protein orthologs but proved helpful in our analyses to keep the best matches between species.

The other post-transcriptional mechanism explored was the interactions between bacterial sRNA microRNA-sized with *C. elegans* mRNAs. In bacteria, *trans*-acting sRNAs (*trans*-asRNAs) share many characteristics with eukaryotic miRNAs (Layton et al., 2020) such as a limited base-pairing mechanism requiring only partial complementarity to their target sequence and a seed region of 7–12 nucleotides (Wachter et al., 2019).

Eukaryotic miRNA function requires the RISC, which is absent in bacteria. However bacterial Hfq protein analogously presents and stabilizes sRNAs (De Lay et al., 2013; Diallo and Provost, 2020). *Trans*-asRNAs are typically 100–1,000 nucleotides, in contrast to microRNAs that are 21–25 nucleotides (Layton et al., 2020), but we know that shorter RNAs in bacteria have been largely overlooked despite the indications that they could be relevant players in interspecies communication (Zhao et al., 2017; Felden and Gilot, 2018; González Plaza, 2020; Layton et al., 2020). Interestingly, we found microRNA-sized fragments in the group of sRNAs responsive to interaction or *induced RNAs*, but not in the constitutive sRNA group. We found that these microRNA-sized transcripts mainly target intestinally expressed host mRNAs. This fact is concordant with our model in which only live bacteria colonizing the intestine can induce the PIDF response. These results allow us to speculate that microRNA-sized RNAs could be affecting locally the transcriptional state of the host at the site of the infection.

The evaluation of splicing regulatory motifs in bacteria–host interaction could appear counterintuitive, given that alternative splicing is not a relevant mechanism of gene expression control in bacteria. However, there is increasing evidence of drastic changes in host splicing regulation during infection (Liang et al., 2016; Kalam et al., 2018; Chauhan et al., 2019). We evaluated the presence of splicing regulatory motifs in bacterial sRNAs and found plenty of predicted enhancer and silencer intron and exon splicing regulators. These motifs were found in both induced and constitutive sRNAs with splicing silencer motifs slightly over-represented in the induced sRNAs. Much needs to be learned on how pathogens target splicing regulation in new scenarios to fully understand the relevance of these motifs in bacterial sRNAs. An open question is whether sRNAs from bacteria can bind to regulatory proteins in the host nematode as they do in bacteria. RNA-binding proteins (RBP) are involved in the processing, stability, and activity of sRNAs, providing precision in sRNA–mRNA base-pairing (Quendera et al., 2020). As newer tools are developed, it would be appropriate to test whether bacterial sRNAs that bind RBP from bacteria could also bind their orthologs in the nematode.

Finally, our analysis highlights tRNAs and tRNA-derived small RNAs (*tsRNAs*) as candidate mediators for interspecies interaction. The challenge of *elpc-3* mutants with *P. aeruginosa* pathogenic bacteria renders increased PIDF, suggesting that tRNA processing is necessary for the wild-type response to infection. tRNAs and tsRNAs have increasingly recognized roles in gene regulation under stress and in intergenerational inheritance (Chen et al., 2016). Indeed, some tsRNAs can be associated with Argonaute proteins and function as miRNAs (Shen et al., 2018). We found that tRNA-derived sRNAs were exclusively *induced sRNAs*, suggesting a role for tRNAs in our paradigm that may transcend their role in protein translation. Moreover, tRNAs decode essential mRNAs for protein synthesis, deliver amino acids to other places in the cell, and under stress conditions can be cleaved to generate signaling molecules or regulate gene expression (Raina and Ibba, 2014; Megel et al., 2015; Fields and Roy, 2018; Oberbauer and Schaefer, 2018; Barraud and Tisné, 2019; Nguyen et al., 2019).

Relevance of RNA–RNA Communication in Behavior and Physiology

Bacterial RNAs are capable of triggering behavioral decisions in their nematode host (Kaletsky et al., 2020; Legüe et al., 2021). How do RNAs from the two species interact with each other in the context of a holobiont? Interspecies interactions and communication involve the bidirectional transport of regulatory molecules. At least two mechanisms are formally possible: sRNA movement through specific RNA transporters and the use of membrane vesicles for their delivery from the intestine to specific tissues. *C. elegans* expresses intestinal dsRNA transporters (SID-1, SID-2, and SID-5) needed for systemic and environmental RNAi (Winston et al., 2002, 2007; Hinas et al., 2012). Theoretically, these transporters could also internalize RNAs derived from colonizing bacteria (Legüe and Calixto, 2019). Bacterial cargo can be released through membrane vesicles of diverse nature. One well-documented example is through OMVs. These vesicles are known to transport a variety of cargoes including RNAs, proteins, and toxins, among other molecules, and play important roles both in pathogenesis (Ghosal et al., 2015; Koeppen et al., 2016; Zhang et al., 2020) and symbiosis (Moriano-Gutierrez et al., 2020). Additional vesicles that can carry RNA cargo are MV resulting from the explosive lysis of bacteria (Turnbull et al., 2016; Toyofuku et al., 2017).

In summary, this study offers a framework to analyze global transcriptomics in the context of survival behaviors against pathogenesis. The complexity of the interplay at the RNA level in interspecies communication underscores that behavioral adaptations are multistep strategies that require the integration of multiple effectors and targets.

DATA AVAILABILITY STATEMENT

The datasets analyzed during the current study are available in the NCBI repository under BioProject no. PRJNA659-467 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA6594-67, Gabaldón et al., 2020) and BioProject no. PRJNA708299 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA708299, Legüe et al., 2021).

AUTHOR CONTRIBUTIONS

ML and AC made significant contributions in the conceptualization of the study, performed the methodology, wrote the original draft, and reviewed and edited the manuscript. AC acquired funding for the study. All authors conducted the investigation.

FUNDING

This study was supported by the Millennium Scientific Initiative ICM-ANID ICN09-022, CINV, Proyecto Apoyo Redes Formación de Centros (REDES180138), CYTED grant P918PTE3, and Fondecyt 1131038 to AC. BA is funded by a doctoral fellowship (21201778) from the National Agency for Research and Development (ANID).

ACKNOWLEDGMENTS

Some strains were provided by the CGC, which is funded by the NIH Office of Research Infrastructure Programs (P40OD010440). The Centro Interdisciplinario de Neurociencia de Valparaiso is a Millennium Institute supported by the Millennium Scientific Initiative of the Chilean Ministry of Economy, Development, and Tourism (P029-022-F).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.649858/full#supplementary-material

Supplementary Figure 1 | (A,B) Tracks in Genome Browser of read location and coverage. gff files loaded onto Wormbase Genome Browser to show the genomic context where the reads are located.

Supplementary Figure 2 | Venn diagram showing shared putative *C. elegans* targets of sRNAs microRNA-sized that are *induced* and expressed in *C. elegans* intestines from *P. aeruginosa*.

Supplementary Table 1 | *Caenorhabditis elegans* genomic annotations mapped by naïve *P. aeruginosa* PAO1 reads. List of *C. elegans* genes mapped by *P. aeruginosa* reads expressed in host-naïve conditions.

Supplementary Table 2 | Host-induced sRNAs from *P. aeruginosa* PAO1. Bacterial sRNAs upregulated in the F1, F2, F5, and F6 *C. elegans* generations compared to naïve bacteria.

Supplementary Table 3 | Constitutively expressed sRNAs from *P. aeruginosa* PAO1. Genes whose expression does not change in intestinal conditions compared to naïve bacteria.

Supplementary Table 4 | All transcripts expressed in intestinal conditions with their lengths. Lists of all genes expressed in any of the F1, F2, F5, and F6 generations of *C. elegans* and further subdivisions as induced and microRNA-sized.

Supplementary Table 5 | Transcripts constitutively expressed with their lengths.

Supplementary Table 6 | *Caenorhabditis elegans* mRNA putative targets of microRNA-sized-induced sRNAs from *P. aeruginosa* PAO1.

Supplementary Table 7 | List of *C. elegans* mRNA targets of *P. aeruginosa* miRNA-like sRNAs expressed intestinally in *C. elegans*.

Supplementary File 1 | All putative interactions of microRNA-sized host-induced sRNAs with *C. elegans* mRNAs downregulated during interaction with *P. aeruginosa* PAO1.

Supplementary File 2 | List of regulatory motifs predicted to be targeted by induced and constitutive sRNAs by RegRNA 2.0.

Supplementary File 3 | Data for construction of the RNA interactome network.

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biol.* 11:R106. doi: 10.1186/gb-2010-11-10-r106

Anders, S., Pyl, P. T., and Huber, W. (2015). HTSeq–a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166–169. doi: 10.1093/bioinformatics/btu638

Angeles-Albores, D., Lee, R. Y. N., Chan, J., and Sternberg, P. W. (2016). Tissue enrichment analysis for *C. elegans* genomics. *BMC Bioinformatics* 17:366. doi: 10.1186/s12859-016-1229-9

Angeles-Albores, D., Lee, R. Y. N., Chan, J., and Sternberg, P. W. (2018). Two new functions in the WormBase enrichment suite. *microPub. Biol.* doi: 10.17912/W25Q2N

Azam, M. S., and Vanderpool, C. K. (2015). Talk among yourselves: RNA sponges mediate cross talk between functionally related messenger RNAs. *EMBO J.* 34, 1436–1438. doi: 10.15252/embj.201591492

Barraud, P., and Tisné, C. (2019). To be or not to be modified: miscellaneous aspects influencing nucleotide modifications in tRNAs. *IUBMB Life* 71, 1126–1140. doi: 10.1002/iub.2041

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Brencic, A., McFarland, K. A., McManus, H. R., Castang, S., Mogno, I., Dove, S. L., et al. (2009). The GacS/GacA signal transduction system of *Pseudomonas aeruginosa* acts exclusively through its control over the transcription of the RsmY and RsmZ regulatory small RNAs. *Mol. Microbiol.* 73, 434–445. doi: 10.1111/j.1365-2958.2009.06782.x

Brenner, S. (1974). The genetics of *Caenorhabditis elegans*. *Genetics* 77, 71–94.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. doi: 10.1186/1471-2105-10-421

Cardin, S. E., and Borchert, G. M. (2017). Viral MicroRNAs, host MicroRNAs regulating viruses, and bacterial MicroRNA-Like RNAs. *Methods Mol. Biol.* 1617, 39–56. doi: 10.1007/978-1-4939-7046-9_3

Cassada, R. C., and Russell, R. L. (1975). The dauerlarva, a post-embryonic developmental variant of the nematode *Caenorhabditis elegans*. *Dev. Biol.* 46, 326–342. doi: 10.1016/0012-1606(75)90109-8

Celluzzi, A., and Masotti, A. (2016). How our other genome controls our epi-genome. *Trends Microbiol.* 24, 777–787. doi: 10.1016/j.tim.2016.05.005

Chang, T. H., Huang, H. Y., Hsu, J. B., Weng, S. L., Horng, J. T., and Huang, H. D. (2013). An enhanced computational platform for investigating the roles of regulatory RNA and for identifying functional RNA motifs. *BMC Bioinformatics* 14(Suppl. 2):S4. doi: 10.1186/1471-2105-14-S2-S4

Chauhan, K., Kalam, H., Dutt, R., and Kumar, D. (2019). RNA splicing: a new paradigm in host-pathogen interactions. *J. Mol. Biol.* 431, 1565–1575. doi: 10.1016/j.jmb.2019.03.001

Chen, Q., Yan, M., Cao, Z., Li, X., Zhang, Y., Shi, J., et al. (2016). Sperm tsRNAs contribute to intergenerational inheritance of an acquired metabolic disorder. *Science* 351, 397–400. doi: 10.1126/science.aad7977

Cleary, J. L., Condren, A. R., Zink, K. E., and Sanchez, L. M. (2017). Calling all hosts: bacterial communication in situ. *Chem* 2, 334–358. doi: 10.1016/j.chempr.2017.02.001

De Lay, N., Schu, D. J., and Gottesman, S. (2013). Bacterial small RNA-based negative regulation: Hfq and its accomplices. *J. Biol. Chem.* 288, 7996–8003. doi: 10.1074/jbc.R112.441386

Diallo, I., and Provost, P. (2020). RNA-sequencing analyses of small bacterial RNAs and their emergence as virulence factors in host-pathogen interactions. *Int. J. Mol. Sci.* 21:1627. doi: 10.3390/ijms21051627

Felden, B., and Gilot, D. (2018). Modulation of bacterial sRNAs activity by epigenetic modifications: inputs from the eukaryotic miRNAs. *Genes* 10:22. doi: 10.3390/genes10010022

Fernandes De Abreu, D. A., Salinas-Giegé, T., Drouard, L., and Remy, J. J. (2020). Alanine tRNAs translate environment into behavior in *Caenorhabditis elegans*. *Front. Cell Dev. Biol.* 8:571359. doi: 10.3389/fcell.2020.571359

Fields, R. N., and Roy, H. (2018). Deciphering the tRNA-dependent lipid aminoacylation systems in bacteria: novel components and structural advances. *RNA Biol.* 15, 480–491. doi: 10.1080/15476286.2017.1356980

Gabaldón, C., Legüe, M., Palominos, M. F., Verdugo, L., Gutzwiller, F., and Calixto, A. (2020). Intergenerational pathogen-induced diapause in *Caenorhabditis*

elegans is modulated by mir-243. *mBio* 11:e01950–20. doi: 10.1128/mBio.01950-20

Ghosal, A., Upadhyaya, B. B., Fritz, J. V., Heintz-Buschart, A., Desai, M. S., Yusuf, D., et al. (2015). The extracellular RNA complement of *Escherichia coli*. *Microbiologyopen* 4, 252–266.

González Plaza, J. J. (2020). Small RNAs as fundamental players in the transference of information during bacterial infectious diseases. *Front. Mol. Biosci.* 7:101. doi: 10.3389/fmolb.2020.00101

Hinas, A., Wright, A. J., and Hunter, C. P. (2012). SID-5 is an endosome-associated protein required for efficient systemic RNAi in *C. elegans*. *Curr. Biol.* 22, 1938–1943. doi: 10.1016/j.cub.2012.08.020

Huang, Y. C., Chen, H. T., and Teng, S. C. (2010). Intragenic transcription of a noncoding RNA modulates expression of ASP3 in budding yeast. *RNA* 16, 2085–2093. doi: 10.1261/rna.2177410

Kalam, H., Singh, K., Chauhan, K., Fontana, M. F., and Kumar, D. (2018). Alternate splicing of transcripts upon *Mycobacterium tuberculosis* infection impacts the expression of functional protein domains. *IUBMB Life* 70, 845–854. doi: 10.1002/iub.1887

Kaletsky, R., Moore, R. S., Vrla, G. D., Parsons, L. R., Gitai, Z., and Murphy, C. T. (2020). *C. elegans* interprets bacterial non-coding RNAs to learn pathogenic avoidance. *Nature* 586, 445–451. doi: 10.1038/s41586-020-2699-5

Koeppen, K., Hampton, T. H., Jarek, M., Scharfe, M., Gerber, S. A., Mielcarz, D. W., et al. (2016). A novel mechanism of host-pathogen interaction through sRNA in bacterial outer membrane vesicles. *PLoS Pathog.* 12:e1005672. doi: 10.1371/journal.ppat.1005672

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

Layton, E., Fairhurst, A. M., Griffiths-Jones, S., Grencis, R. K., and Roberts, I. S. (2020). Regulatory RNAs: a universal language for inter-domain communication. *Int. J. Mol. Sci.* 21:8919. doi: 10.3390/ijms21238919

Lefebvre, F. A., and Lécuyer, E. (2017). Small luggage for a long journey: transfer of vesicle-enclosed small RNA in interspecies communication. *Front. Microbiol.* 8:377. doi: 10.3389/fmicb.2017.00377

Legüe, M. E., Aguila, B., Pollak, B., and Calixto, A. (2021). Bacterial small RNAs and host epigenetic effectors of a transgenerational memory of pathogens in *C. elegans*. *bioRxiv [Preprint]*. doi: 10.1101/2021.03.26.437277

Legüe, M., and Calixto, A. (2019). RNA language in *Caenorhabditis elegans* and bacteria interspecies communication and memory. *Curr. Opin. Syst. Biol.* 13, 16–22. doi: 10.1016/j.coisb.2018.08.005

Leitão, A. L., Costa, M. C., Gabriel, A. F., and Enguita, F. J. (2020). Interspecies communication in holobionts by non-coding RNA exchange. *Int. J. Mol. Sci.* 21:2333. doi: 10.3390/ijms21072333

Liang, G., Malmuthuge, N., Guan, Y., Ren, Y., Griebel, P. J., and Guan, L. L. (2016). Altered microRNA expression and pre-mRNA splicing events reveal new mechanisms associated with early stage *Mycobacterium avium* subspecies paratuberculosis infection. *Sci. Rep.* 6:24964. doi: 10.1038/srep24964

Liu, H., Wang, X., Wang, H. D., Wu, J., Ren, J., Meng, L., et al. (2012). Escherichia coli noncoding RNAs can affect gene expression and physiology of *Caenorhabditis elegans*. *Nat. Commun.* 3:1073. doi: 10.1038/ncomms2071

Liu, S., da Cunha, A. P., Rezende, R. M., Cialic, R., Wei, Z., Bry, L., et al. (2016). The host shapes the gut microbiota via fecal MicroRNA. *Cell Host Microbe* 19, 32–43. doi: 10.1016/j.chom.2015.12.005

Lybecker, M., Bilusic, I., and Raghavan, R. (2014). Pervasive transcription: detecting functional RNAs in bacteria. *Transcription* 5:e944039. doi: 10.4161/21541272.2014.944039

Mann, M., Wright, P. R., and Backofen, R. (2017). IntaRNA 2.0: enhanced and customizable prediction of RNA-RNA interactions. *Nucleic Acids Res.* 45, W435–W439. doi: 10.1093/nar/gkx279

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

Mattick, J. S., and Makunin, I. V. (2006). Non-coding RNA. *Hum. Mol. Genet.* 15, R17–R29. doi: 10.1093/hmg/ddl046

Megel, C., Morelle, G., Lalande, S., Duchêne, A. M., Small, I., and Maréchal-Drouard, L. (2015). Surveillance and cleavage of eukaryotic tRNAs. *Int. J. Mol. Sci.* 16, 1873–1893. doi: 10.3390/ijms16011873

Moënne-Loccoz, Y., Mavingui, P., Combes, C., Normand, P., and Steinberg, C. (2015). "Microorganisms and biotic interactions," in *Environmental Microbiology: Fundamentals and Applications*, eds J. C. Bertrand, P. Caumette, P. Lebaron, R. Matheron, P. Normand, and T. Sime-Ngando (Berlin: Springer), 395–444.

Moreno-Hagelsieb, G., and Latimer, K. (2008). Choosing BLAST options for better detection of orthologs as reciprocal best hits. *Bioinformatics* 24, 319–324. doi: 10.1093/bioinformatics/btm585

Moriano-Gutierrez, S., Bongrand, C., Essock-Burns, T., Wu, L., McFall-Ngai, M. J., and Ruby, E. G. (2020). The noncoding small RNA SsrA is released by Vibrio fischeri and modulates critical host responses. *PLoS Biol.* 18:e3000934. doi: 10.1371/journal.pbio.3000934

Morris, K. V. (2011). The emerging role of RNA in the regulation of gene transcription in human cells. *Semin. Cell Dev. Biol.* 22, 351–358. doi: 10.1016/j.semcdb.2011.02.017

Nguyen, H. A., Hoffer, E. D., and Dunham, C. M. (2019). Importance of a tRNA anticodon loop modification and a conserved, noncanonical anticodon stem pairing in for decoding. *J. Biol. Chem.* 294, 5281–5291. doi: 10.1074/jbc.RA119.007410

Nguyen, T. C., Zaleta-Rivera, K., Huang, X., Dai, X., and Zhong, S. (2018). RNA, action through Interactions. *Trends Genet.* 34, 867–882. doi: 10.1016/j.tig.2018.08.001

Oberbauer, V., and Schaefer, M. R. (2018). tRNA-derived small RNAs: biogenesis, modification, function and potential impact on human disease development. *Genes* 9:607. doi: 10.3390/genes9120607

Palominos, M. F., Verdugo, L., Gabaldon, C., Pollak, B., Ortíz-Severín, J., Varas, M. A., et al. (2017). Transgenerational diapause as an avoidance strategy against bacterial pathogens in *Caenorhabditis elegans*. *mBio* 8, e01234–17. doi: 10.1128/mBio.01234-17

Quendera, A. P., Seixas, A. F., Dos Santos, R. F., Santos, I., Silva, J. P. N., Arraiano, C. M., et al. (2020). RNA-binding proteins driving the regulatory activity of small non-coding RNAs in bacteria. *Front. Mol. Biosci.* 7:78. doi: 10.3389/fmolb.2020.00078

Raina, M., and Ibba, M. (2014). tRNAs as regulators of biological processes. *Front. Genet.* 5:171. doi: 10.3389/fgene.2014.00171

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616

Rosenberg, E., and Zilber-Rosenberg, I. (2013). "Role of microorganisms in adaptation, development, and evolution of animals and plants: the hologenome concept," in *The Prokaryotes–Prokaryotic Biology and Symbiotic Associations*, eds E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, and F. Thompson (Berlin: Springer), 347–358.

Rosenberg, E., and Zilber-Rosenberg, I. (2018). The hologenome concept of evolution after 10 years. *Microbiome* 6:78. doi: 10.1186/s40168-018-0457-9

Shen, Y., Yu, X., Zhu, L., Li, T., Yan, Z., and Guo, J. (2018). Transfer RNA-derived fragments and tRNA halves: biogenesis, biological functions and their roles in diseases. *J. Mol. Med.* 96, 1167–1176. doi: 10.1007/s00109-018-1693-y

Sjöström, A. E., Sandblad, L., Uhlin, B. E., and Wai, S. N. (2015). Membrane vesicle-mediated release of bacterial RNA. *Sci. Rep.* 5:15329. doi: 10.1038/srep15329

Storz, G., Vogel, J., and Wassarman, K. M. (2011). Regulation by small RNAs in bacteria: expanding frontiers. *Mol. Cell* 43, 880–891. doi: 10.1016/j.molcel.2011.08.022

Toyofuku, M., Cárcamo-Oyarce, G., Yamamoto, T., Eisenstein, F., Hsiao, C. C., Kurosawa, M., et al. (2017). Prophage-triggered membrane vesicle formation through peptidoglycan damage in Bacillus subtilis. *Nat. Commun.* 8:481. doi: 10.1038/s41467-017-00492-w

Trapnell, C., Pachter, L., and Salzberg, S. L. (2009). TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25, 1105–1111. doi: 10.1093/bioinformatics/btp120

Trotta, E. (2014). On the normalization of the minimum free energy of RNAs by sequence length. *PLoS One* 9:e113380. doi: 10.1371/journal.pone.0113380

Turnbull, L., Toyofuku, M., Hynen, A. L., Kurosawa, M., Pessi, G., Petty, N. K., et al. (2016). Explosive cell lysis as a mechanism for the biogenesis of bacterial membrane vesicles and biofilms. *Nat. Commun.* 7:11220. doi: 10.1038/ncomms11220

Wachter, S., Bonazzi, M., Shifflett, K., Moses, A. S., Raghavan, R., and Minnick, M. F. (2019). A CsrA-binding, trans-acting sRNA of *Coxiella burnetii* is necessary for optimal intracellular growth and vacuole formation during early infection of host cells. *J. Bacteriol.* 201, e00524–19. doi: 10.1128/JB.00524-19

Ward, N., and Moreno-Hagelsieb, G. (2014). Quickly finding orthologs as reciprocal best hits with BLAT, LAST, and UBLAST: how much do we miss. *PLoS One* 9:e101850. doi: 10.1371/journal.pone.0101850

Waters, L. S., and Storz, G. (2009). Regulatory RNAs in bacteria. *Cell* 136, 615–628. doi: 10.1016/j.cell.2009.01.043

Westermann, A. J., Förstner, K. U., Amman, F., Barquist, L., Chao, Y., Schulte, L. N., et al. (2016). Dual RNA-seq unveils noncoding RNA functions in host-pathogen interactions. *Nature* 529, 496–501. doi: 10.1038/nature16547

Winston, W. M., Molodowitch, C., and Hunter, C. P. (2002). Systemic RNAi in *C. elegans* requires the putative transmembrane protein SID-1. *Science* 295, 2456–2459. doi: 10.1126/science.1068836

Winston, W. M., Sutherlin, M., Wright, A. J., Feinberg, E. H., and Hunter, C. P. (2007). *Caenorhabditis elegans* SID-2 is required for environmental RNA interference. *Proc. Natl. Acad. Sci. U.S.A.* 104, 10565–10570. doi: 10.1073/pnas.0611282104

Zhang, H., Zhang, Y., Song, Z., Li, R., Ruan, H., Liu, Q., et al. (2020). sncRNAs packaged by *Helicobacter* pylori outer membrane vesicles attenuate IL-8 secretion in human cells. *Int. J. Med. Microbiol.* 310:151356. doi: 10.1016/j.ijmm.2019.151356

Zhao, C., Zhou, Z., Zhang, T., Liu, F., Zhang, C. Y., Zen, K., et al. (2017). *Salmonella* small RNA fragment Sal-1 facilitates bacterial survival in infected cells via suppressing iNOS induction in a microRNA manner. *Sci. Rep.* 7:16979. doi: 10.1038/s41598-017-17205-4

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership