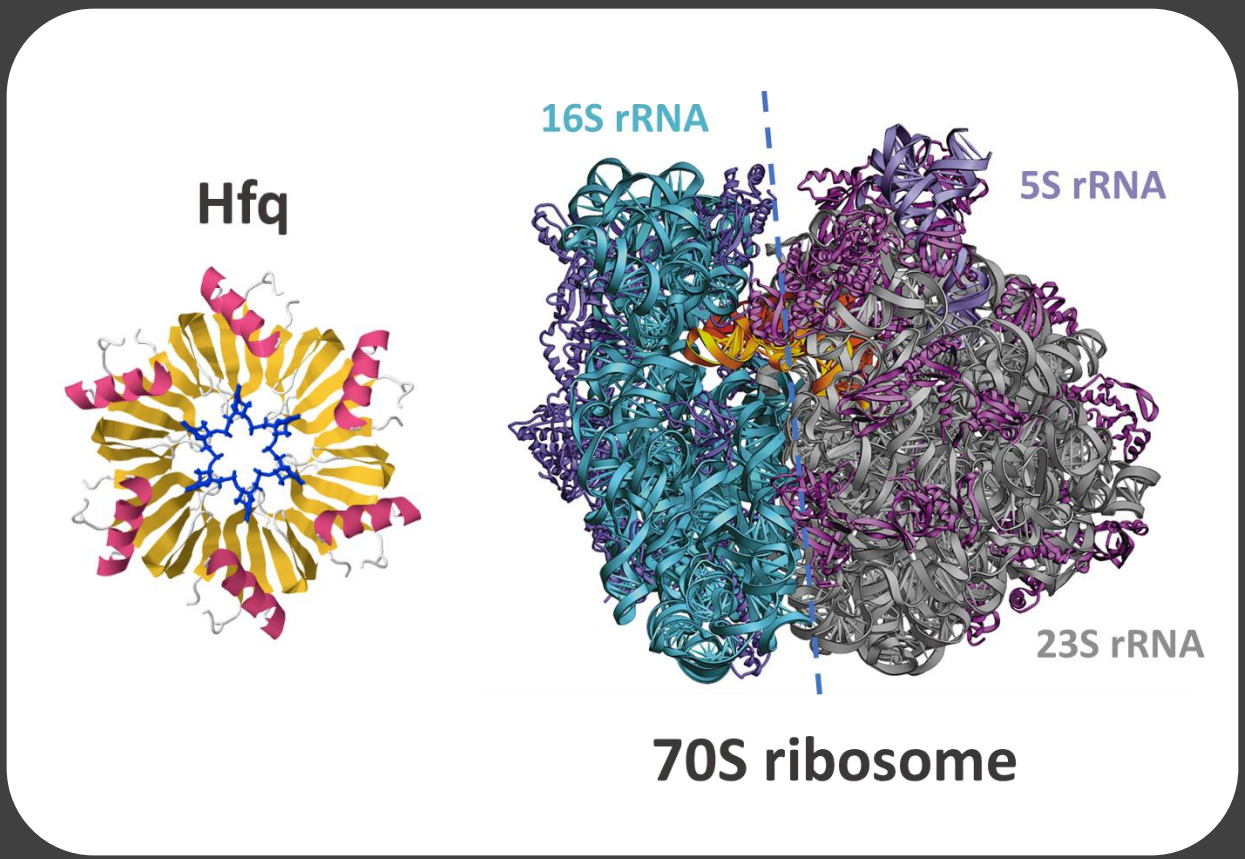
More than an RNA matchmaker: Expanding the roles of Hfq into ribosome biogenesis

Ricardo Filipe da Cruz Duarte dos Santos



Dissertation presented to obtain the Ph.D. degree in Biology

Instituto de Tecnologia Química e Biológica António Xavier | Universidade Nova de Lisboa

Oeiras,
February, 2019

UNIVERSIDADE NOVA DE LISBOA

More than an RNA matchmaker: Expanding the roles of Hfq into ribosome biogenesis

Ricardo Filipe da Cruz Duarte dos Santos

Dissertation presented to obtain the Ph.D. degree in Biology

Instituto de Tecnologia Química e Biológica António Xavier | Universidade Nova de Lisboa

Oeiras, February, 2019

The pictures used in the cover of this doctoral dissertation were adapted from biology.kenyon.edu (Hfq structure) & Yusupov *et al*, 2001 (Ribosome structure)

Supervisor:

Professora Doutora Cecília Maria Pais de Faria de Andrade Arraiano
Investigadora Coordenadora do Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa.

Co-supervisor:

Prof. Doutor José Eduardo Marques Andrade
Investigador FCT do Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa

President of the Jury:

Representante do Presidente do Conselho Científico
Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa.

Examiners:

Prof. Doutor Phillippe Régnier
Professor Catedrático, Université Paris Diderot (Paris 7/ Émérite), et Institut de Biologie Physico Chimique, Paris, France.

Professora Doutora Maria do Carmo Salazar Velez Roque da Fonseca
Professora Catedrática, Faculdade de Medicina da Universidade de Lisboa, e Presidente do Instituto de Medicina Molecular, Lisboa.

Professora Doutora Ana Sofia Bregieiro Eulálio
Investigadora Principal, Centro de Neurociências e Biologia Celular, Universidade de Coimbra.

Prof. Doutor Luís Jaime Gomes Ferreira da Silva Mota
Professor Auxiliar, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa.

Acknowledgments

Acknowledgments

I am thankful to all the people that have directly and indirectly contributed to this work. This work would not have been possible without the support of:

- **Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa (ITQB NOVA)** for providing the optimal work conditions and scientific environment for the execution of this dissertation.
- **Fundação para a Ciência e Tecnologia (FCT)** for the financial support of my PhD fellowship (PD/BD/105733/2014).
- My supervisor, **Dr. Cecília M Arraiano**, for supporting my scientific endeavors and the kindness demonstrated since the beginning.
- My co-supervisor, **Dr. José M. Andrade**, for mentoring me in all aspects and for friendship nurtured along the years. *Sensei, doumo arigatou gozaimasu*.
- **My lab mates**, that accompanied me through this journey making the everyday lab life simpler and funnier.
- **My friends and colleagues** of the MolBioS PhD program class of 2015, shaking brains.
- **My closest friends**, that always looked out for me and kept me (in)sane. A special thanks to Stephanie, Susana, Diana, Daniel, Anup e Inês.
- **My family**, for the unconditional love and support along every single leap I take. To my father's focus, my mother's care and my sister's resolve. You are a bit of me.
- **My beloved**, Sofia, thank you for igniting me. For moving me. For helping me. NFWMB
- **My little one**, David, thank you for the fire. For the smile. For the anchor.

Contents

Abstract

Abstract

Ribosome biogenesis is a complex process involving multiple factors. The work described here is primarily centered in the study of ribosomal RNA, highlighting its central role in translation regulation. We have uncovered new regulators involved in rRNA processing, folding and degradation pathways. For the first time, we demonstrate that the widely conserved RNA chaperone Hfq, mostly known as the sRNA-mRNA matchmaker, acts as a ribosomal assembly factor in *Escherichia coli*, affecting rRNA processing, ribosome levels, translation efficiency and accuracy. This function is suggested to be independent of its activity as sRNA-regulator. Furthermore, Hfq is found to interact with RNase R, a hydrolytic exoribonuclease. These two proteins cooperate not only in a novel RNA quality control pathway that eliminates superfluous rRNA fragments but also in rRNA maturation. Overall, we demonstrate that Hfq and RNase R are critical for ribosome levels and act in previously unrecognized pathways affecting translation. We extended the list of natural substrates of the widely conserved Hfq RNA chaperone, that now includes ribosomal RNA, the most abundant RNA class in the cell. We show that Hfq is a central regulator affecting different levels of ribosome biogenesis, that include not only the rRNA processing but also ribosome assembly. This work provides an additional explanation for the pleiotropic effects of Hfq deletion on bacterial physiology, besides its common role as a sRNA regulator. Moreover, it paves the way to understand the function of Hfq in many bacteria in which this RNA-binding protein is not a major player in sRNA biology. In addition, this work expanded our knowledge on the regulatory features of rRNA. rRNA has been neglected in most high throughput studies, usually considered a contaminant that needs to be removed from samples. We now show that rRNA may act as reservoir of sRNAs, a feature that seems to be conserved throughout life. Taking as a whole, this work reveals new functions for old and widely characterized members of the RNA network and offers a new perspective to control gene expression through modulation of RNA.

Resumo

Resumo

A biogénese de ribossomas é um processo complexo que envolve múltiplos factores. O trabalho aqui descrito centra-se principalmente no estudo do RNA ribossomal, sublinhando os eu papel central na regulação da tradução. Revelámos novos reguladores envolvidos no processamento, *folding* e metabolismo do rRNA. Demonstramos pela primeira vez que a chaperona de RNA Hfq, conhecida como mediadora de interacções entre pequenos RNAs e RNAs mensageiros, actua como um factor de montagem dos ribossomas em *Escherichia coli*, afectando não só o processamento do rRNA mas também os níveis de ribossomas. Os nossos resultados indicam que esta nova função parece ser independente da sua actividade como regulador de sRNAs. Mostramos ainda que Hfq interage com a RNase R, uma exoribonuclease hidrolítica. Estas duas proteínas cooperam não só num novo mecanismo de controlo da qualidade do RNA que elimina fragmentos de rRNA supérfluos mas também na maturação do rRNA. Colectivamente, demonstramos que Hfq e RNase R têm um papel crucial nos níveis de ribossomas e actuam numa via previamente desconhecida afectando a tradução. A lista de substratos naturais Hfq agora incluem que o rRNA, a classe de RNAs mais abundante na célula. Desta forma demonstramos que Hfq é um regulador central que afecta diferentes aspectos da biogénese de ribossomas, incluindo não só o processamento de rRNA mas também a montagem de ribossomas. Este trabalho fornece um explicação adicional para os efeitos pleiotrópicos da deleção da Hfq na fisiologia bacteriana, para além do seu papel como regulador de sRNAs. Adicionalmente, abre caminho para compreender a função da Hfq em várias bactérias nas quais esta proteína de ligação ao RNA não é um principal regulador na biologia de sRNAs. O trabalho aqui descrito expande o conhecimento actual das características regulatórias do rRNA. O rRNA tem vindo a ser negligenciado na maioria dos estudos *high throughput*, onde são considerados contaminantes que a ser eliminados das amostras. Mostramos agora que o rRNA pode actuar como um reservatório de sRNAs, uma característica que parece ser genericamente conservada. Globalmente este trabalho revela novas funções para proteínas previamente conhecidas e largamente caracterizadas que compõem a rede regulatória do RNA, oferecendo novas perspectivas para o controlo da expressão génica através da modulação pelo RNA.

List of publications

List of publications

The work presented in this Dissertation contributed for the following publications.

Andrade JM*, **dos Santos RF***, Chelysheva I, Ignatova Z & Arraiano CM (2018) The RNA-binding protein Hfq is important for ribosome biogenesis and affects translation fidelity. *EMBO J.* **37:** e97631. DOI: 10.15252/embj.201797631
*Joint first authors

This article was highlighted in the News and Views section of the EMBO Journal issue: Sharma IM, Korman A & Woodson SA (2018) The Hfq chaperone helps the ribosome mature. *EMBO J.* **37:** e99616

dos Santos RF, Quendera AP, Boavida S, Seixas AF, Arraiano CM & Andrade JM (2018) Major 3'–5' Exoribonucleases in the Metabolism of Coding and Non-coding RNA. In *Progress in Molecular Biology and Translational Science* **159**: 101–155. Academic Press. DOI: 10.1016/bs.pmbts.2018.07.005

dos Santos, RF, Andrade JM, Pissarra, J, Deutscher MP and Arraiano; Hfq and RNase R interact and cooperate in a novel rRNA quality control-pathway. (in final preparation)

Matos RG, Casinhas J, Bárria C, **dos Santos RF**, Silva IJ & Arraiano CM (2017) The Role of Ribonucleases and sRNAs in the Virulence of Foodborne Pathogens. *Front. Microbiol.* **8:** 910. DOI: 10.3389/fmicb.2017.00910

Domingues S, Moreira RN, Andrade JM, **dos Santos RF**, Bárria C, Viegas SC & Arraiano CM (2015) The role of RNase R in trans-translation and ribosomal quality control. *Biochimie* **114:** 113–8. DOI: 10.1016/j.biochi.2014.12.012

Thesis outline

Thesis outline

This dissertation is divided into five main chapters:

Chapter 1 is a general introduction which focuses mainly on the role of RNA-binding proteins, specifically on the three major 3'-5' exoribonucleases and the RNA chaperone Hfq, exploring their interconnection with ribosomal RNA maturation and degradation.

The results of this Doctoral work are presented in Chapters 2, 3 and 4. Each one of these chapters contains its own Abstract, Introduction, Results, Discussion, Materials and Methods, References and Supplementary information sections.

Chapter 2 describes the work that led to the discovery of a new function for the widely known RNA chaperone Hfq. It is shown that Hfq impacts the maturation status of the 16S rRNA and 30S subunit and when it is absent a decrease in the functional 70S pool, translation efficiency and translation accuracy.

Chapter 3 explores a novel interaction between Hfq and the exoribonuclease RNase R. Results demonstrate that both proteins cooperate in the removal of deleterious rRNA fragments which arise from initial endonucleolytic cleavages of the 16S and 23S rRNAs. Moreover, it is reported that neither one of the two other major 3'-5' exoribonucleases can compensate for the lack of Hfq and RNase R. Additionally, it is shown that the simultaneous deletion of Hfq and RNase R also impacts the processing of both 16S rRNA and 23S rRNA changing the ribosome pool.

Chapter 4 reports the existence of two rRNA fragments (rRF) that are not promptly degraded and are predicted to be functional. This work shows that both rRFs are present in exponentially growing cells and can bind a wide variety of possible targets. Specifically, we provide predictions of mRNAs targeted by these rRFs, as well as sRNAs whose action can be modulated by both rRFs simultaneously.

Chapter 5 is the final discussion that weights on the results of the previous three chapters, connecting the main results produced by this Dissertation.

Chapter 1

Introduction

This chapter was based on:

dos Santos RF, Quendera AP, Boavida S, Seixas AF, Arraiano CM & Andrade JM (2018) Major 3'–5' Exoribonucleases in the Metabolism of Coding and Non-coding RNA. In *Progress in Molecular Biology and Translational Science* **159**: 101–155. Academic Press. DOI: 10.1016/bs.pmbts.2018.07.005

Matos RG, Casinhas J, Bária C, **dos Santos RF**, Silva IJ & Arraiano CM (2017) The Role of Ribonucleases and sRNAs in the Virulence of Foodborne Pathogens. *Front. Microbiol.* **8:** 910. DOI: 10.3389/fmicb.2017.00910

Domingues S, Moreira RN, Andrade JM, **dos Santos RF**, Bárria C, Viegas SC & Arraiano CM (2015) The role of RNase R in trans-translation and ribosomal quality control. *Biochimie* **114:** 113–8. DOI: 10.1016/j.biochi.2014.12.012

1. Chapter 1: Introduction

1.1. General introduction

The ribosome is a macromolecular complex responsible for the orchestration of an essential cellular process – protein synthesis. It is an essential ribonucleoprotein complex necessarily present in all domains of life. Although the ribosome has been extensively studied, the research mainly focused on understanding the mechanism behind the process of translation (Wimberly *et al*, 2000; Yusupov *et al*, 2001; Schmeing & Ramakrishnan, 2009). Meaning that the biology underlying the biogenesis of the ribosome and its subunits was somewhat overlooked.

Ribosome production involves a series of intricate steps, like the synthesis, modification and processing of both the RNA and protein components, that are likely to occur simultaneously within the cell. Hence, the need for extra-ribosomal factors that aid in the assembly of all the parts that make up the ribosomal subunits. These are ribosome biogenesis factors that act to assist during various stages of the ribosome production pathway. Many of these factors are proteins known to act on a set of steps, forming an intricate network of ribosome biogenesis regulation (Kaczanowska & Rydén-Aulin, 2007; Shajani *et al*, 2011). Structurally, ribosomes are roughly two-thirds RNA and not surprisingly ribosome biogenesis factors are often RNA-binding proteins (Steitz & Moore, 2003). Nevertheless, given the complexity of ribosome production, our current knowledge about ribosome biogenesis and its regulators is still limited.

This introduction will focus on the bacterial ribosome biogenesis of the model organism *Escherichia coli*, highlighting the rRNA processing step. RNA-binding proteins with known or potential functions regarding rRNA processing or metabolism will also be addressed.

1.2. Ribosomes

Ribosomes are responsible for the translation of genetic information contained in the base sequence of messenger RNAs into the amino acid sequence of proteins. Protein synthesis is carried out in all organisms using the same molecular principles, meaning that all ribosomes share universal features. Namely, functional ribosomes consist of two asymmetrical subunits, one smaller and one larger, comprising in total more than fifty specialized ribosomal proteins that bind the ribosomal RNA molecules. These unequal ribonucleoprotein complexes have different functions: the small subunit harbor the decoding center and is responsible for initiating translation by binding the messenger RNA, whereas the larger subunit contains the peptidyl transferase center responsible for protein synthesis *per se*, accommodating the three tRNA binding sites (A – aminoacyl; P – peptidyl; E – exit) involved in the production of the nascent polypeptide chain (Nierhaus, 2006). The most well-known and extensively studied ribosomal particle is the bacterial ribosome from *Escherichia coli*. Its relative simplicity and the absence of more complex feedback functions that are present in other multicellular organisms makes it an ideal model for ribosome biogenesis studies. In fact, it has been used as a reference for ribosome-related research over the years.

In bacteria, the two asymmetrical subunits are referred to as the 30S and 50S subunits – small subunit (SSU) and large subunit (LSU), respectively. Each subunit is an intricately folded RNA molecule with a vast subset of specialized proteins attached. Specifically, the 30S subunit is formed by the 16S ribosomal RNA (rRNA) and 21 ribosomal proteins (r-proteins) whilst the 50S subunit comprises two rRNA molecules, 23S and 5S, with 33 additional r-proteins (Figure 1) (Kaczanowska & Rydén-Aulin, 2007; Shajani *et al*, 2011; Thurlow *et al*, 2016). Although each subunit exists independently, the functional ribosomal particle is required for protein synthesis to occur, and so both subunits assemble into a ~2.5 MDa supramolecular machine, the 70S bacterial ribosome (Williamson, 2003; Laursen *et al*, 2005).

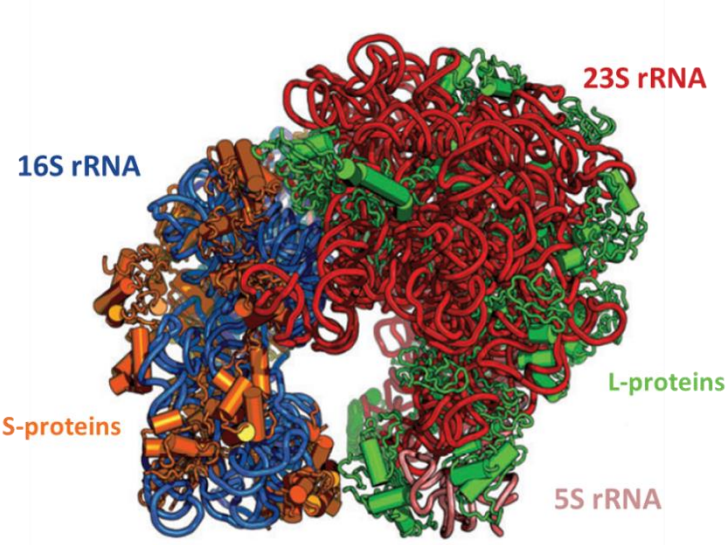


Figure 1 – Structural representation of the bacterial ribosome and its components (adapted from Shajani et al, 2011).

This large ribonucleoprotein complex carries out the translation process using messenger RNA as the template and aminoacyl-transfer RNAs as substrates. Bacterial translation can be subdivided into three steps: initiation, elongation and termination (Figure 2). Successful initiation of bacterial protein synthesis relies on the action of three initiation factors (IF1-3) as well as on the interaction between the Shine-Dalgarno sequence of the mRNA and the anti-Shine-Dalgarno region of the 16S rRNA (Laursen *et al*, 2005). The first is positioned 6-9 nucleotides before the initiation codon, whereas the is found near the 3' end of the 16S rRNA molecule. Hence, correct ribosome positioning is accomplished by the SD/anti-SD interactions that allow the action of the initiation factors in accommodating the fMet-tRNA^{fMet} to the P site of the 50S subunit (Schmeing & Ramakrishnan, 2009). The elongation step consists in the sequential addition of amino acids to the nascent polypeptide chain through mRNA decoding. Aminoacyl-tRNAs are brought to the A site by the elongation factor Tu (EF-Tu) which after GTP hydrolysis releases the aminoacyl end of the tRNA, allowing for peptidyl transfer to occur in the nascent chain (Kaczanowska & Rydén-Aulin, 2007). The decoding process ensures that only the correct tRNA binds the A site which is accomplished by codon/anti-codon interactions between mRNA and tRNAs as well as through the interaction of two 16S rRNA

residues (A1492 and A1493) with the minor groove of the mRNA/tRNA mini helix (Ogle & Ramakrishnan, 2005). Elongation factor G (EF-G) is required for the translocation of the tRNA from the A to the P site, while the deacylated tRNA moves from the P to the E site, to be released. This leaves the A site empty and ready for the next round of ribosome translocation. Finally, translation termination takes place when a stop codon enters the A site and is recognized by the release factors 1 or 2. RF1 recognizes UAA and UAG while RF2 recognizes UAA and UGA stop codons (Schmeing & Ramakrishnan, 2009). This triggers hydrolysis of the polypeptide from the peptidyl-tRNA, releasing it from the newly synthesized protein from the ribosome. RF3 promotes dissociation of RF1/RF2 from the A site while the ribosome recycling factor (RRF) and EF-G dissociate the subunits and clear the P site. IF3 enters the P site and stimulates the release of the mRNA, preparing the ribosome for the next initiation step (Laursen *et al*, 2005; Schmeing & Ramakrishnan, 2009).

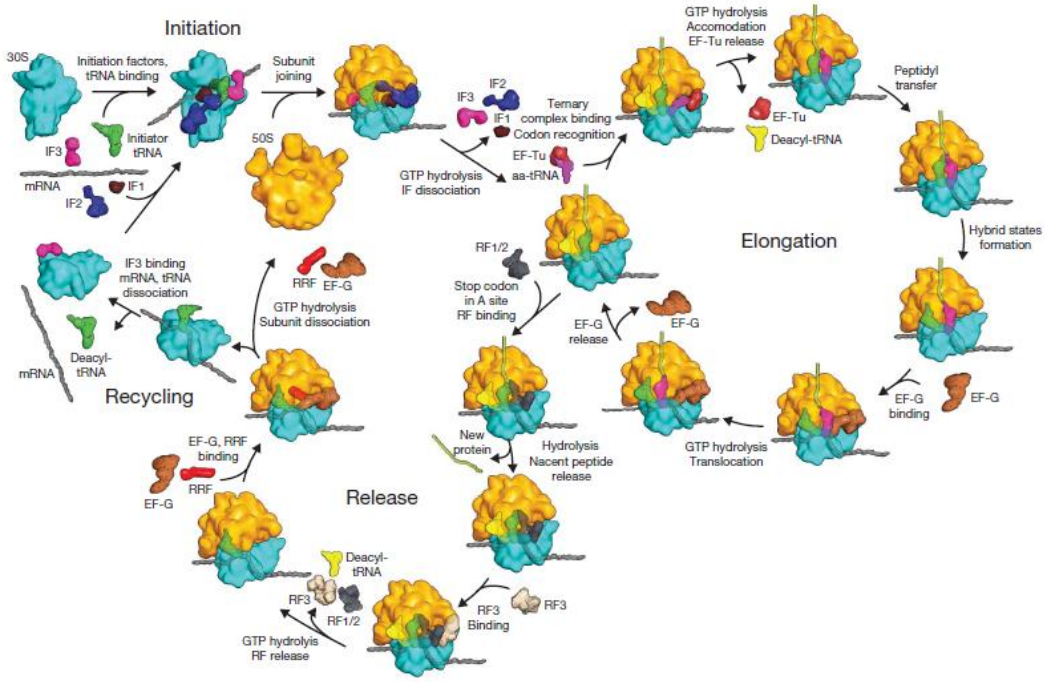


Figure 2 – Overview of the translation cycle (adapted from Schmeing & Ramakrishnan, 2009). aa-tRNA, aminoacyl-tRNA; EF elongation factor; IF, initiation factor; RF, release factor.

The growth rate of the bacterial cell is intimately related to its overall capacity for protein synthesis. Curiously, the ratio between actively translating and idle ribosomes is constant. Therefore, an increase in protein synthesis capacity is only achieved by an increase in the number of ribosomes. Accordingly, log phase cells exhibit ribosome levels of approximately ~70,000 particles, whether cells growing at slower rates reduce this number to as low as 2,000 (Neidhardt *et al*, 1990). In fact, ribosomes can account for more than half of the dry weight of fast-growing *E. coli* cells. Ribosome production is therefore highly demanding to the bacterial cell since it involves the biosynthesis of high quantities of large RNA molecules and numerous proteins. Accordingly, ribosome biogenesis is a multi-step complex process that consumes over one third of the total energy production in fast growing cells (Chen & Williamson, 2013). The close relation between protein synthesis and growth rate means that ribosome biogenesis is essential for cell fitness and must be highly coordinated with other cellular processes to keep homeostasis.

1.2.1. Ribosome biogenesis

The assembly of the bacterial ribosome is a thorough and well-coordinated biological process that because of its energy requirements, must occur as smoothly and efficiently as possible. Ribosome biogenesis comprises a series of events that lead to the correct formation of both subunits. These include transcription, processing and nucleotide modification of the rRNA as well as translation and modification of the r-proteins (Figure 3). Followed by the hierarchically binding of the latter to their corresponding rRNA. This confers the overall folding of each subunit and consolidates its structural characteristics essential for their function.

The genes encoding the rRNA are typically present in several copies across any given genome. In the case of *E. coli* the three rRNA genes are encoded in the same operon, which is repeated 6 times, in a total of 7 rDNA copies (*rrnA-E*, *rrnG* and *rrnH* operons). Transcription from these operons during exponential phase accounts for over 50% of the cell's RNA synthesis (Condon *et al*, 1995a). Ribosomal rRNAs are transcribed by the RNA polymerase complex (RNAP) as one long primary transcript containing all three rRNA genes. RNAP is a multi-subunit complex composed of four subunits ($\alpha_2\beta\beta'$) plus the

associated sigma factors (σ) that recognize promoter sequences. The common order of rRNA genes spaced by internal transcribed sequences (ITS) is as follows: 16S – ITS-1 – (either $tRNA^{Ala}$ and $tRNA^{Ile}$ or $tRNA^{Glu}$) – ITS-2 – 23S – ITS-3 – 5S. The *rrn* operons possess two tandemly arranged promoters (P1 and P2) that are both recognized by σ^{70}, the sigma factor involved in the transcription of house-keeping and growth-related genes (Condon *et al*, 1995b). Many regulatory layers are in place to ensure constitutively high levels of transcription of the *rrn* operons, including the presence of UP elements recognized by the RNAP α subunit, FIS-mediated activation and the presence of antiterminator *nut*-like sequences present in the external and internal transcribed sequences (ETS and ITS, respectively) (Condon *et al*, 1992; Gyorfy *et al*, 2015). However, transcription from rDNA operons must also be negatively regulated in order to ensure a correct modulation according to cellular growth and stress response. For example, when an uncharged tRNA enters the A site, the ribosome-associated RelA protein responds by synthetizing (p)ppGpp, two small nucleotides that act as signaling molecules (Wilson & Nierhaus, 2007). This initiates the stringent response due to amino acid deprivation which triggers a ser of pleiotropic responses, including the activation of protein degradation, amino acid synthesis and, most importantly, the immediate shutdown of rRNA and tRNA synthesis (Neidhardt *et al*, 1990). Moreover, rDNA operons in *E. coli* are located in different parts of the genome and its operons have functional differences (Condon *et al*, 1992), which raises the possibility of a differential regulation of rRNA synthesis. There are only small variations in the endogenous sequence of the rRNA genes. However, this heterogeneity was shown to be important in the response to nutrient limitation. Upregulation of synthesis from the *rrnH* operon during this stress triggers a modulation of ribosome function. This operon encodes a *rrsH* (16S rRNA) baring ten sequence variations that allow *rrsH*-containing ribosomes have the ability to alter the expression of a set of specific stress-response genes (Kurylo *et al*, 2018). Hence, ribosome heterogeneity adds yet another layer of ribosome regulation that impacts cell physiology (Byrgazov *et al*, 2013).

The genes encoding ribosomal proteins need to be transcribed, and these mRNA later translated. In order to maintain correct levels of ribosomal components, expression of the 19 operons containing r-proteins is subject to regulation by the level of rRNA

transcription. Ten r-proteins were found to bind not only to rRNA but also to their own mRNA, leading to an autogenous translation repression. When in excess, S1, S4, S7, S8, S15, S20, L1, L4, L10, and L20 negatively regulate their transcripts. Moreover, operons containing r-proteins tend to cluster other important ribosome-related proteins, like elongation factors, or even RNAP core components (Neidhardt *et al*, 1990; Nomura, 1999; Kaczanowska & Rydén-Aulin, 2007; Shajani *et al*, 2011). Additionally, a set of r-proteins – including L7/L12 or S12 – are post-translationally modified, for example, by the addition of methyl groups or incorporation of a methylthio-aspartic acid residues (Arnold & Reilly, 1999). Although the purpose of r-protein modifications is not clear these can serve to alter r-protein binding efficiency or to optimize the binding of extra-ribosomal factors and translation ligands. Nevertheless, this adds to the complexity of the biosynthesis of the ribosomal components.

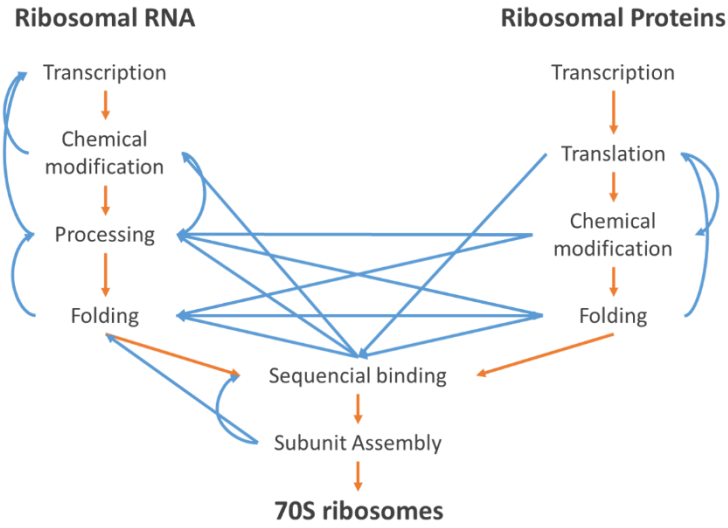


Figure 3 – Overview of the multiple steps of ribosome biogenesis in *E. coli*. Orange arrows are a simplistic representation of the flow of maturation, while blue arrows represent the co-occurrence of the multiple steps within the cell.

Similarly, the 16S and 23S rRNA molecules that mainly constitute the ribosomal subunits are also subjected to several modifications. The 16S rRNA contains 11 modifications while the 23S comprises 25 modifications. These modifications include the isomerization of uridine to pseudouridine as well as the addition of methyl or carbonyl

groups (Del Campo & Ofengand, 2004). Although many of the modified ribonucleotides are highly conserved, the function behind each modification is still elusive. However, since some modifications are added to naked rRNAs while others are added later after r-protein binding, the first may be a way to fine-tune rRNA folding while the latter may enable maturation checkpoints to maintain a correct structural conformation. Strikingly, some modifications and initial r-protein binding occurs even before the transcription of the ribosomal DNA (rDNA) is complete (Figure 3).

1.2.2. Ribosomal RNA processing

Ribosomal RNA processing is an essential step of ribosome biogenesis that carves the mature rRNA sequences out of the primary transcript and respective rRNA precursors (Figure 4). Misprocessed rRNA molecules cause severe defects during ribosome production, ultimately leading to structural defects and dysfunctional ribosomal particles (Connolly & Culver, 2009; Shajani *et al*, 2011; Leong *et al*, 2013). The enzymes responsible for this processing are called ribonucleases (RNases) as they catalyze the nucleolytic cleavage of RNA molecules. Nevertheless, other non-RNase proteins are known to be required for the correct maturation of rRNAs. Although the molecular mechanism behind this requirement is largely unknown, these proteins like RNA helicases, RNA chaperones or modification enzymes can bind the pre-ribosomal particle facilitating rRNA refolding and r-protein binding (Sharpe Elles *et al*, 2009; Sharma *et al*, 2018). These steps can therefore lead to the exposure of the RNA structure determinants necessary for RNase recruitment and action (Wilson & Nierhaus, 2007; Shajani *et al*, 2011).

The processing of rRNA primary transcript starts with the endoribonuclease – an RNase that cleaves RNA internally – RNase III, releasing the precursors rRNA molecules which contain in themselves the sequence for the mature rRNA as well as additional extra nucleotides at each end (Deutscher, 2009). Thereby, the precursor for 16S rRNA is termed 17S rRNA and contains an extra 115 nts at the 5’ end and 33 nts at the 3’ end. The 5’ end is matured firstly by RNase E that removes 49 nts, followed by RNase G that removes the remaining 66 nts (Li *et al*, 1999b). Several RNases can participate in the 3’ end maturation of the 17S rRNA, namely 4 exoribonucleases – RNases that cleave RNA from one end to

the opposite direction – RNase R, RNase II, PNPase and RNase PH and one recently discovered endoribonuclease, YbeY (Davies *et al*, 2010; Jacob *et al*, 2013; Sulthana & Deutscher, 2013). Similarly, the precursor for 23S rRNA, the pre-23S rRNA, contains 3 to 7 extra nts at the 5' end and 7 to 9 nts at the 3'end. Processing of both ends is dependent on RNase III cleavages, but RNase G also participates in the 5' end maturation, while RNase T along with RNase PH participate in the 3' end trimming (Li *et al*, 1999a; Song *et al*, 2011; Gutgsell & Jain, 2012). Finally, the precursor for 5S rRNA, the 9S rRNA, has 84 additional nts at the 5' end and 42 at the 3' end. RNase E is responsible for processing at both ends leaving only 3 additional nts at each end. RNase T finishes the trimming at the 3' end while the enzyme responsible for the final 5' end maturation is still unknown (Roy *et al*, 1983; Li & Deutscher, 1995).

Processing of rRNA molecules is a key step during ribosome biogenesis, since the presence of the extra flanking sequences can cause kinetic traps to arise slowing or even hindering the correct rRNA folding with detrimental consequences to the cell (Woodson, 2008; Connolly & Culver, 2009; Clatterbuck Soper *et al*, 2013). For example, the presence of the 5' end extra nucleotide sequence of the 17S rRNA precursor was shown to form alternative structures with the first nucleotides of helix 1, leading to incorrect folding of this helix and central-pseudoknot ultimately affecting translation fidelity (Brink *et al*, 1993; Roy-Chaudhuri *et al*, 2010). This is substantiated by the fact that 30S subunits containing the 17S rRNA precursor are rendered inactive by the basepairing between the extra nucleotides at each end. In fact, there is a quality control mechanism in place that senses 70S ribosomes that contain immature 30S subunits, targeting them for degradation mediated the ribonucleases YbeY and RNase R (Jacob *et al*, 2013; Warner, 2013). Moreover, the structure of mature 30S subunits with 16S rRNA incorporated clearly presents both 5' and 3' ends apart from each other. In this way, the processing of 17S rRNA to mature 16S rRNA within 30S subunits triggers their activation. This final processing step ensures that only correctly assembled and matured 30S subunits are active and can therefore engage in translation.

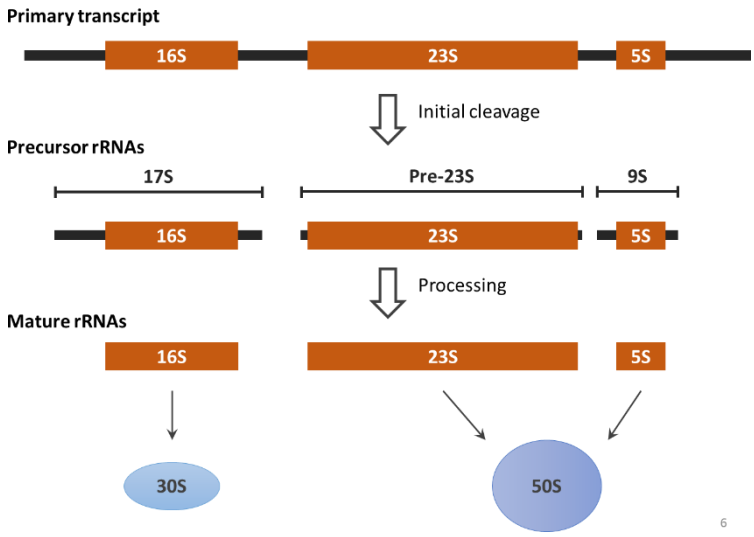


Figure 4 – Depiction of rRNA processing steps that give raise to rRNA precursors, matured forms and corresponding subunits.

1.2.3. Ribosome biogenesis factors

In order to meet the cellular requirements for active ribosomes the series of steps that lead to the formation of ribosomes must be fine-tuned to high efficiency. Notably, all the 57 components that make up the ribosome – except for tetrameric L7/L12 – are required in one copy per ribosomal particle, which raises considerable implications for the biogenesis process (Wimberly *et al*, 2000; Yusupov *et al*, 2001; Schuwirth *et al*, 2005; Yonath 2009).

Remarkably, each ribosomal subunit can be reconstituted *in vitro* using rRNA molecules and purified r-proteins. This implies that all the information needed for the correct formation of the ribosome is comprised within its structural components. These seminal studies leaded by Nomura and colleagues on the assembly of the 30S subunit as well as the ones leaded by Nierhaus and colleagues on the 50S subunit assembly, provided invaluable insights into how r-proteins bind and shape the ribosome (Traub & Nomura, 1968; Mizushima & Nomura, 1970; Nierhaus & Dohme, 1974). Altogether, this early research recognized that that some r-proteins directly bind naked rRNAs, while others depend either of specific rRNA sites that are only made available by binding of the first r-proteins or of the interaction with other r-proteins to be recruited to the pre-ribosomal

particle (Williamson, 2003; Wilson & Nierhaus, 2007; Shajani *et al*, 2011). An assembly map that represents the interdependencies of r-proteins for their incorporation in each of the subunits summarizes a highly hierarchy and cooperative process. Furthermore, reconstitution intermediate particles were also detected, which needed a heat activation step to proceed with the binding to the next subset of r-proteins. This means that the formation of the mature subunit *in vivo* must be facilitated by non-ribosomal molecules that lower the energy of the activation steps (Davis & Williamson, 2017).

Additionally, biosynthesis and assembly of all constituents must be carefully coordinated to ensure cell fitness and avoid superfluous energy costs. Accordingly, we find an extensive and complex network of regulatory circuits controlling the multiple stages of ribosome biogenesis (Connolly & Culver, 2009; Shajani *et al*, 2011; Davis & Williamson, 2017). Although the ribosome and its components possess an intrinsic self-assembly character, cells evolved ways to proficiently accelerate ribosome production by maintaining a "toolbelt" of proteins that assist during the various assembly steps. These proteins are extra-ribosomal biogenesis factors, which are critical for a smooth and accurate production of ribosomes (Table 1).

The main component of ribosomes is the long and highly structured molecules of RNA. Therefore, during folding of such extensive RNAs there is a higher probability of alternative secondary structures to occur, many of which are inactive and possible very stable (Woodson, 2008; Sharma *et al*, 2018). Transitioning from such kinetic traps into the native form is a slow process, since interactions must be broken and remodeled. The action of ribosome biogenesis factors allows for this remodeling to occur rapidly and therefore promote the maturation of the ribosomal subunits. Various factors have already been identified, however, our knowledge about them and their molecular mechanisms is still scarce. The list of known factors includes GTPases, RNA helicases and protein chaperones (Table 1) (Charollais *et al*, 2004; Sharma *et al*, 2005; Jiang *et al*, 2006).

Table 1 – Examples of *E. coli* proteins implicated in ribosome assembly (adapted from Gibbs & Fredrick, 2018).

Biogenesis factor	Type	Ribosomal subunit
RbfA	RNP binding	30S
RimJ	RNP binding	30S
RimM	RNP binding	30S
RimP	RNP binding	30S
YhbY	RNP binding	50S
KsgA (RsmA)	Modification enzyme	30S
RsmC	Modification enzyme	30S
RlmA (RrmA)	Modification enzyme	50S
RlmE (RrmJ)	Modification enzyme	50S
RluB	Modification enzyme	50S
RluC	Modification enzyme	50S
RluD	Modification enzyme	50S
DeaD (RhlD)	Helicase	50S
DbpA (RhlC)	Helicase	50S
SrmB (RhlA)	Helicase	50S
RhlE	Helicase	50S
DnaK/DnaJ/GrpE	Protein chaperone	30S, 50S
GroES/GroEL	Protein chaperone	50S
Era	GTPase	30S
RsgA (YjeQ)	GTPase	30S
LepA (EF4)	GTPase	30S
Der (EngA)	GTPase	50S
YihA (EngB)	GTPase	50S
BipA (TypA)	GTPase	50S

Many of these were found to suppress phenotypes arising after rRNA mutations or r-protein inactivations. For example, the ribosome-binding factor A (RbfA) was found to suppress the cold-sensitive phenotype of C23U mutation in the 16S rRNA (Dammel & Noller, 1995). The C23 residue is part of the helix 1, just after the central pseudoknot, located near the 5' of the mature 16S rRNA molecule. In fact, RbfA was found to facilitate the refolding of the 16S leader sequence as well as the conformational changes necessary for the formation of the central pseudoknot (Dammel & Noller, 1995). This ribosome biogenesis factor is essential for growth at low temperatures, where rRNA folding and

ribosome production is slower (Xia *et al*, 2003). Moreover, RbfA was found to associate with free 30S subunits and not with 70S ribosomes, meaning that it dissociates from the subunit after proper maturation is achieved (Dammel & Noller, 1995; Shajani *et al*, 2011). Cells that that lack RbfA exhibit 17S rRNA precursor accumulation and aberrant ribosome profiles, both hallmarks of ribosome biogenesis defects (Jones & Inouye, 1996).

The ribosome maturation factor M (RimM) is another ribosome biogenesis factor involved in the maturation of the 30S subunit. Similarly to Δ*rbfA* strains, RimM depletion leads to accumulation of 17S rRNA and free subunits accompanied by a reduction in the polysome fraction (Bylund *et al*, 1997; Leong *et al*, 2013). However, RimM acts on the opposite side of the 30S subunit, specifically on the 3' domain, through interaction with r-protein S19 facilitating. RimM is thought to facilitate the binding of S19 and S13 to helices 31 and 33b, respectively, as well as the binding of other 3' domain r-proteins (Guo *et al*, 2013). RimM influences the acetylation status of S18 and assists the fold of helices 33 and 43 (Clatterbuck Soper *et al*, 2013). Notably, overexpression of RbfA partially suppresses the slow growth phenotype of Δ*rimM* strains, but 16S maturation is only marginally increased (Bylund *et al*, 1998). This is one example of the interconnections and cooperation between ribosome biogenesis factors, that is also observed among other 30S and 50S maturation factors. Structural analysis of the pre-30S particles isolated from Δ*rbfA* and Δ*rimM* strains revealed common features between them, including unprocessed 5' and 3' ends, unfolding of the central pseudoknot and undocked helix 44 (which harbors the decoding center) (Clatterbuck Soper *et al*, 2013). Hence, RbfA and RimM act on the 5' and 3' domains, respectively, acting later in the ribosome biogenesis cascade to assist maturation of critical 30S features (Table 1) (Thurlow *et al*, 2016).

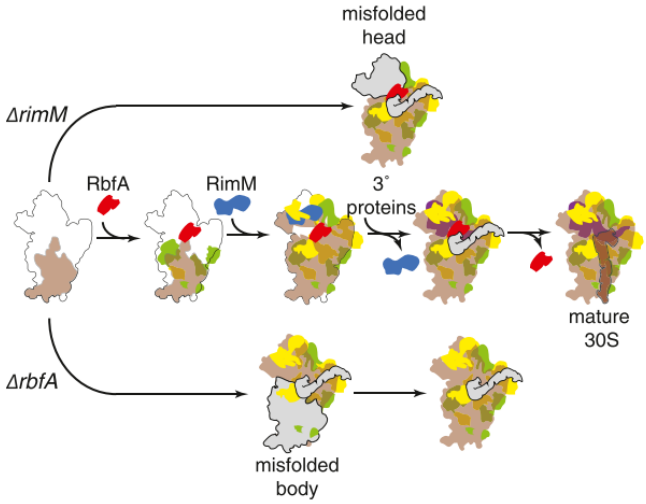


Figure 5 – Overview of a model for RbfA and RimM action on 30S maturation (adapted from Clatterbuck Soper *et al*, 2013). Correctly folded rRNA is in brown and solid gray represents nonnative structures. RimM (blue) facilitates proper folding of the 16S 3' domain during assembly of the head (top row), while RbfA (red) acts on the 5' domain assembly and promotes formation of the central pseudoknot (bottom row).

The cold-shock DEAD-box protein A (CsdA or DeaD) is an RNA helicase involved in 50S subunit assembly that was found to be essential for growth at low temperatures (Charollais *et al*, 2004). RNA helicases are thought to rescue the extensive 23S rRNA molecule from conformational traps. This is substantiated by the fact that *in vitro* reconstitution of 50S subunits requires heat activation to allow conformational rearragments (Nierhaus & Dohme, 1974; Herold & Nierhaus, 1987). Moreover, analysis of the r-protein composition of the pre-50S intermediate that accumulates in an Δ*csdA* mutant reveals that it may act later on the assembly of the 50S subunit, as many of the necessary r-proteins are already bound (Peil *et al*, 2008). Overexpression of other DEAD-box helicases, like RhlE were shown to suppress the growth defect of a CsdA deletion mutant (Awano *et al*, 2007), which further highlights the importance of parallel pathways for correct maturation of the ribosomal particles.

Overall, we find that ribosome biogenesis is a complex landscape with numerous factors involved, many of which are interdependable. Ranging from transcription initiation regulation, through rRNA processing by ribonucleases and into the maturation

of ribosomal subunits by ribosome biogenesis factors, all must be carefully orchestrated. Cooperation between all factors and enzymes involved is absolutely critical to achieve a highly coordinated biological process capable of making this supramolecular machine. In this sense, ribonucleases play a critical role in ribosome biogenesis as they catalyze RNA processing from rRNA precursors to their mature functional forms (Deutscher, 2009). Multiple 3'-5' exoribonucleases have been implicated in the processing of the 16S rRNA (Sulthana & Deutscher, 2013). As is often the case in other RNA biology processes, these processive enzymes tend to overlap functions, being able to compensate the lack of each other which denotes an obvious advantage in maintaining an accurate ribosome production (Deutscher, 2009; Arraiano *et al*, 2010; Andrade *et al*, 2009b; dos Santos *et al*, 2018). Moreover, 3'-5' exoribonucleases are important for rRNA quality control mechanisms that enables ribosome turnover, elimination of aberrant ribosomes and recycling of nucleotides during stress conditions (Cheng & Deutscher, 2003; Basturea *et al*, 2011; Jacob *et al*, 2013).

1.3. Exoribonucleases

The level at which any RNA is expressed is determined by its rate of transcription and its rate of degradation. Cells are constantly producing new RNAs while discarding old or defective transcripts in order to regulate gene expression. Coding RNA/messenger RNA (mRNA) decay is essential because of its role in destruction of defective mRNAs producing aberrant proteins, recycling of ribonucleotides, and adaptation to changes in the environment. Moreover, RNA transcripts usually have to be processed and/or modified during their maturation. The structure and chemically reactive nature of RNA make this molecule more unstable than DNA. However, RNAs are not all degraded at the same rate or in the same manner. Consequently, cells must control precisely the various enzymatic activities involved in RNA degradation and processing, to avoid elimination of important RNA molecules. The enzymes that cleave and trim newly made transcripts to mature them, and rapidly destroy unnecessary cellular RNAs are called ribonucleases. These enzymes are described as endoribonucleases if they cleave RNA molecules internally, or exoribonucleases if they attack RNAs from one of their extremities.

In prokaryotes, RNA decay is primarily initiated by endonucleolytic cleavages, followed by the action of exoribonucleases that rapidly eliminate the intermediary products (Carpousis *et al*, 2009; Mackie, 2013). However, exoribonucleases can also act on primary transcripts (Andrade *et al*, 2009b). Gram-negative bacteria do not encode 5'-3' exoribonucleases, which seem to be less abundant than 3'-5' exoribonucleases (Durand & Condon, 2018). Bacterial RNases have been most extensively studied in *Escherichia coli* that encodes about twenty RNases, of which eight are 3'-5' exoribonucleases (Mohanty & Kushner, 2018). In *E. coli*, most of the degradation processes that required this type of ribonucleases are mainly carried out by three: PNPase, RNase II and RNase R (Arraiano *et al*, 2010). PNPase (encoded by *pnp*) is a 3'-5' phosphorolytic enzyme member of PDX family of enzymes. The RNase II (*rnb*) and RNase R (*rnr*) are paralogue proteins that belong to the RNase II family of enzymes that hydrolytically degrade RNA in the 3' to 5' direction in a processive manner (Zuo & Deutscher, 2001). These enzymes have homologues in all domains of life. PNPase differs from RNase II and RNase R not only in the mode of action but also in its structure. The homotrimeric PNPase is assembled into a ring-like structure with a central channel, composed by six RNase PH-like domains (Symmons *et al*, 2000). PNPase homologues are found in eukaryotic organelles, with a similar domain organization and structure. Moreover, this overall organization is similar to those of archaeal and human multiprotein exosome core complexes, indicating evolutionary links in structure and function between PNPase and exosomes. The RNase II/RNase R enzymes are characterized by the presence of a well-conserved RNB domain which has a unique $\alpha\beta$-fold (Frazão *et al*, 2006). The RNB catalytic domain was named after the *E. coli rnb* gene which encodes ribonuclease B, widely known as ribonuclease II (RNase II). RNase II from *E. coli* is the prototype of this family that also includes bacterial RNase R and eukaryotic Rrp44/Dis3 proteins. Some organisms like *Saccharomyces cerevisiae* may present only one member, while others may present up to three paralogues. In Bacteria, the γ division *Proteobacteria* often contain both RNase II and RNase R homologues; however, outside of this lineage, the single family member present more closely resembles RNase R than RNase II (Zuo & Deutscher, 2001). In Archaea and Eukarya, classification of the homologues members follows the generic designation of RNase II-like

or RNase II/RNase R enzymes. These enzymes are also important in several stress responses and have been associated to pathogenesis (Andrade *et al*, 2009b).

PNPase, RNase II and RNase R may display some functional redundancy, as they share common substrates. However, they do not show the same degree of ability to compensate each other's activity. Cells must express PNPase together with at least one of the hydrolytic enzymes RNase II or RNase R, as double mutants *pnp rnb* or *pnp rnr* are lethal (Donovan & Kushner, 1986; Cheng *et al*, 1998). However, the double *rnb rnr* strain is viable. This highlights the importance of PNPase in cell biology. These 3'-5' enzymes are active against a wide range of substrates, including coding (mRNA) and non-coding RNAs (ribosomal RNA/rRNA, transfer RNA/tRNA, small RNA/sRNA). PNPase and RNase II are mainly involved in mRNA decay but also participate in tRNA and rRNA maturation and degradation (Donovan & Kushner, 1986; Li & Deutscher, 1994; Cheng & Deutscher, 2003). PNPase is the main enzyme affecting small non-coding RNAs stability (Andrade & Arraiano, 2008; De Lay & Gottesman, 2011; Andrade *et al*, 2012). RNase R is more important in quality control mechanisms that assure the degradation of aberrant rRNA and tRNA molecules, though it has also been related to mRNA decay (Li *et al*, 2002; Andrade *et al*, 2006; Cheng & Deutscher, 2003; Andrade *et al*, 2009b). A common substrate requirement for all the major 3'-5' exoribonucleases is the need for a linear overhang in the 3' end of the RNA molecule to promote binding of the enzyme. Hence, the degradation of RNA molecules by PNPase, RNase II and RNase R is promoted by polyadenylation, the addition of poly(A) stretches at the 3' end of the RNA (Régnier & Hajnsdorf, 2013).

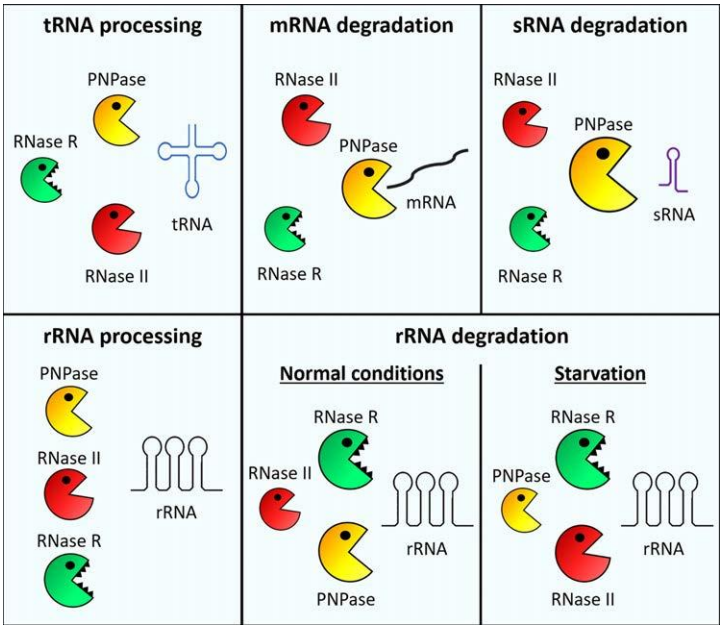


Figure 6 – Substrate specificities of each one of the three major 3'-5' exoribonucleases in *E. coli* (adapted from dos Santos et al, 2018). RNases more active against a specific substrate are represented closer to that substrate.

1.3.1. PNPase

Polynucleotide phosphorylase (PNPase) is one of the most remarkable enzymes involved in RNA metabolism. This enzyme is unique amongst exoribonucleases as it presents dual activity: it works either as a 3'-5' exoribonuclease in RNA degradation or as a polymerase in RNA synthesis, depending on the inorganic phosphate (Pi) concentration available (Mohanty & Kushner, 2000). The seminal work by Marianne Grunberg-Manago led to the discover of PNPase as the first enzyme able to synthesize RNA *in vitro* (Grunberg-Manago *et al*, 1955). This activity was essential for elucidation of the genetic code, leading to Severo Ochoa's Nobel prize in Physiology or Medicine in 1959. Nonetheless, today it is widely recognized that its main activity in cells is RNA degradation. In recent years, PNPase has become a key enzyme in the regulation of non-coding RNAs across species. PNPase is widely conserved from bacteria to plants and metazoans (Zuo & Deutscher, 2001; Bermúdez-Cruz *et al*, 2005). Studies on this enzyme have contributed enormously, and still do, to a better understanding of RNA genetics and its regulation.

1.3.1.1. PNPase function and regulation

PNPase is a member of the PDX family, along with RNase PH found in bacteria, and the core of the exosome found in archaea and eukaryotes (Zuo & Deutscher, 2001). The main activity of PNPase in cells is the degradation of RNA. However, depending on inorganic phosphate concentration, this enzyme can act as a synthetic enzyme leading to polymerization of single-stranded RNA (ssRNA) from nucleoside diphosphates (dNDPs). When acting on RNA, Mg^{2+} is an essential cofactor (Andrade *et al*, 2009b). Another reaction less explored is its ability to exchange the β-phosphate group of dNDPs and free orthophosphate (Grunberg-Manago, 1963). Nevertheless, PNPase is also involved in single-stranded DNA (ssDNA) degradation and DNA repair pathways, reactions that are mostly dependent on the presence of Mn^{2+} ions (Cardenas *et al*, 2009). Therefore, PNPase is a multifunctional protein active against a wide range of substrates (Figure 6).

PNPase catalyzes the processive 3'-5' phosphorolytic degradation of RNA, releasing nucleoside diphosphates in the presence of high inorganic phosphate concentrations. To be able to bind to an RNA substrate, the RNA molecule must have a minimal 3' overhang of 7 to 10 unpaired nucleotides (nts) (Py *et al*, 1996; Cheng & Deutscher, 2005). A wide-genome analysis showed that PNPase is the 3'-5' exoribonuclease that significantly affects more transcripts in *E. coli* (Pobre & Arraiano, 2015). Deletion of PNPase resulted in the upregulation of 59% of transcripts (compared to only 29% and 41% in deletion strains of RNase II and RNase R, respectively). Moreover, many of these transcripts were stable RNAs (tRNAs, rRNAs and sRNAs). In fact, PNPase seems to be particularly relevant for stability control of non-coding RNAs (Andrade & Arraiano, 2008; Andrade *et al*, 2012, 2013). PNPase is involved not only in the degradation but also in the processing and maturation of a variety of substrates PNPase acts in a processive way and is sequence independent, though its action on folded RNAs can be stimulated by 3' polyadenylation (Xu & Cohen, 1995; Py *et al*, 1996; Spickler & Mackie, 2000). However, the processive degradation of RNA by PNPase is blocked by double-stranded RNA (dsRNA) structures (Spickler & Mackie, 2000). To overcome such physical

blocks on RNA, PNPase can associate with other proteins in a variety of multiprotein complexes.

At low inorganic phosphate (Pi) concentrations, however, PNPase catalyzes the polymerization of ssRNA from dNDPs in a template-independent manner (Godefroy, 1970; Sulewski *et al*, 1989). Firstly identified *in vitro*, this reaction was further observed to occur *in vivo* as PNPase was shown to add polyribonucleotide tails in *E. coli* strains lacking PAP I, the major polyadenylating enzyme (Mohanty & Kushner, 2000). Contrarily to the homopolymeric adenosine tails added by PAP I, the PNPase-biosynthesized tails are heteropolymeric containing all four nts (Slomovic *et al*, 2008). Rho-dependent transcription terminators were suggested to be modified by the polymerase activity of PNPase, while Rho-independent transcription terminators act as polyadenylation signals for PAP I (Mohanty & Kushner, 2006). Although PAP I is the major polyadenylating enzyme in *E. coli*, responsible for modification of more than 90% of the transcripts in exponentially growing bacteria, PNPase seems to be the main polyadenylating enzyme in *Streptomyces coelicolor*, spinach chloroplasts and cyanobacteria (Yehudai-Resheff *et al*, 2001; Rott *et al*, 2003; Sohlberg *et al*, 2003; Jones & Mackie, 2013). The archaeal exosome, structurally very similar to PNPase, has also been demonstrated to be responsible for the addition of heteropolymeric tails in the archaeon *Sulfolobus solfataricus* (Portnoy *et al*, 2005). Hence, rounds of addition of heteropolymeric tails and exoribonucleolytic digestion by PNPase can promote the degradation of RNA in different organisms.

The RNA chaperone Hfq is a mediator of sRNA-mRNA interactions that has been shown to display other relevant functions in the cell. Hfq was also found to regulate PNPase activity. The biosynthetic activity of PNPase in the *hfq* single mutant is enhanced and it becomes the primary polynucleotide polymerase, adding heteropolymeric tails almost exclusively to 3' truncated mRNAs (Mohanty *et al*, 2004). Previous observations have hinted at a relationship between PNPase activity and cellular metabolism. For instance, PNPase can be regulated by nucleotides; the binding of ATP to PNPase results in inhibition of both the phosphorolytic and the polymeric activities of this exoribonuclease (Del Favero *et al*, 2008). On the other hand, c-di-GMP activates PNPase activity in a reaction that is dependent of oxygen and controlled by its association with DosC and DosP

direct oxygen sensor proteins (Tuckerman *et al*, 2011). Earlier crystallographic studies of *E. coli* PNPase revealed the presence of citrate from the crystallization buffer at the active site (Nurmohamed *et al*, 2009). When the putative role of citrate in PNPase activity was further explored, it was observed that its enzymatic activity was inhibited in the presence of magnesium-chelated citrate (Nurmohamed *et al*, 2011). Interestingly, human PNPase (hPNPase) regulates RNA import to the mitochondria where the Krebs cycle takes place and citrate is present at low concentrations (Wang *et al*, 2010). Overall, data show that PNPase activity is dependent on the metabolic state of the cell. Curiously, hPNPase and the archaeal exosome complex from *Sulfolobus solfataricus* activities are also inhibited by the presence of citrate, suggesting a conserved link between RNA degrading activity and cellular metabolism throughout evolution (Stone *et al*, 2017).

PNPase is encoded by the *pnp* gene which is transcribed by two promoters (Portier & Regnier, 1984). PNPase expression is regulated at the post-transcriptional level by many factors. Two of the most important and the first to be identified were PNPase itself and RNase III (Robert-Le Meur & Portier, 1992). A stem-loop structure found upstream of the RBS sequence in the 5'UTR makes the *pnp* mRNA very stable. In normal conditions, RNase III cleaves at this double-stranded region, originating an RNA helix with a protruding 3' end that is then preferentially removed by PNPase. The processed transcript is then rapidly degraded (Robert-Le Meur & Portier, 1992, 1994). PNPase is thus autogenous regulated and the stability of the *pnp* mRNA is inversely correlated with the amount of active PNPase in the cell (Jarrige *et al*, 2001). Later, it was shown that PNPase can act as a translational repressor of its own expression independently of RNase III cleavage (Carzaniga *et al*, 2015). In the absence of RNase III, PNPase can inhibit its own translation possibly by competing for *pnp* primary transcript with the ribosomal protein S1. Polyadenylation can indirectly affect PNPase autoregulation and thus PNPase levels. The higher the levels of polyadenylated transcripts are, the more likely PNPase binds to them instead of the 5' end of its own transcripts (Mohanty & Kushner, 2002). Interestingly, RNase II also regulates PNPase expression since in the absence of RNase II, there is an overexpression of PNPase (Zilhão *et al*, 1996a).

The RNA-binding protein CsrA was also found to repress PNPase translation (Park *et al*, 2015). CsrA protein is a component of the Csr system that is widely involved in regulation of virulence, motility, biofilm formation, quorum sensing, carbon metabolism and stringent response (Vakulskas *et al*, 2015). After *pnp* 5' UTR stem-loop is processed by RNase III and PNPase, CsrA binds to two sites of the transcript, where one CsrA site overlaps the *pnp* Shine Dalgarno sequence, preventing ribosome binding and leading to the repression of *pnp* translation (Park *et al*, 2015). CsrA-mediated repression of PNPase was observed throughout growth, but it was greater during stationary phase. Being part of the Csr system, CsrA is negatively regulated by sRNAs CsrB and CsrC, which suggests that alterations in CsrB/C levels would translate into alterations in PNPase levels.

PNPase levels are induced in cells exposed to a temperature downshift. At optimal temperature, PNPase is hardly fundamental to *E. coli*, unless either RNase II or RNase R are missing (Donovan & Kushner, 1986; Cheng *et al*, 1998). Interestingly, a strain lacking both RNase II and RNase R is viable, meaning that PNPase can replace their activity in those conditions (Andrade *et al*, 2006). Nonetheless, at low temperatures, PNPase is indispensable for *E. coli* growth (Piazza *et al*, 1996; Zangrossi *et al*, 2000), and has been identified as a cold shock protein (Jones *et al*, 1987). During the cold acclimation phase there is a 10-fold increase in *pnp* transcripts (Zangrossi *et al*, 2000) and a 2-fold increase in PNPase protein level (Mathy *et al*, 2001). This increase in transcripts appears to be due mainly to their stabilization (Zangrossi *et al*, 2000; Mathy *et al*, 2001). However, RNase III processing of *pnp* transcripts is not affected by the temperature downshift. Since the *pnp* mRNA is not efficiently translated during cold acclimation, the stabilization of the transcripts appears to compensate for this poor translation to maintain or increase the protein levels (Zangrossi *et al*, 2000). After adaptation to cold, *pnp* mRNA levels decrease most probably due to RNase R activity (Zhang *et al*, 2018).

Not only proteins are regulators of PNPase expression. *In vitro* assays have showed that the small non-coding SraG RNA (located upstream of the *pnp* gene) binds with *pnp* mRNA and induces new sites that can be recognized and cleaved by RNase III within the *pnp* mRNA region base paired with SraG (Fontaine *et al*, 2016). Furthermore, the SraG-*pnp* mRNA basepairing prevents the *pnp* translational initiation *in vitro* because

it blocks the 30S subunit docking onto the RBS sequence of the transcript. Moreover, SraG-dependent regulation of PNPase expression was found to occur either in the presence or absence of RNase III (Fontaine *et al*, 2016). It is proposed that such regulation can be relevant after induction of the stringent response, which results in the 2-fold increase in the levels of the regulatory non-coding RNA SraG.

1.3.1.2. PNPase complexes

PNPase is sensitive to RNA secondary structures and the processive degradation of RNA by PNPase is blocked by double-stranded RNA regions (Spickler & Mackie, 2000). PNPase can associate with different proteins which form multiprotein complexes helpful to overcome this physical challenge. The most well-known PNPase-based complex is the degradosome, a multiprotein complex that is mainly composed by the endoribonuclease RNase E, which provides the scaffold for the complex, the DEAD-box RhlB helicase and the glycolytic enzyme enolase (Miczak *et al*, 1996; Py *et al*, 1996; Vanzo *et al*, 1998). The degradosome composition is dynamic and varies in response to environmental changes. For instance, at cold temperatures the helicase RhlB is replaced by the helicase CsdA (Prud'homme-Généreux *et al*, 2004).

The degradosome complex (or variations of it) is found in many different organisms. In the Gram-negative *Caulobacter crescentus*, an RNA degradosome complex was also identified (Hardwick *et al*, 2011). It is composed by PNPase, RNase E as the scaffold, a DEAD-box helicase, and, surprisingly, the Krebs cycle enzyme aconitase. This is another proof of the connection between RNA degradation and cellular metabolism, as mentioned above. Based on bacterial two-hybrid system and copurification assays, an RNA degradosome has also been proposed to exist in the Gram-positive *B. subtilis.* In the absence of RNase E, it is RNase Y that provides the scaffold for the complex, which includes PNPase, RNase J1, the DEAD-box helicase CshA and two glycolytic enzymes, enolase and phosphofructokinase (Lehnik-Habrink *et al*, 2010; Newman *et al*, 2012; Salvo *et al*, 2016). Nevertheless, *B. subtilis* degradosome partners have not yet been purified in complex. Furthermore, only few components colocalized: RNase Y is located at the membrane, RNase J1/J2 and CshA are close to the RNA bulk in the cytoplasm, while

PNPase and the glycolytic enzymes are uniformly dispersed in the cytoplasm (Cascante-Estepa *et al*, 2016). These results do not exclude completely the interactions between the components but suggest that these interactions could be only transitory.

Like in *B. subtilis*, the components of the RNA degradosome are the same in *Staphylococcus aureus*. RNase Y interacts with the RNA helicase CshA and enolase, although there is no confirmed interaction between PNPase and RNase Y (Roux *et al*, 2011). Recently, it has been suggested that RnpA, a subunit of RNase P, could also have a role in this degradosome (Wang *et al*, 2017). There is competition between PNPase and enolase for RnpA, though PNPase catalytic activity was unaltered when bound to RnpA, and enolase activity was slightly upregulated (Wang *et al*, 2017). On the other hand, in *Staphylococcus epidermis* it was observed that PNPase forms a complex with RNase J1 and RNase J2 (Raj *et al*, 2018). These interactions have not changed either catalytic activities of PNPase, which could mean that they could be more important to increase degradation of structured RNA in certain environmental or cellular contexts.

PNPase is able to associate with RhlB independently of the degradosome, and this complex was found to be functional for the *in vitro* degradation of dsRNAs (Lin & Lin-Chao, 2005; Liou *et al*, 2002). This PNPase-RhlB interaction was later observed to be important for cysteine regulation *in vivo*, since its disruption led to the increase of intracellular content of cysteine and enhanced anti-oxidative resistance (Tseng *et al*, 2015). Additionally, PNPase can associate with Hfq and/or PAP I. Such complexes are important for the poly(A)-dependent decay of RNAs. Moreover, as previously stated, Hfq stimulates PNPase polymerase activity (Mohanty *et al*, 2004).

1.3.1.3. PNPase structure

Pure PNPase was isolated from *E. coli* as a homotrimer of 78 kDa subunits (Portier, 1975). The X-ray crystal structures of *E. coli*, *Streptomyces antibioticus* and *C. crescentus* PNPase reveal a homotrimeric subunit organization with a ring-like architecture (Figure 7) (Symmons *et al*, 2000; Shi *et al*, 2008; Nurmohamed *et al*, 2009; Hardwick *et al*, 2012). Each monomer exhibits a five-domain arrangement: at the N-terminus, two RNase PH domains (first and second core domains) are linked by an α-helical domain (Symmons *et*

al, 2002); at the C-terminal end, two RNA-binding domains named KH and S1 (Mattaj, 1993). In the quaternary structure, the KH and S1 domains are found together in one face of the trimer, while the active site is in the opposite side. The three subunits associate via trimerization interfaces of the core domains, forming a central channel where catalysis occurs. The existence of conserved basic residues in the neck region, along with a proper constriction in the channel, are pivotal in capturing RNA for processive degradation (Shi *et al*, 2008). Two constriction points have been identified in the channel, and the structure of PNPase in complex with RNA clearly indicates that the pathway followed by the RNA molecule is along the central pore in the direction of the active site (Symmons *et al*, 2000; Shi *et al*, 2008; Nurmohamed *et al*, 2009). The dynamic translocation of RNA by the enzyme depends on the conformational changes which occur in the opening at the central channel and its neighboring regions (Nurmohamed *et al*, 2009). Kinetic studies have established that the presence of distinct and separated RNA-binding sites causes, in part, the processivity of PNPase. The trimer channel probably contributes both to processivity and to the regulation of PNPase activity by RNA structural elements (Symmons *et al*, 2000).

The first core RNase PH domain has a conserved FFRR loop interacting with RNA, approximately 20 Å away from the putative catalytic site in the second core (Carzaniga *et al*, 2014). Two *E. coli* PNPase mutants in the FFRR loop (R79A and R80A) showed a great increase in K_m for ADP/Pi binding. Moreover, the first core domain was suggested to be involved in the guanosine pentaphosphate (pppGpp) synthetase activity of *S. antibioticus* PNPase, absent in the majority of PNPases. The second core RNase PH domain binds tungstate, a phosphate analogue, which suggests that the catalytic center is located here. This region may also have been separately adapted as a second active site in PNPases in general (Symmons *et al*, 2000). Most of the conserved residues are clustered in the second core domain (which contains the putative catalytic center) and in a small part of the first core domain. Only a few are in the KH or S1 RNA-binding domains. Mutations in amino acids located around the tungstate binding site abolish or severely decrease all catalytic activities of the enzyme, suggesting that these mutations affect the catalytic site directly (Jarrige *et al*, 2002). *E. coli* PNPase crystals obtained in the presence of Mn^{2+} showed that

the co-factor is coordinated by the conserved residues D486, D492 and K494 (Nurmohamed *et al*, 2009). Moreover, mutation of D492 eliminates both phosphorolytic and polymeric activities (Jarrige *et al*, 2002).

The α-helical domains, which connect the first and second core domains, have been implicated in the catalytic activity of *E. coli* PNPase (Briani *et al*, 2007), though they are the least conserved domains (Jarrige *et al*, 2002; Bermúdez-Cruz *et al*, 2005). In *E. coli* PNPase, mutating adjacent residues the α-helical domain seems to affect the catalytic activity (Briani *et al*, 2007). Analysis of the *S. antibioticus* structure demonstrated that helix α1 faces the putative phosphate binding site and could thus be a component of the catalytic center (Symmons *et al*, 2000; Briani *et al*, 2007). These results along with studies in spinach chloroplast and human PNPases (Yehudai-Resheff *et al*, 2003; Sarkar *et al*, 2005) suggest that the catalytic site of *E. coli* PNPase is probably composed of structural elements in both the first and second core domains.

On the other hand, the RNA-binding domains have limited effects on phosphorolysis, as PNPase mutants lacking either domains are catalytically active. RNA electrophoretic mobility shift assays showed that both KH and S1 domains are required for a proper binding (Matus-Ortega *et al*, 2007) and for PNPase autoregulation (Wong *et al*, 2013). Interestingly, PNPase truncated in these two domains can still bind RNA, though with weaker affinities, which confirms that the catalytic core has intrinsic RNA-binding activity (Stickney *et al*, 2005). Also, these truncated enzymes are usually stuck with their substrate, and the number of processed molecules is very reduced. A two-step model was proposed to explain this, based on the indirect facilitation of PNPase activity by both KH and S1 domains, allowing substrate binding and product release. The model proposes that there are two RNA binding surfaces on each PNPase monomer: one in the core catalytic domain and another formed by KH and S1 domains. First, it is suggested that ssRNA molecules interact weakly with one of the tandem KH-S1 domains. Then this bound ssRNA interacts strongly with the catalytic site, where it is subjected to phosphorolysis, until a stem-loop appears and the enzyme stalls because the substrate remains bound to the core RNA binding surface. Finally, another molecule could bind to other unoccupied

tandem KH-S1 domain, migrate to the core RNA binding site and allow the displacement of the stalled molecule (Stickney *et al*, 2005).

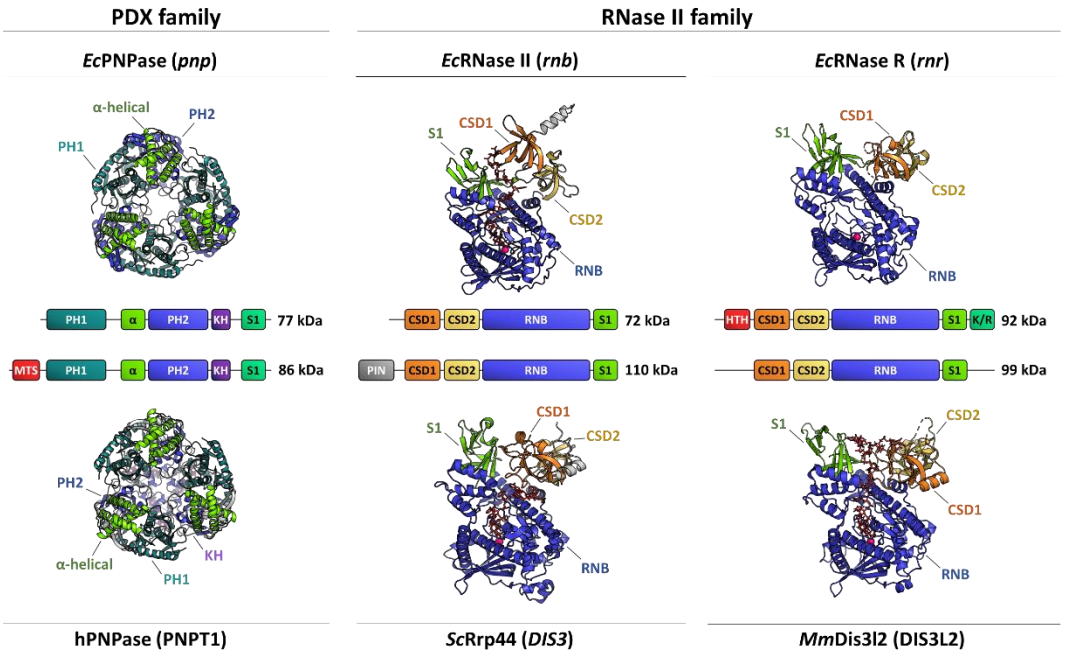


Figure 7 – Schematic representation of the domain organization and crystal structures of representative members of the PDX and RNase II protein families. The name of each protein along with the corresponding gene between brackets is shown next to each crystal structure. The domain organization followed by the approximate molecular weight of each prokaryotic and eukaryotic member (top and bottom panels, respectively) is aligned and displayed on top of each other for easy comparison. The specific domains present in each 3D crystal structure are also depicted. The Mg^{2+} cofactor is represented as a sphere within the catalytic centers, when applicable. Organisms and PDB identification codes are the following: EcPNPase, *Escherichia coli* PNPase (PDBID: 3GCM); hPNPase, *Homo sapiens* PNPase (PDBID: 3U1K); EcRNase II, *Escherichia coli* RNase II (PDBID: 2IX1); ScRrp44, *Saccharomyces cerevisiae* Rrp44 (PDBID: 2VNU); EcRNase R, *Escherichia coli* RNase R (PDBID: 5XGU); MmDis3L2, *Mus musculus* Dis3L2 (PDBID: 4PMW).

1.3.1.4. PNPase substrates

Whole-genome transcriptomic studies identified that PNPase as the exoribonuclease that most significantly affects *E. coli* transcripts (Mohanty & Kushner, 2003; Bernstein *et al*, 2004; Pobre & Arraiano, 2015). In a recent RNA-seq study, a total of 226 differentially expressed transcripts were reported in the deletion pnp mutant strain (Pobre & Arraiano, 2015). Not surprisingly, PNPase was found to affect many different pathways in the cell. However, this study highlighted the importance of PNPase in

contributing to cell motility in *E. coli*. The pnp mutant was significantly less motile and formed less biofilms than the parental strain, although the genes under PNPase regulation responsible for these phenotypes were not totally clear (Pobre & Arraiano, 2015). Nonetheless, these findings in *E. coli* are in agreement with reports in *B. subtilis*. Inactivation of PNPase in *B. subtilis* also triggered a deficient swarming motility phenotype (Liu *et al*, 2014, 2016). This was shown to be dependent on the PNPase-mediated degradation of the slrA mRNA. SlrA protein controls the levels of fla/che operon and, indirectly, sigD regulon, controlling the major determinants of cell motility (Liu *et al*, 2016). Hence, regulation of RNA turnover is a major determinant of motility gene expression and help explain PNPase impact on the virulence of different pathogenic bacteria.

PNPase is key in a surveillance mechanism that removes oxidized RNA. The exposure of cells to oxidative stress results in the formation of an oxidized form of guanine, the 8-oxoguanine, which is particularly harmful as it can be used as substrate by RNA polymerase. PNPase can discriminate 8-oxoguanine-containing RNA from normal RNA; it binds specifically to 8-oxoguanine-containing RNAs and removes these faulty molecules from the translational machinery (Hayakawa *et al*, 2001). As consequence, PNPase was found to protect both *E. coli* and HeLa cells from oxidative stress (Khidr *et al*, 2008; Wu *et al*, 2009; Tseng *et al*, 2015). This constitutes another example that PNPase contributes to a high fidelity of translation.

In addition to its role in mRNA turnover, PNPase is involved not only in the degradation but also in the processing of a variety of stable RNAs (tRNAs, rRNAs and sRNAs) in different organisms. In *E. coli*, all tRNAs are encoded with extra nucleotides at both 5' and 3' ends. Many of these transcripts terminate in a Rho-independent way, generating a stem-loop at the 3' end. PNPase is involved in the maturation and repair of the tRNAs 3'-terminus (Reuven *et al*, 1997; Maes *et al*, 2012; Mohanty *et al*, 2016). In the case of tRNA(Leu5), the Rho-independent terminator of its primary transcript *leuX* is removed by PNPase (Mohanty & Kushner, 2010). Moreover, elimination of defective tRNA precursors and tRNA degradation are other two important tasks of PNPase (Li *et al*, 2002).

PNPase, along with RNase R, is further essential for rRNA quality control, being responsible for the elimination of RNA molecules that appear after the incorrect maturation of 16S and 23S rRNAs (Cheng & Deutscher, 2003). Additionally, PNPase is implicated in 3'-end maturation of 16S rRNA (Sulthana & Deutscher, 2013). Recently, it was verified that PNPase can also process the 3'-end of purified pre-small subunits of 16S rRNA *in vitro* (Smith *et al*, 2018). During stationary phase, in *Deinococcus radiodurans*, rRNA decay is performed by RYPER, a ribonucleic complex formed by PNPase, the RNA-binding protein Rsr (a Ro autoantigen ortholog) and the small noncoding Y RNA (Wurtmann & Wolin, 2010; Chen *et al*, 2013). Y RNA acts as the scaffold and blocks the KH/S1 domains of PNPase, reducing its ability to interact with ssRNA, thus making the more structured rRNAs the newly preferable targets. Interestingly, a similar complex associating PNPase, Rsr and a ncRNA was also found in *S.* Typhimurium, an evolutionary distant species of *D. radiodurans;* however, its RNA degradation properties are still elusive (Chen *et al*, 2013). Also, putative Y RNAs were identified in more than 250 bacteria and phages encoding a Ro ortholog (Chen *et al*, 2014), likely indicating their potential to build RYPER-like complexes for rRNA degradation.

PNPase seems to be particularly relevant for stability control of non-coding RNAs across many distinct species. PNPase-dependent degradation of sRNAs was initially detected in *Salmonella* Typhimurium and *E. coli* (Viegas *et al*, 2007; Andrade & Arraiano, 2008). The chromosomal encoded MicA and RybB sRNAs that regulate outer membrane proteins are destabilized by PNPase, particularly in the stationary phase (Andrade & Arraiano, 2008). In strong contrast, RNase II and RNase R were not showed to play any role on this regulation. Also, it was shown that PNPase-mediated activity against sRNA do not necessarily require association within the degradosome. Even more, in the absence of its target *ompA*, the small RNA MicA is still degraded by PNPase (Andrade & Arraiano, 2008). In particular, it was found that sRNAs free of Hfq binding are preferably degraded by PNPase (Andrade *et al*, 2012, 2013). In bacteria, the RNA chaperone Hfq stabilizes sRNAs by protecting them from the attack of ribonucleases (Vogel & Luisi, 2011). Surprisingly, inactivation of Hfq and PNPase results in the strong increase of MicA, GlmY, RyhB and SgrS sRNAs levels (Andrade *et al*, 2012). This high affinity for non-coding RNAs

is conserved and the hPNPase was also found to selectively and preferentially degrade a microRNA in human melanoma cells (Das *et al*, 2010). PNPase is also responsible for the maturation of the CRISPR non-coding RNA RliB in *Listeria monocytogenes* (Sesto *et al*, 2014). Without PNPase activity, RliB CRISPR is not correctly processed and is not functional. Therefore, PNPase contributes to the DNA interference activity of RliB CRISPR. This work highlighted a new and unexpected function for PNPase in the processing of CRISPR non-coding RNAs. It has also been demonstrated that PNPase can also stabilize some sRNAs. A *pnp* mutant in *E. coli* presented destabilized sRNAs CyaR, RyhB and MicA (De Lay & Gottesman, 2011; Cameron & De Lay, 2016). It was proposed that PNPase probably protects Hfq-bound sRNAs by limiting the access of RNase E to them. PNPase was shown to protect sRNAs from degradation in the presence of Hfq (Andrade *et al*, 2012; Bandyra *et al*, 2016). *In vitro* studies showed that the enzyme degrades sRNAs in the absence of Hfq, but binds and is unable to degrade sRNAs in its presence, forming a tertiary complex with Hfq and sRNAs (Bandyra *et al*, 2016).

1.3.2. RNase II

RNase II is the prototype of the widespread RNB family of enzymes. Most interestingly, it is common to find different RNase II paralogues in a same organism (Zuo & Deutscher, 2001). As examples, two RNase II-like enzymes (RNase II and RNase R) are found in *E. coli* and three (hDis3, hDis3L and hDis3L2) can be found in human cells. The different RNase II homologues, although similar in domain organization, may exhibit different substrate specificities, expression profiles or even function in different cell compartments. Nevertheless, all RNase II-like enzymes display high activity against single-stranded RNA. RNase II is responsible for most of the hydrolytic activity on polyadenylated RNA in *E. coli* crude extracts and RNase II-like enzymes are responsible for the catalytic activity in archaeal and eukaryotic exosomes. However, unlike the other major 3'-5' exoribonucleases RNase R and PNPase, RNase II is stalled by RNA secondary structures. Strikingly, this trait contributes to an apparent paradox: RNase II can in fact protect several RNAs from degradation (Marujo *et al*, 2000). This happens because RNase II efficiently removes any single-stranded region downstream an RNA stem-loop, impairing then the

accessibility of other exoribonucleases to those substrates. Undoubtedly, RNase II is an important player in RNA metabolism, affecting not only RNA stability but also RNA maturation. Consequently, RNase II-like enzymes are involved in a multitude of regulatory pathways including those controlling several diseases.

1.3.2.1. RNase II function and regulation

RNase II is a 3'-5' exoribonuclease that degrades mRNA through a hydrolytic mechanism, in a processive manner yielding 5'-nucleoside monophosphates. It can disrupt short double-duplexes, but its activity is generally blocked by stem-loop structures (Cannistraro & Kennell, 1999). The preferable substrate is the homopolymer poly(A) and the target-transcripts must be longer than 10-15 nucleotides (nts), otherwise RNase II degrades mRNA distributively (Frazão *et al*, 2006). The end-product released has 4 nts, which are posteriorly destroyed by Oligoribonuclease (Andrade *et al*, 2009b). It is a highly active enzyme in the degradation of single-stranded RNA molecules, being responsible for 90% of the hydrolytic activity on poly(A) RNA observed in cell extracts (Figure 6) (Deutscher & Reuven, 1991). Strikingly, a major role for RNase II in cells seems to be the protection of mRNA (Marujo *et al*, 2000; Mohanty & Kushner, 2000). The single-stranded stretches present at the 3' end of RNAs are rapidly degraded by RNase II due to its high affinity to these substrates; as consequence, it impairs the access of other exoribonucleases since there is no accessible 3' single-stranded RNA sequence to which they could bind to initiate degradation (Mohanty & Kushner, 2003).

The RNase II family of enzymes is composed by a large number of homologue 3'-5' exoribonucleases found across all domains of life (Zuo & Deutscher, 2001). Organisms may contain only one or more members of the family, which include the close related bacterial RNase II and RNase R and eukaryotic Rrp44/Dis3 proteins. Due to the high homology between RNase II and RNase R, eukaryotic members of the same family are generally referred to as RNase II-like or RNase II/RNase R homologues.

The Rrp44/Dis3 is a yeast RNase II enzyme. It is a 110 kDa protein, that shares 19% homology in its sequence with the *E. coli* RNase II and is capable of degrading single-stranded RNA substrates endo- and exonucleolytically in a processive way (Dziembowski

et al, 2007). On the other hand, humans have three RNase II-like enzyme paralogues: hDis3, hDis3L1 and hDis3L2. hDis3 and hDis3L1 (also known as hDis3L) have different cellular localizations: hDis3, which is enriched in the nucleus and hDis3L, which is exclusively cytoplasmic (Tomecki *et al*, 2010; Staals *et al*, 2010; Łabno *et al*, 2016). These two proteins possess an additional PIN-domain in the N-terminus, a feature that is absent in the third eukaryotic homologue Dis3L2 (Łabno *et al*, 2016).

Acetylation was found to modify RNase II activity. The Lysine residue K501 of RNase II can be acetylated, leading to a decrease in substrate binding of RNase II. This residue is close to the catalytic center and its acetylation partially blocks the RNA channel affecting the enzyme activity (Song *et al*, 2016). Curiously, the acetylation status of RNase II decreases its enzymatic activity but not its stability. Acetylation of K501 is a reversible modification that is post-translationally controlled by Pka and CobB, an acetyltransferase and a deacetylase, respectively. High levels of RNase II acetylation are observed during slow growing cells.

RNase R is also regulated by acetylation, but only at the stability and not at the activity level (Liang *et al*, 2011). However, it is not known if Archaea and Eukarya members could be under a similar regulation.

RNase II is a monomer of 72 kDa and 644 amino acids (aa), encoded by the *rnb* gene from *E. coli*. The *rnb* gene is transcribed by two promotors, P1 and P2, and terminates in a Rho-independent terminator 10 nts downstream of the *rnb* stop codon (Zilhão *et al*, 1996b). Unless PNPase is absent, RNase II is not an essential enzyme in *E. coli* (Zilhão *et al*, 1996a), but no other RNase can substitute for RNase II to maintain cell survival during nutritional starvation (Sulthana *et al*, 2017). RNase II expression is controlled at the transcriptional and post-transcriptional levels. PNPase cleaves and degrades *rnb* mRNA, modulating its expression (Zilhão *et al*, 1996a). The endonucleases RNase III and RNase E perform an indirect control of RNase II levels. In a mutant strain for RNase III, the *pnp* mRNA is not properly cleaved, affecting the levels of PNPase and, therefore, the levels of RNase II. In a strain mutant for RNase E, there is an increase in both levels of *rnb* mRNA and RNase II itself (Zilhão *et al*, 1995). Additionally, it was found that the protein Gmr, which possesses a PAS domain that functions as an environmental

sensor, controls the stability of RNase II. It was observed that RNase II is twice more stable in the absence of *gmr* when compared to the wild-type (Cairrão *et al*, 2001). However, the molecular mechanism underlying this regulation remains mostly elusive.

1.3.2.2. RNase II complexes

In eukaryotes, a multiprotein complex termed exosome is composed of 9-11 protein subunits that cleave RNA molecules, one nucleotide at a time, in a 3' to 5' direction. The exosome is responsible mainly for mRNA decay but also in the processing of small RNAs like snoRNA and snRNAs and in rRNAs (Mitchell *et al*, 1997; Liu *et al*, 2006; Lorentzen *et al*, 2008). The human exosome core is composed of six proteins (hRrp41, hRrp42, hRrp43, hRrp45, hRrp46 and hMtr3) that form a hexameric ring similar in structure to PNPase. Above this hexameric ring are three proteins that contain KH and S1 RNA binding domains (hRrp4, hRrp40 and Csl4). Remarkably, all these enzymes are catalytically inactive (Liu *et al*, 2006). The catalytic active subunit is in fact an RNase II-like enzyme, either Dis3 or Dis3L1 (depending if it is the nuclear exosome or the cytoplasmic exosome).

As stated above, bacteria can have a multiprotein complex involved in RNA degradation, the degradosome. PNPase is usually the exoribonuclease found in such complex. Similar to what was observed in eukaryotic exosomes, in which an RNase II-like enzyme was responsible for the catalytic activity, it was suggested that bacterial RNase II might associate with the RNA degradosome (Lu & Taghbalout, 2014). However, this interaction only occurs after binding of PNPase and RhlB to the C-terminal domain of RNase E, which then results on a conformational change that allows RNase II to access the degradosome scaffold. Despite this finding, it remains to be elucidated the impact of such association. It is still not clear if their activities in the degradosome are coordinated or exclusive (Lu & Taghbalout, 2014).

1.3.2.3. RNase II structure

The elemental domain organization composed by RNA-binding domains (two cold shock domains (CDS) in the N-terminus and S1 domain in the C-terminus) flanking the RNB

catalytic domain can be found in all members of the RNase II family (Zuo *et al*, 2006; Frazão *et al*, 2006). However, additional domains may exist which confer different biochemical properties to its members. For example, unlike *E. coli* RNase II, the eukaryotic Rrp44/Dis3 homologue contains a PIN domain in the N-terminus that has endonuclease activity (Figure 7) (Lebreton *et al*, 2008; Schaeffer *et al*, 2009; Schneider *et al*, 2009). Even though bacterial RNase II and RNase R share a high sequence homology, the presence of additional domains specifically found in RNase R, confer to this enzyme the intrinsic capability to degrade structured RNAs that RNase II lacks.

The catalytic region of all RNase II-like enzymes is the RNB domain and it can stretch of about 400 aa in the central region and adopts a typical $\alpha\beta$-fold (Frazão *et al*, 2006). Analysis of the crystallography structure of *E. coli* RNase II revealed that nucleotides 1-5 of an RNA fragment interact with the anchor region, formed by CSD1, CSD2 and S1 domains, and the final 9-13 nts are clamped in the RNB domain between residues F358 and Y253 (Frazão *et al*, 2006). This explains why transcripts must be longer than 10 nts to be processively degraded, being the Y253 residue the one thought to be responsible for setting the size of RNase II end-product. The substitution of this residue alters the size of the end product from 4 to 10 nts, causing loosening of RNA substrate at the catalytic site (Barbas *et al*, 2008).

The structure of RNase II revealed that RNA degradation involves four highly conserved aspartate residues (D201, D207, D209 and D210) that coordinate the recruitment of two Mg^{2+} ions for catalysis. Substitution of the conserved residue E542 by an alanine, thought to help the removal of the exiting nucleotide upon phosphodiester cleavage, gave rise to a "super-enzyme", with extraordinary catalysis and binding activities. This conformational modification in the RNB domain and reorganization of the RNA-binding interface displayed 100-fold increase in exonucleolytic activity and 20-fold increase in RNA-binding affinity (Barbas *et al*, 2009). Later, it was suggested that RNase II binds to the cytoplasmic membrane via an amino-terminal amphipathic α-helix (NTH) motif that functions as a membrane binding anchor (Lu & Taghbalout, 2013).

Studies on the radiation resistant bacterium *D. radiodurans*, identified that the DrR63 protein is an RNase II-like enzyme. Although it possesses an RNB catalytic domain

and C-terminal S1 domain like RNase II, it lacks the two CSD in the N-terminal region. Instead DrR63 possesses a helix-turn-helix (HTH) motif. This HTH motif interacts with an extension of C-terminal S1 domain and it was observed that this makes DrR63 able to approach closer to an RNA duplex than *E. coli* RNase II. The length of the RNA binding path is similar to the observed for RNase II, however the latter has a narrow clamp whereas DrR63 displays an open architecture and a truncated N-terminus (Schmier *et al*, 2012).

1.3.2.4. RNase II substrates

An RNA-seq analysis comparing the wild-type strain with a deletion mutant of RNase II, identified 187 transcripts differentially expressed. Most of the transcripts were down-regulated, about 67%, compared to the up-regulated, only 29% (Pobre & Arraiano, 2015). This is probably related to the protection that RNase II can exert on RNA molecules by efficiently removing the 3' end linear sequence that are required for binding of the other major exoribonucleases (Marujo *et al*, 2000). Interestingly, most transcripts affected by RNase II are related to motility and flagellum assembly. The transcript with a higher fold in the Δ*rnb* mutant was antigen-43 or *flu*, known to promote aggregation and inhibit bacterial motility (Pobre & Arraiano, 2015). Surprisingly, RNase II shares more transcripts with PNPase than with RNase R (Pobre & Arraiano, 2015). Additionally, RNase II is involved in the mRNA degradation of stalled mRNAs that lack a stop codon. This process involves the tmRNA-SmpB complex which provides a stop codon to a stalled ribosome in order to recycle the ribosome, in a mechanism that is called *trans*-translation (Janssen & Hayes, 2012). In *E. coli*, prolonged translational arrest allows mRNA degradation into the A-site of stalled ribosomes. The enzyme that cleaves the A-site codon is not known, but its activity requires RNase II to degrade mRNA downstream of the ribosome (Janssen *et al*, 2013).

Even though its major role is the degradation of mRNA, RNase II also functions in the processing of tRNA and other stable RNAs. However, RNase II role in these reactions seems very limited. A whole-genome transcriptional study revealed that inactivation of RNase II marginally affected stable RNAs, including tRNAs and sRNAs (Pobre & Arraiano, 2015). In fact, RNase II has not generally been implicated in the degradation of regulatory

RNAs (Andrade *et al*, 2012, 2013). Endoribonucleases are thought to be involved in generating the mature 5' termini while maturation of the 3' end requires the action of exoribonucleases (Li & Deutscher, 1996). Any one of a number of exoribonucleases, including RNases T, D, PH, II, R and PNPase can carry out the trimming reaction of tRNAs *in vivo*, although RNases T and PH are the most effective (Li *et al*, 1998). Hence, in normal conditions, RNase II activity is not critical for tRNA processing. Similar observations can be made for rRNA processing. Multiple exoribonucleases can catalyze the maturation of pre-16S rRNA (Sulthana & Deutscher, 2013). However, RNase R and PNPase are the most important for the elimination of aberrant rRNA fragments under normal growth conditions (Cheng & Deutscher, 2003; Basturea *et al*, 2011). In contrast, under starvation, RNase II and RNase R are the most important exoribonucleases for rRNA fragment removal (Basturea *et al*, 2011). RNase II was found to control RNase PH activity under starvation (Sulthana *et al*, 2017). RNase PH is very unstable during starvation and stationary conditions with levels of the enzyme decreasing as much as 70%. Yet, the mechanism by which RNase II controls and regulates RNase PH is not fully understood until today but it is suggested to correlate with RNase II mediated degradation of ribosomes (Sulthana *et al*, 2017).

1.3.3. RNase R

RNase R is unique amongst the 3'-5' exoribonucleases in its remarkable capability to degrade RNA secondary structures on its own. Unlike RNase II, RNase R is not blocked by RNA stem-loops and in contrast to PNPase it does not require association with other proteins (e.g. helicases) to overcome such structures. Furthermore, RNase R is a stress regulator induced in different environmental stimuli being particularly important in the cold shock response which leads to a strong stabilization of RNA structures. Not surprising, RNase R is highly effective against a wide variety of substrates. RNase R is a key enzyme involved in different surveillance mechanisms that work to eliminate unnecessary mRNAs, defective tRNA, aberrant rRNA molecules and damaged ribosomes (Li *et al*, 2002; Jacob *et al*, 2013; Cheng & Deutscher, 2003). Overall, RNase R arises as an important quality control enzyme for translation. Much work on RNase R elucidated a novel mode of

regulation of RNA degradative enzymes. Additionally, RNase R is now used as a routinely tool in molecular biology, a feature particularly important for the study of circular RNAs (circRNAs). The continuous study of this unique enzyme will certainly continue to expand the knowledge of RNA-dependent pathways.

1.3.3.1. RNase R function and regulation

RNase R is a hydrolytic exoribonuclease belonging to the RNase II family of proteins. RNase R is a highly processive and sequence-independent enzyme that degrades RNA releasing nucleoside monophosphates, having a dinucleotide as the end product of digestion (Cheng & Deutscher, 2002). Although it shares the family typical RNB catalytic domain, RNase R comprises distinctive features from RNase II. For instance, its ability to degrade structured RNA substrates making use of an intrinsic helicase activity requiring only a single-stranded 3' overhang of at least 7 nts for effective binding (Vincent & Deutscher, 2006). Substrates with 3' overhangs as shorter as 3 nts can also be degraded, although not as efficiently (Hossain *et al*, 2016). RNase R is involved in several quality control mechanisms that act to remove defective RNAs. This exoribonuclease is effective against a wide range of substrates, from highly structured mRNAs (as those containing REP sequences) to highly structured ribosomal RNAs (Figure 6). RNase R homologues are widespread in nature and ubiquitously important in various organisms (Andrade *et al*, 2009b). Comprising both helicase and exoribonuclease activities make it a unique enzyme in the RNA degradation landscape. In recent years such outstanding features have led for RNase R to be utilized as a molecular biology tool, specifically in the study of eukaryotic circular RNAs (circRNA). The 5' and 3' ends of circRNAs are not free, instead they are covalently bound together, hence yielding a circular RNA molecule (Ebbesen *et al*, 2016). The expression of circRNAs is generally developmental stage, cell-type and tissue specific, which arouse great interest on the research community. Since circRNAs have no unprotected extremities they are essentially exoribonuclease-resistant. RNase R, was found to be the best exoribonuclease for the methodical digestion of eukaryotic linear RNAs providing an enrichment in circRNAs for further downstream experiments (Suzuki *et al*, 2006; Suzuki & Tsukahara, 2014).

The *rnr* gene which encodes RNase R, is part of an operon containing: *nsrR* (transcriptional regulator), *rnr* (exoribonuclease), *rlmB* (rRNA methyltransferase) and *yjfI* (unknown function) (Cheng *et al*, 1998). Operon transcription is thought to be driven by a σ^{70} promoter and transcript processing and decay is mainly regulated by RNase E (Cairrão & Arraiano, 2006). RNase R levels change in response to different physiological conditions, namely there is a 3- to 10-fold increase during various stresses such as cold-shock or stationary phase of growth (Cairrao *et al*, 2003; Chen & Deutscher, 2005; Andrade *et al*, 2006). However, overexpression of RNase R becomes hazardous to the bacterial cell, possibly because of its ability to degrade virtually all cellular RNAs (Cheng & Deutscher, 2002). Hence, it was hypothesized that a tightly regulated mechanism for controlling RNase R levels was in place. It was firstly discovered that such regulation was accomplished by variations in RNase R stability and that the stabilization correlated with increasing protein levels in response to stresses (Chen & Deutscher, 2010; Cairrao *et al*, 2003; Chen & Deutscher, 2005; Andrade *et al*, 2006).

Strikingly, the variations in RNase R stability were found to be controlled by the post-translational acetylation status of a single lysine residue (K544). RNase R becomes highly unstable when K544 is acetylated and is degraded; in contrast, when K544 is not modified RNase R is very stable (Liang *et al*, 2011). This nicely correlates with the fact that RNase R from exponential growing cells was shown to be acetylated, whereas stationary phase RNase R is not. This acetylation is dependent on the product of the *yfiQ* gene which encodes lysine acetyltransferase Pka. In fast growing cells, Pka carries out K544 acetylation destabilizing RNase R. However, during stationary phase or cold-shock conditions Pka is absent and acetylation activity decreases dramatically (Liang & Deutscher, 2012a).

Acetylation of the K544 residue stimulates the binding of tmRNA-SmpB which in turn promotes proteolysis through the recruitment of Lon and HslUV proteases to the N-terminal region of RNase R (Liang & Deutscher, 2012b). Conversely, tmRNA-SmpB was shown to bind only weakly to RNase R during stationary phase of growth, when K544 is not acetylated. This leads to a protein stabilization and consequently increases RNase R levels, allowing its crucial action during these stress conditions. RNase R was further found

to be highly associated with ribosomes during exponential growth (Liang & Deutscher, 2013). RNase R binding to ribosomes fully stabilizes the protein whereas its free form turns over extremely fast (~2 min). Intriguingly, tmRNA-SmpB is required for both RNase R association to ribosomes and turnover of the free form of the enzyme, although the two processes are independent. This allows the free form of RNase R to carry out its well-known functions as a stress-response post-transcriptional regulator by acting on specific RNA substrates.

1.3.3.2. RNase R complexes

As RNase R has intrinsic helicase activity, there is no specific requirement for association with RNA helicases or other proteins in multiprotein complexes to succeed in the degradation of double stranded RNAs. In fact, RNase R is not found in "normal" degradosomes. However, RNase R replaces PNPase as the exoribonuclease present in "specialized" degradosomes, formed under stress conditions. RNase R interacts with the endoribonuclease RNase E in the degradosome of the cold-adapted bacterium *Pseudomonas syringae* Lz4W (Purusharth *et al*, 2007). Curiously, RNase R is also found in the stationary-phase degradosome in *E. coli* (Carabetta *et al*, 2010). We could speculate that other conditions that result in the induction of RNase R expression may lead to its incorporation in similar multiprotein complexes. In addition, RNase R was reported to interact with the 30S subunit, specifically with the ribosomal protein S12 (Strader *et al*, 2013). However, the importance of this contact remains mostly unknown to date.

1.3.3.3. RNase R structure

RNase R is a large (~92kDa) multidomain protein with a modular arrangement like that of the RNase II family members. The catalytic RNB domain is flanked by the RNA-binding domains, the cold-shock domains CSD1 and CSD2 at the N-terminus and the S1 domain at the C-terminus (Figure 7) (Barbas *et al*, 2008). Strikingly, RNase R is bigger than RNase II since it comprises at least two additional domains: a K/R-rich domain and a helix-turn-helix domain (HTH) at the C- and N-terminal regions, respectively (Hossain *et al*, 2015; Chu *et al*, 2017). Mutational studies identified key residues for the nuclease activity

located in the active center of the RNB domain. Of notice, a D280N mutant (similar to the D209N mutant of RNase II previously reported exhibited no exonucleolytic activity (Frazão *et al*, 2006). Furthermore, the Y324 residue was identified as essential for keeping a dinucleotide as the final product of digestion (Matos *et al*, 2009). A distinctive feature of RNase R is the much stronger binding of the catalytic channel for the RNA substrates in comparison to RNase II, consequently providing a stronger pull on the RNA molecule (Vincent & Deutscher, 2009). Strikingly, a mutant RNase R expressing only the RNB catalytic domain can degrade a fully complementary double-stranded RNA molecule, whilst the presence of the RNA-binding domains restores the requirement for a linear 3' overhang. Therefore, it was suggested that the RNA-binding domains (CSD1, CSD2 and S1) would function as discriminatory domains to select RNA molecules with 3' extensions as RNase R substrates (Matos *et al*, 2009). Moreover, the flanking RNA-binding domains are responsible for the intrinsic helicase activity of RNase R.

Two additional regions were meanwhile identified in RNase R: a Walker A motif within the C-terminal K/R rich patch and a Walker B motif within the N-terminal CSD2 domain. Both the Walker A and B motifs are responsible for forming an ATP binding site when near each other. Interestingly, since both motifs are in opposite extremities of the protein, this means that the N- and C-terminal regions must come together. A conformational change is thought to occur upon dsRNA binding, bringing the Walker A and B motifs together thus forming the ATP pocket. ATP binding, but not its hydrolysis, then stabilizes the conformational rearrangement and triggers the helicase activity, thus stimulating RNA strand binding to CSD and S1 domains (Hossain *et al*, 2015).

Crystallization of full-length RNase R proved to be a challenging task. Nonetheless, a truncated form lacking the N-terminal helix-turn-helix domain and the C-terminal K/R rich region (RNase R ΔHTH-K mutant) yielded crystals that allowed the resolution of its structure (Chu *et al*, 2017). RNase R ΔHTH-K variant exhibited only a slightly slower *in vitro* activity towards a dsRNA with a 10 nucleotide 3' overhang when compared to the full-length protein, thus confirming that truncation did not significantly affect the overall enzymatic activity (Matos *et al*, 2009; Vincent & Deutscher, 2009; Chu *et al*, 2017). The crystallized protein also bared a spontaneous mutation (A131V) that enhanced

crystallization, but why this was the case remains unclear since the residue is in a disordered loop. The three-dimensional structure of other members of the RNase II family had already been determined for Rrp44, Dis3L2 and RNase II itself, which served as basis for comparative studies (Faehnle *et al*, 2014; Frazão *et al*, 2006; Lorentzen *et al*, 2008). Superimposition of the RNB domain of RNase R with those of RNase II and Rrp44 revealed that they share an overall similar structure. Strikingly, the auxiliary RNA-binding domains adopted a different position and orientation, suggesting either that these domains are flexible or, perhaps more likely, that the RNases have different RNA-binding modes.

Insights into the hydrolysis mechanism confirmed it is identical to that of RNase II (Frazão *et al*, 2006). Specifically, a two-metal ion catalytic mechanism where one Mg^{2+} is bound by residues D272 and D281 in the RNB domain, whereas a second Mg^{2+} binds only after RNA binding in the active site and is therefore absent from the 3D structure. Moreover, residues Y324 and Y383 match two tyrosine residues at the same position of RNase II that stack the 3'-end nucleotides (Frazão *et al*, 2006). It is likely that Y324 and Y383 are, therefore, responsible for setting the end-product of the digestion as a dinucleotide, corroborating previous biochemical data (Matos *et al*, 2009; Vincent & Deutscher, 2009; Chu *et al*, 2017).

Another important feature of RNase R structure is a tri-helix region located in the RNB domain that serves as a wedge for the unwinding of RNA. Curiously, in the RNA-bound crystal structure reports of the RNase II family members, the RNB tri-helix region was found to interact with the substrate RNA on Dis3L2 and Rrp44 (Faehnle *et al*, 2014; Vuković *et al*, 2016), however this interaction was not observed for the RNase II structure, which argued for the importance of the tri-helix region for the unwinding activity. In fact, mutational and biochemical experiments revealed that an RNase R deletion mutant lacking the whole tri-helix region is ineffective in unwinding its substrate (Chu *et al*, 2017). The energy for the unwinding of the RNA is therefore proposed to derive from the hydrolysis of the RNA itself, and not from ATP hydrolysis as in other RNA helicases. The hydrolysis of substrate further pulls the 3'-end of the molecule to the active site while the wedge region helps unwind the RNA (Chu *et al*, 2017; Hossain *et al*, 2015).

Published data showed that both RNase II and Dis3L2 bind a ssRNA molecule through the S1 and CSD1 motifs forming a top channel that caps the RNB catalytic motif, whereas the crystal structure of Rrp44 bound to ssRNA revealed a side channel formed by CSD1 and RNB. Strikingly, the crystal structure of RNase R revealed that it possesses two open channels: one top channel (S1 and CSD1) and one side channel (CSD1 and RNB), which is a unique feature among other members of the RNase II family (Chu *et al*, 2017). Based on the crystallographic, data two models for RNase R action can be proposed, depending on which one of the two channels is responsible for threading the RNA to the active site or serve as an exit for the non-sessile 5'-end that is unwound during degradation. One model is based on the Rrp44-RNA structure and contemplates the side channel as the entry channel, whereas a second model based on the RNase II-RNA structure suggests the top channel as the entrance. The latter is more likely to be true since measurements of both RNase R channels indicate that the top channel's width is sufficient for accommodating dsRNA, contrasting with the narrower side channel where only ssRNA could bind. Undoubtedly the resolution of RNase R crystal structure is a milestone in understanding the complex behavior of this remarkable enzyme. Nonetheless, a co-crystallization of RNase R with a structured RNA substrate would certainly allow for a refined model of its action, as happened with other members of the RNase II family of enzymes.

1.3.3.4. RNase R substrates

Under normal laboratorial conditions in bacteria grown to exponential phase, an RNA transcriptomic study revealed that the single inactivation of RNase R affects only ~6% of *E. coli* open reading frames (ORFs) (Pobre & Arraiano, 2015). Nevertheless, RNase R variants with mutations within or close to the entrance of the nuclease channel were shown to suppress the mRNA turnover defects of an RNase E mutant, which raises the question if RNase R could be more important for mRNA decay than previously thought (Hammarlöf *et al*, 2015). RNase R has been shown to participate in the elimination of REP-containing mRNAs and in the polyadenylated-dependent degradation of the *rpsO* mRNA (Cheng & Deutscher, 2005; Andrade *et al*, 2009a). RNase R is induced in response to

several stresses (Cairrao *et al*, 2003; Chen & Deutscher, 2005; Andrade *et al*, 2006). This enzyme seems particularly important in the degradation of mRNAs under these harsh conditions, namely in adaptation of cells to temperature downshift (Zhang *et al*, 2018).

An important bacterial strategy to survival is adaptation to temperature fluctuations, which involves a rapid and extensive reprogramming of gene expression. Bacteria exposed to significantly sub-optimal temperatures develop a cold-shock response that is characterized by an acclimation phase during which cell growth temporarily stops. During this phase, even though translation is drastically reduced, some cold shock proteins are produced, which helps cells to adapt to low temperature (Phadtare & Severinov, 2010). RNase R is one of these cold shock proteins and its induction occurs through mRNA and protein stabilization and increased translation after the temperature downshift (Cairrão & Arraiano, 2006; Liang *et al*, 2011). Inactivation of RNase R leads to growth deficiency at low temperatures in different bacteria (Söderberg & Cianciotto, 2010; Cairrão & Arraiano, 2006). One of the main consequences of cold shock is the stabilization of RNA secondary structures which correlates with lower translation levels. For a long time, it was suspected that RNase R was important in mRNA decay at low temperature. In *Legionella pneumophila*, inactivation of RNase R results in the accumulation of structured RNAs specifically in cold shock (Charpentier *et al*, 2008). *E. coli* RNase R, but not RNase II or PNPase, can substitute for the RNA helicase CsdA, an essential gene involved in mRNA decay in low temperature (Awano *et al*, 2007). These two enzymes were suggested to share some common target mRNAs (Phadtare, 2012). The role of RNase R in mRNA degradation during acclimation was recently confirmed using global methods to monitor the genome-wide changes in translation (Zhang *et al*, 2018). This study demonstrated that RNase R is part of a surveillance system that eliminates the poorly translated structured mRNAs that are stabilized by low temperature and may compete to access ribosomes, thus playing a key role for translation recovery after cold shock.

It is noticeable the wide range of RNA substrates that are targeted by this exoribonuclease. However, because of its natural ability towards the degradation of RNA secondary structures, RNase R has major roles in the quality control and degradation of

rRNA (Zundel *et al*, 2009; Cheng & Deutscher, 2003). Bacterial ribosomes are extremely stable in exponential growing cells. However, under stresses that lead to slow growth, there is a lower demand for functional ribosomes because of a global decrease in translation. Therefore, there is a need for recycling the constituents of the macromolecules to ensure energy efficiency and cell survival. Endoribonucleases break down the large rRNA molecules internally to smaller fragments that are subsequently degraded by 3'-5' exoribonucleases. It has been shown that elimination of intermediate rRNA fragments during starvation conditions is carried out by RNase R and RNase II. Accumulation of these fragments is hazardous to the cell since there is no recycling of ribonucleotides (Kaplan & Apirion, 1975; Zundel *et al*, 2009).

As mentioned above, an RNase R and PNPase double mutant is synthetic lethal. This inviable phenotype was shown to arise shortly after 16S and 23S rRNA fragments accumulated to high levels in a conditional RNase R^- $PNPase^-$ double mutant. These fragments were also suggested to be generated after initial endonucleolytic cleavages (Cheng & Deutscher, 2003). Both RNase R and PNPase were thus identified as being part of the same quality control mechanism that eliminates aberrant rRNAs during steady-state growth as soon as they are generated. Failing to do so interferes with correct ribosome maturation leading to cell death. Additionally, RNase R was described to take part in a novel late-stage quality control mechanism that also involves YbeY, a newly identified endoribonuclease (Jacob *et al*, 2013). Depletion of YbeY leads to accumulation of misprocessed 16S rRNA precursor and immature 30S subunits, which argues for a role in the processing of rRNA. Strikingly, YbeY and RNase R were found to mediate the specific *in vitro* degradation of 70S ribosomes baring defective 30S subunits. The immature small 30S subunits were found to be the trigger that allowed YbeY to distinguish between defective and non-defective ribosomes. The latter are thus thought to be initially targeted for endonucleolytic cleavages by YbeY followed by RNase R unwinding and processive digestion (Jacob *et al*, 2013). The requirement for both YbeY/RNase R for this quality control mechanism is substantiated by the observation that *ybeY* and *rnr* genes were shown to have a strong genetic interaction, which could also explain the high conservation of both proteins throughout evolution (Davies *et al*, 2010).

The role of RNase R in rRNA and ribosome homeostasis is further underlined by the fact that it participates in the 3'-end maturation of the 17S rRNA (16S rRNA precursor), a functional overlap with PNPase, RNase II and RNase PH (Sulthana & Deutscher, 2013). This finding places RNase R not only as part of degradation and surveillance machinery, but also on the list of 16S rRNA processing enzymes (Smith *et al*, 2018). A role in the maturation of the 3'-end of rRNAs had been previously described for RNase R homologues in the Gram-negative bacteria *P. syringae* during cold-shock and in the eukaryotic organism *Arabidopsis thaliana*, which argues in favor of a conserved widespread function (Bollenbach *et al*, 2005; Purusharth *et al*, 2007).

RNase R also participates in tRNA quality control by degrading defective tRNAs in *E. coli* (Li *et al*, 2002; Vincent & Deutscher, 2006). Addition of either poly(A) or CCACCA tails to the tRNA stimulates the RNase R activity against tRNA (Wellner *et al*, 2018). This exoribonuclease seems to have a more significant role as a degradative enzyme rather than processing enzyme of tRNA in *E. coli*. However, RNase R was suggested to play an important function in tRNA processing in other bacteria. Shorter $tRNA^{Cys}$ species accumulated only when RNase R was also absent, suggesting that this RNase R has a principal role in the quality control of tRNAs in *B. subtilis* (Campos-Guillén *et al*, 2010). RNase R was also shown to be the responsible enzyme to remove tRNA 3'-trailers and generate mature 3'-ends in tRNAs from *Mycoplasma genitalium* (Alluri & Li, 2012).

RNase R is also required for the correct processing of the tmRNA in *E. coli* during cold shock and in the degradation of the tmRNA in *C. crescentus* during cell cycle (Russell & Keiler, 2009; Cairrao *et al*, 2003). This RNA molecule holds the unique bifunctional ability of acting as both a tRNA and mRNA, and in concert with the protein SmpB directs the stalled ribosome to its own ORF tagging the aberrant protein for proteolysis. Moreover, RNase R was shown to copurify with the tmRNA-SmpB complex (Karzai & Sauer, 2001; Venkataraman *et al*, 2014). RNase R is suggested to degrade the aberrant nonstop mRNA and thus play a key role in the *trans*-translation mechanism of protein quality control (Richards *et al*, 2006; Ge *et al*, 2010). Surprisingly, RNase R is not found to greatly affect the stability of small RNAs (Andrade & Arraiano, 2008; Andrade *et al*, 2012). Nevertheless, it was reported that RNase R is involved in the degradation of both the sRNA

SR4 and its target *bsr*G mRNA of *B. subtilis* (Jahn *et al*, 2012). We anticipate that the continuous study of sRNA stability pathways will highlight the role of RNase R in the metabolism of many more regulatory RNAs.

1.4. Hfq

The bacterial Hfq is an abundant and phylogenetically conserved RNA chaperone that can remodel RNA secondary structures. Although it has been reported to bind ATP, requires no energy input to carry out its molecular function involving RNA transactions (Santiago-Frangos & Woodson, 2018). Hfq belongs to an extensive RNA-binding protein family, the Sm/Lsm superfamily, with homologues found in almost every organism (Sauer, 2013; Wilusz & Wilusz, 2013). Its N-terminal region contains two motifs – Sm1 and Sm2 – which are characteristic of this family. Conversely, the C-terminal region is more variable and disordered, and its functions have only now begun to be understood (Link *et al*, 2009; Vogel & Luisi, 2011; Santiago-Frangos *et al*, 2017). Initially, Hfq was identified in the late 1960's as an essential host factor for RNA bacteriophage Qβ infection in *E. coli*, as it was found to promote the melting a 3' secondary structure of the viral RNA necessary for efficient replication (Franze de Fernandez *et al*, 1968). Since then, the bacterial Hfq has been implicated in many cellular processes, mainly through the post-transcriptional regulation of gene expression. Additionally, the eukaryotic and archaeal Hfq counterparts were also shown to be involved in various biological pathways.

Hfq pleiotropic functions were substantially uncovered upon disruption of the *hfq* gene in *E. coli*, which leaded to the rising diverse defective phenotypes (Tsui *et al*, 1994). Namely, decreasing bacterial growth rate, changing mutagenesis rate, increase in UV sensitivity, oxidant sensitivity and osmo-sensitivity. However, how this RNA chaperone was able to regulate so many different pathways was a mystery. To date, Hfq is widely known as the RNA-binding protein responsible for catalyzing the basepairing between small non-coding RNAs and mRNA.

1.4.1. Hfq function and regulation

Small non-coding RNAs are small RNA molecules with regulatory functions that typically act on messenger RNAs to alter their stability or translation profile (Waters & Storz, 2009; Storz *et al*, 2011; Andrade *et al*, 2012, 2013). sRNAs that are encoded in the opposite strand of its mRNA target are termed *cis*-encoded sRNAs. Opposingly, sRNAs genes can be far apart from their regulated targets and are therefore referred to as *trans*-encoded sRNAs. The first necessarily exhibit a perfect complementarity to their target mRNA, whilst the latter often display a limited and imperfect basepairing ability. Additionally, *trans*-encoded sRNAs can regulate multiple mRNA targets, which in turn code for functionally diverse proteins, involved in an array of cellular pathways (Vogel & Luisi, 2011; Kavita *et al*, 2018).

Depending on the binding specificities of the sRNA to its target, the post-transcriptional regulation can either repress or activate mRNA expression. Accordingly, the same sRNA can repress or activate different mRNA targets, which adds a layer of complexity to this regulatory circuit (Figure 8). Positive mechanisms of sRNA-mediated regulation typically involve sRNA annealing to the 5' end of the mRNA target that lead to the remodel of inhibitory RNA secondary structures that occlude the ribosome binding site (RBS) (Soper *et al*, 2010). Hence, an exposure of the RBS leads to an enhanced translation of the target mRNA. Alternatively, sRNAs can bind and sequester RNase cleavage sites present in their target mRNAs enhancing their stability (Fröhlich *et al*, 2013). Inversely, sRNAs can negatively regulate their targets, which is the most observed sRNA-mediated type of gene expression regulation. sRNAs can basepair to or near the RBS, obstructing ribosome binding and repressing translation (De Lay *et al*, 2013). Moreover, sRNAs can directly target the mRNA for degradation by RNase E (Morita *et al*, 2005). Yet another type of regulation uses sRNAs as a sponge molecule that binds other sRNAs or function as a sponge for RNA-binding proteins, leading to an indirect regulation of gene expression (Figueroa-Bossi *et al*, 2009; Lalaouna *et al*, 2015).

Hfq is one of the RNA-binding proteins that was found to catalyze the imperfect basepairing between *trans*-encoded sRNAs and their targets. Consequently, Hfq is able to participate in this variety of mechanisms of post-transcriptional gene regulation. Because

of its chaperoning activity, Hfq was shown to be critical for the transient stabilization of several of these basepair interactions (Santiago-Frangos & Woodson, 2018). These sRNAs often regulate more than just one target mRNA while the same mRNA can be regulated by different sRNAs. This creates a complex circuitry of sRNA regulators that oftentimes overlap, whether synergistically for the same outcome or in antagonizing pathways (Wang *et al*, 2015). The cellular levels of Hfq are thought to be maintained within a limit range because of an autoregulatory mechanism where Hfq is able to bind to its own 5' UTR, masking the RBS and consequently inhibiting translation (Vecerek *et al*, 2005). This means that cellular RNAs compete among themselves for Hfq binding, and the outcome of specific regulatory pathway is ultimately dictated by the amount of the available RNA (Santiago-Frangos & Woodson, 2018). Hfq-mediated regulation through sRNAs impacts various biological processes many times involved in stress response (Papenfort & Vogel, 2009). Therefore, the pleiotropic phenotype of Hfq inactivation in *E. coli* is largely attributed to a disruption in the sRNA regulation network (Vogel & Luisi, 2011; Hajnsdorf & Boni, 2012; Updegrove *et al*, 2016). Nevertheless, sRNA-mediated gene regulation. can still occur in the absence of Hfq, although at slower rates. In fact, nearly two thirds of the sRNAs identified in *E. coli* do not rely on Hfq for their function (Olejniczak & Storz, 2017). Moreover, Hfq is not involved in sRNA-mediated gene regulation in many bacteria, including *Bacillus subtilis*, *Listeria monocytogenes* and *Caulobacter crescentus*, suggesting the existence of alternative RNA chaperone. (Christiansen *et al*, 2006; Rochat *et al*, 2015; Fröhlich *et al*, 2018).

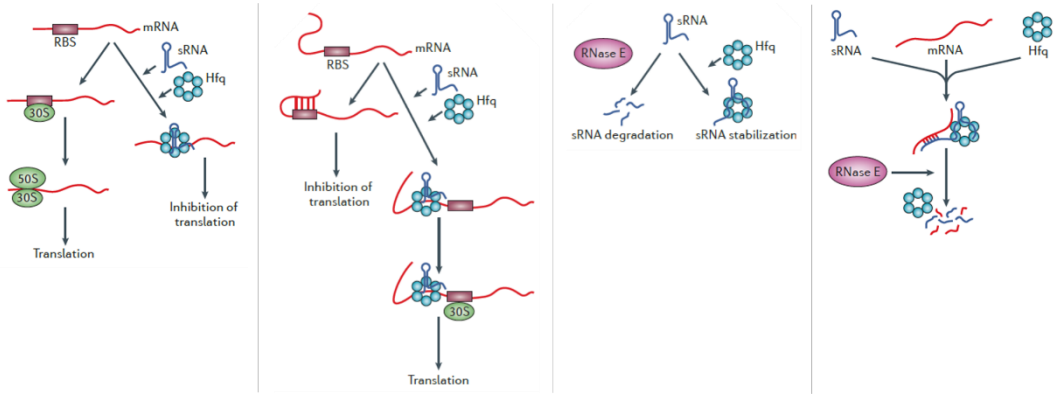


Figure 8 – Various mechanisms of regulation by sRNA/Hfq (adapted from Vogel & Luisi, 2011).

The recent discovery of ProQ has confirmed that Hfq does not stand alone in the midst of sRNA regulatory networks. ProQ is an RNA-binding protein of the ProQ/FinO family that has been demonstrated to bind several transcripts, including sRNAs, with an impact on cellular physiology (Smirnov *et al*, 2016). Moreover, *Salmonella enterica* ProQ was found to stabilize the ProQ-dependent sRNA RaiZ and to promote its basepairing with the RBS of the *trans*-encoded *hupA* target. The stabilization of these interaction inhibits 30S loading into the transcript effectively downregulating its translation, in a mechanism reminiscent of those already described for some Hfq-associated sRNAs (Smirnov *et al*, 2017).

1.4.2. Hfq structure and RNA binding surfaces

Hfq makes use of different RNA-binding surfaces and disordered domains to recognize a wide variety of RNA substrates. Structurally, Hfq assembles into a homohexameric ring with a central pore (Link *et al*, 2009; Vogel & Luisi, 2011). This architecture exposes different protein surfaces (Figure 9). Namely, the proximal face – as the surface where the N-terminal α-helix is located –, the distal face – in the opposite side of the proximal face. Additionally, the lateral rim also engages in substrate binding, as well as the C-terminal tail (Panja *et al*, 2013; Schu *et al*, 2015; Santiago-Frangos *et al*, 2017). Accordingly, each surface interacts differently with the RNA molecules. The proximal face, conserved among the Sm/Lsm proteins, interacts preferentially with single-stranded U-rich sequences, typically found at the 3' end of sRNAs (Vogel & Luisi, 2011; Sauer, 2013; Kavita *et al*, 2018). Conversely, the distal face binds A-rich sequences, specifically ARN repeats, commonly found in many mRNA molecules. In addition, the lateral rim surface exhibits basic patches that are also able to bind complementary regions of Hfq-bound RNAs, often harboring AU-rich sequences. Finally, the C-terminal tail, with its acidic nature, has been suggested to promote de releasing of RNAs from Hfq core, which could be nonspecifically bound (Santiago-Frangos *et al*, 2017). Moreover, this model proposes that the acidic tip allows Hfq to rapidly bind and release RNAs, cycling through the great

variety of cellular substrates, until it finds a suitable basepair interaction (Woodson *et al*, 2018).

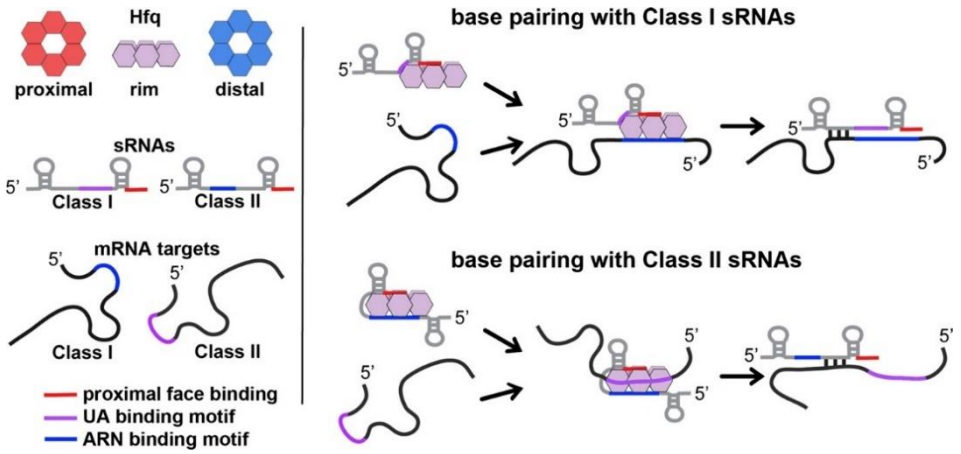


Figure 9 – Representation of the preferential binding of three different Hfq RNA binding surfaces towards a specific class of sRNAs (adapted from Schu et al, 2015).

Hfq was found to interact with different RNAs, including sRNAs, mRNAs and rRNA (de Haseth & Uhlenbeck, 1980; Muffler *et al*, 1996; Fröhlich *et al*, 2013; Updegrove *et al*, 2016; Kavita *et al*, 2018). Given the sequence diversity of Hfq interactors, a subset of sRNAs were also shown to bind to the distal face, disrupting the canonical view for sRNA-Hfq-mRNA binding. Moreover, Hfq was found to autoregulate its levels directly binding its mRNA 5' UTR specifically through the distal face, in a sRNA-independent manner (Morita & Aiba, 2018). Making use of a various *hfq* alleles bearing specific point mutations, Schu and colleagues were able to disrupt either the proximal, rim or distal RNA binding surfaces and assess the loss-of-binding from known Hfq-dependent sRNAs. As a result, sRNAs were differentiated into two classes, depending on which of Hfq surface they preferentially bind. Class I sRNAs depend on the proximal face and lateral rim and are the prominent class, while class II sRNAs interact with the distal face and are less often found (Schu *et al*, 2015). Curiously, different class sRNAs tend to act on mRNA targets that harbor different Hfq binding motifs. Accordingly, Class I sRNAs regulate mRNAs with the ARN motif that

bind the distal face, whereas class II sRNAs regulate rim-binding mRNAs bearing the AU-rich motif (Schu *et al*, 2015; Zhang *et al*, 2013; Santiago-Frangos & Woodson, 2018).

The role of Hfq as an RNA matchmaker involved in sRNA-mediated gene expression regulation is firmly established. Nevertheless, the observation that other sRNA-binding proteins, like ProQ, are also implicated in this regulatory network and that Hfq is not involved in sRNA-mediated gene regulation in many bacteria raises the possibility for previously unexplored roles for this RNA chaperone.

1.4.3. Other functions for Hfq beyond sRNA-mediated regulation

Hfq discovery was related to its RNA remodeling activity on a viral RNA, without which, no efficient replication could occur (Franze de Fernandez *et al*, 1968). Additionally, the pleiotropic phenotypes observed upon Hfq inactivation raised significant interest on the protein (Tsui *et al*, 1994). Combined with the fact that Hfq was shown to be a major regulator of sRNA-mediated control of gene expression, attention from other potentially important functions for Hfq were diverted. Supporting the notion that Hfq actions extend beyond the sRNA-mRNA matchmaker task is the fact that Hfq was found to interact with other proteins of the RNA degradation machinery. Namely, Hfq could copurify with the endoribonuclease RNase E in a complex that was proposed to be guided by sRNA to the mRNA target for degradation (Morita *et al*, 2005; Ikeda *et al*, 2011; Morita & Aiba, 2011). RNase E is also the central component of the bacterial degradosome, a multi-protein complex involved in RNA degradation (Vanzo *et al*, 1998; Arraiano *et al*, 2010; Mohanty & Kushner, 2016). RNase E provides the scaffold for binding of additional proteins, including the RhlB RNA helicase and the exoribonuclease PNPase (Bruce *et al*, 2018). Hfq was shown to copurify specifically with RNase E and not with other components of the degradosome. However, another study suggests an interaction between Hfq and PNPase (Mohanty *et al*, 2004). This may indicate that Hfq interacts with free PNPase and not when this RNase is complexed with the degradosome. The biological implications for this interactions are still elusive but Hfq can help PNPase degrade sRNAs, which usually contain strong secondary structures (Andrade & Arraiano, 2008; De Lay & Gottesman, 2011; Andrade *et al*, 2012, 2013; Bandyra *et al*, 2016). Additionally, Hfq also interacts with the poly(A) polymerase I

(PAPI) stimulating the addition of poly(A) tails on mRNAs (Hajnsdorf & Régnier, 2000; Le Derout *et al*, 2003; Mohanty *et al*, 2004; Régnier & Hajnsdorf, 2013). Moreover, Hfq is also suggested to impact RNA transcription since it was shown to interact with the RNA polymerase complex and the Rho terminator protein (Sukhodolets & Garges, 2003; Rabhi *et al*, 2011).

Hfq interaction with many components of the RNA biology landscape expands the Hfq impact beyond sRNA regulatory networks. Interestingly, Hfq was shown to be one of the nucleoid associated proteins that shape the bacterial chromosome (Azam & Ishihama, 1999). Notably, *in vivo* evidence was revealed by the sub-cellular localization of Hfq showing that approximately 20% was complexed with DNA (Diestra *et al*, 2009). Hfq mediates DNA compaction presumably through bridging DNA segments (Jiang *et al*, 2015). This is function depends on Hfq distal face and its C-terminal domain (Updegrove *et al*, 2010). The latter region exhibits an amyloid-like structure that can assemble Hfq on the chromosome and the DNA seems to induce the organization of Hfq into filaments, ultimately leading to DNA compaction (Cech *et al*, 2016; Malabirade *et al*, 2018). This is substantiated by the fact that Hfq was found to form fiber-like structures as observed through transmission electron microscopy assays (Cech *et al*, 2016). However, alteration of DNA topology previously observed in Hfq deletions strains is now believed to be an interdict effect through regulation of other proteins, since Hfq alone cannot alter de DNA topology (Malabirade *et al*, 2018). In addition, Hfq has been suggested to be directly involved in the regulation of cellular replication and transposition of several transposon systems (Cech, 2016). Although the exact mechanism that allows Hfq regulation of DNA compaction is still elusive, it is clear that the widely known RNA-binding protein also binds DNA and is an important factor in chromosome structure.

Hfq impacts a wide variety of aspects in cellular physiology that are not directly linked to its well characterized sRNA-mRNA matchmaker role. Altogether, these observations demonstrate Hfq versatility and impact on many biological processes and argue in favor of a functional interaction with the diverse set of substrates and protein partners. Notably, in the early 1980's, Hfq was shown to bind 16S rRNA *in vitro* but a possible function underlying this interaction was not investigated (de Haseth &

Uhlenbeck, 1980). Moreover, a proteomic study identified Hfq as a possible interacting protein of r-protein S12 (Strader *et al*, 2013). Since it seems able to interact with ribosomal components, is it possible that this extensively studied protein still holds undiscovered functions?

1.5. References

Alluri RK & Li Z (2012) Novel one-step mechanism for tRNA 3'-end maturation by the exoribonuclease RNase R of *Mycoplasma genitalium*. *J. Biol. Chem.* **287:** 23427–33

Andrade JM & Arraiano CM (2008) PNPase is a key player in the regulation of small RNAs that control the expression of outer membrane proteins. *RNA* **14:** 543–51

Andrade JM, Cairrão F & Arraiano CM (2006) RNase R affects gene expression in stationary phase: regulation of ompA. *Mol. Microbiol.* **60:** 219–28

Andrade JM, Hajnsdorf E, Régnier P & Arraiano CM (2009a) The poly(A)-dependent degradation pathway of *rpsO* mRNA is primarily mediated by RNase R. *RNA* **15:** 316–26

Andrade JM, Pobre V & Arraiano CM (2013) Small RNA modules confer different stabilities and interact differently with multiple targets. *PLoS One* **8:** e52866

Andrade JM, Pobre V, Matos AM & Arraiano CM (2012) The crucial role of PNPase in the degradation of small RNAs that are not associated with Hfq. *RNA* **18:** 844–55

Andrade JM, Pobre V, Silva IJ, Domingues S & Arraiano CM (2009b) The role of 3'-5' exoribonucleases in RNA degradation. *Prog. Mol. Biol. Transl. Sci.* **85:** 187–229

Arnold RJ & Reilly JP (1999) Observation of *Escherichia coli* ribosomal proteins and their posttranslational modifications by mass spectrometry. *Anal. Biochem.* **269:** 105–12

Arraiano CM, Andrade JM, Domingues S, Guinote IB, Malecki M, Matos RG, Moreira RN, Pobre V, Reis FP, Saramago M, Silva IJ & Viegas SC (2010) The critical role of RNA processing and degradation in the control of gene expression. *FEMS Microbiol. Rev.* **34:** 883–923

Awano N, Xu C, Ke H, Inoue K, Inouye M & Phadtare S (2007) Complementation analysis of the cold-sensitive phenotype of the *Escherichia coli* csdA deletion strain. *J. Bacteriol.* **189:** 5808–15

Azam TA & Ishihama A (1999) Twelve species of the nucleoid-associated protein from *Escherichia coli*. Sequence recognition specificity and DNA binding affinity. *J. Biol. Chem.* **274:** 33105–13

Bandyra KJ, Sinha D, Syrjanen J, Luisi BF & De Lay NR (2016) The ribonuclease polynucleotide phosphorylase can interact with small regulatory RNAs in both protective and degradative modes. *RNA* **22:** 360–372

Barbas A, Matos RG, Amblar M, López-Viñas E, Gomez-Puertas P & Arraiano CM (2008) New insights into the mechanism of RNA degradation by ribonuclease II: Identification of the residue responsible for setting the RNase II end product. *J. Biol. Chem.* **283:** 13070–13076

Barbas A, Matos RG, Amblar M, López-Viñas E, Gomez-Puertas P & Arraiano CM (2009) Determination of key residues for catalysis and RNA cleavage specificity. One mutation turns RNase II Into a 'super-enzyme'. *J. Biol. Chem.* **284:** 20486–20498

Basturea GN, Zundel MA & Deutscher MP (2011) Degradation of ribosomal RNA during starvation: comparison to quality control during steady-state growth and a role for RNase PH. *RNA* **17:** 338–45

Bermúdez-Cruz RM, Fernández-Ramírez F, Ramírez F, Kameyama-Kawabe L & Montañez C (2005) Conserved domains in polynucleotide phosphorylase among eubacteria. *Biochimie* **87:** 737–45

Bernstein JA, Lin P-H, Cohen SN & Lin-Chao S (2004) Global analysis of *Escherichia coli* RNA degradosome function using DNA microarrays. *Proc. Natl. Acad. Sci.* **101:** 2758–2763

Bollenbach TJ, Lange H, Gutierrez R, Erhardt M, Stern DB & Gagliardi D (2005) RNR1, a 3'-5' exoribonuclease belonging to the RNR superfamily, catalyzes 3' maturation of chloroplast ribosomal RNAs in *Arabidopsis thaliana*. *Nucleic Acids Res.* **33:** 2751–63

Briani F, Del Favero M, Capizzuto R, Consonni C, Zangrossi S, Greco C, De Gioia L, Tortora P & Dehò G (2007) Genetic analysis of polynucleotide phosphorylase structure and functions. *Biochimie* **89:** 145–57

Brink MF, Verbeet MP & de Boer HA (1993) Formation of the central pseudoknot in 16S rRNA is essential for initiation of translation. *EMBO J.* **12:** 3987–96

Bruce HA, Du D, Matak-Vinkovic D, Bandyra KJ, Broadhurst RW, Martin E, Sobott F,

Shkumatov A V & Luisi BF (2018) Analysis of the natively unstructured RNA/protein-recognition core in the *Escherichia coli* RNA degradosome and its interactions with regulatory RNA/Hfq complexes. *Nucleic Acids Res.* **46:** 387–402

Bylund GO, Persson BC, Lundberg LA & Wikström PM (1997) A novel ribosome-associated protein is important for efficient translation in *Escherichia coli*. *J. Bacteriol.* **179:** 4567–74

Bylund GO, Wipemo LC, Lundberg LA & Wikström PM (1998) RimM and RbfA are essential for efficient processing of 16S rRNA in *Escherichia coli*. *J. Bacteriol.* **180:** 73–82

Byrgazov K, Vesper O & Moll I (2013) Ribosome heterogeneity: another level of complexity in bacterial translation regulation. *Curr. Opin. Microbiol.* **16:** 133–9

Cairrão F & Arraiano CM (2006) The role of endoribonucleases in the regulation of RNase R. *Biochem. Biophys. Res. Commun.* **343:** 731–7

Cairrão F, Chora Â, Zilhão R, Carpousis AJ & Arraiano CM (2001) RNase II levels change according to the growth conditions: Characterization of gmr, a new *Escherichia coli* gene involved in the modulation of RNase II. *Mol. Microbiol.* **39:** 1550–1561

Cairrao F, Cruz A, Mori H, Arraiano CM, Cairrão F, Cruz A, Mori H, Arraiano CM, Cairrao F, Cruz A, Mori H & Arraiano CM (2003) Cold shock induction of RNase R and its role in the maturation of the quality control mediator SsrA/tmRNA. *Mol. Microbiol.* **50:** 1349–60

Cameron TA & De Lay NR (2016) The phosphorolytic exoribonucleases polynucleotide phosphorylase and RNase PH stabilize sRNAs and facilitate regulation of their mRNA targets. *J. Bacteriol.* **198:** 3309–3317

Del Campo M & Ofengand J (2004) Modified Nucleosides of *Escherichia coli* Ribosomal RNA. *EcoSal Plus* **1:**

Campos-Guillén J, Arvizu-Gómez JL, Jones GH & Olmedo-Alvarez G (2010) Characterization of tRNA(Cys) processing in a conditional *Bacillus subtilis* CCase mutant reveals the participation of RNase R in its quality control. *Microbiology* **156:** 2102–11

Cannistraro VJ & Kennell D (1999) The reaction mechanism of ribonuclease II and its interaction with nucleic acid secondary structures. *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.* **1433:** 170–187

Carabetta VJ, Silhavy TJ & Cristea IM (2010) The response regulator SprE (RssB) is required for maintaining poly(A) polymerase I-degradosome association during stationary phase. *J. Bacteriol.* **192:** 3713–21

Cardenas PP, Carrasco B, Sanchez H, Deikus G, Bechhofer DH & Alonso JC (2009) *Bacillus subtilis* polynucleotide phosphorylase 3'-to-5' DNase activity is involved in DNA repair. *Nucleic Acids Res.* **37:** 4157–69

Carpousis AJ, Luisi BF & McDowall KJ (2009) Endonucleolytic initiation of mRNA decay in *Escherichia coli*. *Prog. Mol. Biol. Transl. Sci.* **85:** 91–135

Carzaniga T, Dehò G & Briani F (2015) RNase III-Independent Autogenous Regulation of *Escherichia coli* Polynucleotide Phosphorylase via Translational Repression. *J. Bacteriol.* **197:** 1931–1938

Carzaniga T, Mazzantini E, Nardini M, Regonesi ME, Greco C, Briani F, De Gioia L, Dehò G & Tortora P (2014) A conserved loop in polynucleotide phosphorylase (PNPase) essential for both RNA and ADP/phosphate binding. *Biochimie* **97:** 49–59

Cascante-Estepa N, Gunka K & Stülke J (2016) Localization of Components of the RNA-Degrading Machine in *Bacillus subtilis*. *Front. Microbiol.* **07:** 1492

Cech GM, Szalewska-Pałasz A, Kubiak K, Malabirade A, Grange W, Arluison V & Węgrzyn G (2016) The *Escherichia coli* Hfq Protein: An Unattended DNA-Transactions Regulator. *Front. Mol. Biosci.* **3:** 36

Charollais J, Dreyfus M & Iost I (2004) CsdA, a cold-shock RNA helicase from *Escherichia coli*, is involved in the biogenesis of 50S ribosomal subunit. *Nucleic Acids Res.* **32:** 2751–9

Charpentier X, Faucher SP, Kalachikov S & Shuman HA (2008) Loss of RNase R induces competence development in Legionella pneumophila. *J. Bacteriol.* **190:** 8126–36

Chen C & Deutscher MP (2005) Elevation of RNase R in response to multiple stress conditions. *J. Biol. Chem.* **280:** 34393–6

Chen C & Deutscher MP (2010) RNase R is a highly unstable protein regulated by growth phase and stress. *RNA* **16:** 667–72

Chen SS & Williamson JR (2013) Characterization of the ribosome biogenesis landscape in *E. coli* using quantitative mass spectrometry. *J. Mol. Biol.* **425:** 767–79

Chen X, Sim S, Wurtmann EJ, Feke A & Wolin SL (2014) Bacterial noncoding Y RNAs are widespread and mimic tRNAs. *RNA* **20:** 1715–1724

Chen X, Taylor DW, Fowler CC, Galan JE, Wang H-W & Wolin SL (2013) An RNA Degradation Machine Sculpted by Ro Autoantigen and Noncoding RNA. *Cell* **153:** 166–177

Cheng Z-F & Deutscher MP (2003) Quality control of ribosomal RNA mediated by polynucleotide phosphorylase and RNase R. *Proc. Natl. Acad. Sci. U. S. A.* **100:** 6388–93

Cheng Z-F & Deutscher MP (2005) An important role for RNase R in mRNA decay. *Mol. Cell* **17:** 313–8

Cheng Z-FF & Deutscher MP (2002) Purification and characterization of the *Escherichia coli* exoribonuclease RNase R. Comparison with RNase II. *J. Biol. Chem.* **277:** 21624–9

Cheng ZF, Zuo Y, Li Z, Rudd KE & Deutscher MP (1998) The vacB gene required for virulence in *Shigella flexneri* and *Escherichia coli* encodes the exoribonuclease RNase R. *J. Biol. Chem.* **273:** 14077–80

Christiansen JK, Nielsen JS, Ebersbach T, Valentin-Hansen P, Søgaard-Andersen L & Kallipolitis BH (2006) Identification of small Hfq-binding RNAs in *Listeria monocytogenes*. *RNA* **12:** 1383–96

Chu L-Y, Hsieh T-J, Golzarroshan B, Chen Y-P, Agrawal S & Yuan HS (2017) Structural insights into RNA unwinding and degradation by RNase R. *Nucleic Acids Res.* **45:** 12015–12024

Clatterbuck Soper SF, Dator RP, Limbach PA & Woodson SA (2013) In vivo X-ray footprinting of pre-30S ribosomes reveals chaperone-dependent remodeling of late assembly intermediates. *Mol. Cell* **52:** 506–16

Condon C, Liveris D, Squires C, Schwartz I & Squires CL (1995a) rRNA operon multiplicity in *Escherichia coli* and the physiological implications of *rrn* inactivation. *J. Bacteriol.* **177:** 4152–6

Condon C, Philips J, Fu ZY, Squires C & Squires CL (1992) Comparison of the expression of the seven ribosomal RNA operons in *Escherichia coli*. *EMBO J.* **11:** 4175–85

Condon C, Squires C & Squires CL (1995b) Control of rRNA transcription in *Escherichia coli*.

Microbiol. Rev. **59:** 623–45

Connolly K & Culver G (2009) Deconstructing ribosome construction. *Trends Biochem. Sci.* **34:** 256–63

Dammel CS & Noller HF (1995) Suppression of a cold-sensitive mutation in 16S rRNA by overexpression of a novel ribosome-binding factor, RbfA. *Genes Dev.* **9:** 626–37

Das SK, Sokhi UK, Bhutia SK, Azab B, Su Z -z. Z-ZZ -z., Sarkar D & Fisher PB (2010) Human polynucleotide phosphorylase selectively and preferentially degrades microRNA-221 in human melanoma cells. *Proc. Natl. Acad. Sci. U. S. A.* **107:** 11948–53

Davies BW, Köhrer C, Jacob AI, Simmons LA, Zhu J, Aleman LM, Rajbhandary UL & Walker GC (2010) Role of *Escherichia coli* YbeY, a highly conserved protein, in rRNA processing. *Mol. Microbiol.* **78:** 506–18

Davis JH & Williamson JR (2017) Structure and dynamics of bacterial ribosome biogenesis. *Philos. Trans. R. Soc. B Biol. Sci.* **372:** 20160181

Le Derout J, Folichon M, Briani F, Dehò G, Régnier P & Hajnsdorf E (2003) Hfq affects the length and the frequency of short oligo(A) tails at the 3' end of *Escherichia coli rpsO* mRNAs. *Nucleic Acids Res.* **31:** 4017–23

Deutscher MP (2009) Maturation and degradation of ribosomal RNA in bacteria. *Prog. Mol. Biol. Transl. Sci.* **85:** 369–91

Deutscher MP & Reuven NB (1991) Enzymatic basis for hydrolytic versus phosphorolytic mRNA degradation in *Escherichia coli* and *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U. S. A.* **88:** 3277–80

Diestra E, Cayrol B, Arluison V & Risco C (2009) Cellular Electron Microscopy Imaging Reveals the Localization of the Hfq Protein Close to the Bacterial Membrane. *PLoS One* **4:** e8301

Donovan WP & Kushner SR (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:** 120–4

Durand S & Condon C (2018) RNases and Helicases in Gram-Positive Bacteria. *Microbiol. Spectr.* **6:**

Dziembowski A, Lorentzen E, Conti E & Séraphin B (2007) A single subunit, Dis3, is

essentially responsible for yeast exosome core activity. *Nat. Struct. Mol. Biol.* **14:** 15–22

Ebbesen KK, Kjems J & Hansen TB (2016) Circular RNAs: Identification, biogenesis and function. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1859:** 163–168

Faehnle CR, Walleshauser J & Joshua-Tor L (2014) Mechanism of Dis3l2 substrate recognition in the Lin28-let-7 pathway. *Nature* **514:** 252–256

Del Favero M, Mazzantini E, Briani F, Zangrossi S, Tortora P & Dehò G (2008) Regulation of *Escherichia coli* polynucleotide phosphorylase by ATP. *J. Biol. Chem.* **283:** 27355–9

Figueroa-Bossi N, Valentini M, Malleret L, Fiorini F & Bossi L (2009) Caught at its own game: regulatory small RNA inactivated by an inducible transcript mimicking its target. *Genes Dev.* **23:** 2004–15

Fontaine F, Gasiorowski E, Gracia C, Ballouche M, Caillet J, Marchais A & Hajnsdorf E (2016) The small RNA SraG participates in PNPase homeostasis. *RNA* **22:** 1560–1573

Franze de Fernandez MT, Eoyang L & August JT (1968) Factor fraction required for the synthesis of bacteriophage Qbeta-RNA. *Nature* **219:** 588–90

Frazão C, McVey CE, Amblar M, Barbas A, Vonrhein C, Arraiano CM & Carrondo MA (2006) Unravelling the dynamics of RNA degradation by ribonuclease II and its RNA-bound complex. *Nature* **443:** 110–4

Fröhlich KS, Förstner KU & Gitai Z (2018) Post-transcriptional gene regulation by an Hfq-independent small RNA in *Caulobacter crescentus*. *Nucleic Acids Res.* **46:** 10969–10982

Fröhlich KS, Papenfort K, Fekete A & Vogel J (2013) A small RNA activates CFA synthase by isoform-specific mRNA stabilization. *EMBO J.* **32:** 2963–79

Ge Z, Mehta P, Richards J & Karzai AW (2010) Non-stop mRNA decay initiates at the ribosome. *Mol. Microbiol.* **78:** 1159–70

Gibbs MR & Fredrick K (2018) Roles of elusive translational GTPases come to light and inform on the process of ribosome biogenesis in bacteria. *Mol. Microbiol.* **107:** 445–454

Godefroy T (1970) Kinetics of polymerization and phosphorolysis reactions of *Escherichia*

coli polynucleotide phosphorylase. Evidence for multiple binding of polynucleotide in phosphorolysis. *Eur. J. Biochem.* **14:** 222–31

Grunberg-Manago M (1963) Enzymatic synthesis of nucleic acids. *Prog. Biophys. Mol. Biol.* **13:** 175–239

Grunberg-Manago M, Oritz PJ & Ochoa S (1955) Enzymatic synthesis of nucleic acidlike polynucleotides. *Science* **122:** 907–10

Guo Q, Goto S, Chen Y, Feng B, Xu Y, Muto A, Himeno H, Deng H, Lei J & Gao N (2013) Dissecting the in vivo assembly of the 30S ribosomal subunit reveals the role of RimM and general features of the assembly process. *Nucleic Acids Res.* **41:** 2609–20

Gutgsell NS & Jain C (2012) Role of precursor sequences in the ordered maturation of *E. coli* 23S ribosomal RNA. *RNA* **18:** 345–53

Gyorfy Z, Draskovits G, Vernyik V, Blattner FF, Gaal T & Posfai G (2015) Engineered ribosomal RNA operon copy-number variants of *E. coli* reveal the evolutionary trade-offs shaping rRNA operon number. *Nucleic Acids Res.* **43:** 1783–1794

Hajnsdorf E & Boni I V (2012) Multiple activities of RNA-binding proteins S1 and Hfq. *Biochimie* **94:** 1544–53

Hajnsdorf E & 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:** 1501–5

Hammarlöf DL, Bergman JM, Garmendia E & Hughes D (2015) Turnover of mRNAs is one of the essential functions of RNase E. *Mol. Microbiol.* **98:** 34–45

Hardwick SW, Chan VSY, Broadhurst RW & Luisi BF (2011) An RNA degradosome assembly in *Caulobacter crescentus*. *Nucleic Acids Res.* **39:** 1449–59

Hardwick SW, Gubbey T, Hug I, Jenal U & Luisi BF (2012) Crystal structure of *Caulobacter crescentus* polynucleotide phosphorylase reveals a mechanism of RNA substrate channelling and RNA degradosome assembly. *Open Biol.* **2:** 120028–120028

de Haseth PL & Uhlenbeck OC (1980) Interaction of *Escherichia coli* host factor protein with Q beta ribonucleic acid. *Biochemistry* **19:** 6146–51

Hayakawa H, Kuwano M & Sekiguchi M (2001) Specific binding of 8-oxoguanine-containing RNA to polynucleotide phosphorylase protein. *Biochemistry* **40:** 9977–82

Herold M & Nierhaus KH (1987) Incorporation of six additional proteins to complete the

assembly map of the 50 S subunit from *Escherichia coli* ribosomes. *J. Biol. Chem.* **262:** 8826–33

Hossain ST, Malhotra A & Deutscher MP (2015) The Helicase Activity of Ribonuclease R Is Essential for Efficient Nuclease Activity. *J. Biol. Chem.* **290:** 15697–706

Hossain ST, Malhotra A & Deutscher MP (2016) How RNase R Degrades Structured RNA: ROLE OF THE HELICASE ACTIVITY AND THE S1 DOMAIN. *J. Biol. Chem.* **291:** 7877–87

Ikeda Y, Yagi M, Morita T & Aiba H (2011) Hfq binding at RhlB-recognition region of RNase E is crucial for the rapid degradation of target mRNAs mediated by sRNAs in *Escherichia coli*. *Mol. Microbiol.* **79:** 419–432

Jacob AI, Köhrer C, Davies BW, RajBhandary UL & Walker GC (2013) Conserved bacterial RNase YbeY plays key roles in 70S ribosome quality control and 16S rRNA maturation. *Mol. Cell* **49:** 427–38

Jahn N, Preis H, Wiedemann C & Brantl S (2012) BsrG/SR4 from *Bacillus subtilis*--the first temperature-dependent type I toxin-antitoxin system. *Mol. Microbiol.* **83:** 579–98

Janssen BD, Garza-Sánchez F & Hayes CS (2013) A-site mRNA cleavage is not required for tmRNA-mediated ssrA-peptide tagging. *PLoS One* **8:** e81319

Janssen BD & Hayes CS (2012) The tmRNA ribosome rescue system. *Adv. Protein Chem. Struct. Biol.* **86:** 151–91

Jarrige A, Bréchemier-Baey D, Mathy N, Duché O & Portier C (2002) Mutational analysis of polynucleotide phosphorylase from *Escherichia coli*. *J. Mol. Biol.* **321:** 397–409

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

Jiang K, Zhang C, Guttula D, Liu F, van Kan JA, Lavelle C, Kubiak K, Malabirade A, Lapp A, Arluison V & van der Maarel JRC (2015) Effects of Hfq on the conformation and compaction of DNA. *Nucleic Acids Res.* **43:** 4332–41

Jiang M, Datta K, Walker A, Strahler J, Bagamasbad P, Andrews PC & Maddock JR (2006) The *Escherichia coli* GTPase CgtAE is involved in late steps of large ribosome assembly. *J. Bacteriol.* **188:** 6757–70

Jones GH & Mackie GA (2013) Streptomyces coelicolor Polynucleotide phosphorylase can polymerize nucleoside diphosphates under phosphorolysis conditions, with

Implications for the degradation of structured RNAs. *J. Bacteriol.* **195:** 5151–5159

Jones PG & Inouye M (1996) RbfA, a 30S ribosomal binding factor, is a cold-shock protein whose absence triggers the cold-shock response. *Mol. Microbiol.* **21:** 1207–18

Jones PG, VanBogelen RA & Neidhardt FC (1987) Induction of proteins in response to low temperature in *Escherichia coli*. *J. Bacteriol.* **169:** 2092–5

Kaczanowska M & Rydén-Aulin M (2007) Ribosome biogenesis and the translation process in *Escherichia coli*. *Microbiol. Mol. Biol. Rev.* **71:** 477–94

Kaplan R & Apirion D (1975) The fate of ribosomes in *Escherichia coli* cells starved for a carbon source. *J. Biol. Chem.* **250:** 1854–63

Karzai AW & Sauer RT (2001) Protein factors associated with the SsrA.SmpB tagging and ribosome rescue complex. *Proc. Natl. Acad. Sci. U. S. A.* **98:** 3040–4

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

Khidr L, Wu G, Davila A, Procaccio V, Wallace D & Lee WH (2008) Role of SUV3 helicase in maintaining mitochondrial homeostasis in human cells. *J. Biol. Chem.* **283:** 27064–27073

Kurylo CM, Parks MM, Juette MF, Zinshteyn B, Altman RB, Thibado JK, Vincent CT & Blanchard SC (2018) Endogenous rRNA sequence variation can regulate stress response gene expression and phenotype. *Cell Rep.* **25:** 236–248.e6

Łabno A, Warkocki Z, Kuliński T, Krawczyk PS, Bijata K, Tomecki R & Dziembowski A (2016) Perlman syndrome nuclease DIS3L2 controls cytoplasmic non-coding RNAs and provides surveillance pathway for maturing snRNAs. *Nucleic Acids Res.* **44:** 10437–10453

Lalaouna D, Carrier M-C, Semsey S, Brouard J-S, Wang J, Wade JTT & Massé E (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

Laursen BS, Sørensen HP, Mortensen KK & Sperling-Petersen HU (2005) Initiation of protein synthesis in bacteria. *Microbiol. Mol. Biol. Rev.* **69:** 101–23

De Lay N & Gottesman S (2011) Role of polynucleotide phosphorylase in sRNA function in *Escherichia coli*. *RNA* **17:** 1172–1189

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

Lebreton A, Tomecki R, Dziembowski A & Séraphin B (2008) Endonucleolytic RNA cleavage by a eukaryotic exosome. *Nature* **456:** 993–996

Lehnik-Habrink M, Pförtner H, Rempeters L, Pietack N, Herzberg C & Stülke J (2010) The RNA degradosome in *Bacillus subtilis*: identification of CshA as the major RNA helicase in the multiprotein complex. *Mol. Microbiol.* **77:** 958–71

Leong V, Kent M, Jomaa A & Ortega J (2013) *Escherichia coli rimM* and *yjeQ* null strains accumulate immature 30S subunits of similar structure and protein complement. *RNA* **19:** 789–802

Li Z & Deutscher MP (1994) The role of individual exoribonucleases in processing at the 3' end of *Escherichia coli* tRNA precursors. *J. Biol. Chem.* **269:** 6064–71

Li Z & Deutscher MP (1995) The tRNA processing enzyme RNase T is essential for maturation of 5S RNA. *Proc. Natl. Acad. Sci. U. S. A.* **92:** 6883–6

Li Z & Deutscher MP (1996) Maturation pathways for *E. coli* tRNA precursors: a random multienzyme process in vivo. *Cell* **86:** 503–12

Li Z, Pandit S & Deutscher MP (1998) 3' exoribonucleolytic trimming is a common feature of the maturation of small, stable RNAs in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **95:** 2856–61

Li Z, Pandit S & Deutscher MP (1999a) Maturation of 23S ribosomal RNA requires the exoribonuclease RNase T. *RNA* **5:** 139–46

Li Z, Pandit S & Deutscher MP (1999b) RNase G (CafA protein) and RNase E are both required for the 5' maturation of 16S ribosomal RNA. *EMBO J.* **18:** 2878–85

Li Z, Reimers S, Pandit S & Deutscher MP (2002) RNA quality control: degradation of defective transfer RNA. *EMBO J.* **21:** 1132–1138

Liang W & Deutscher MP (2012a) Post-translational modification of RNase R is regulated by stress-dependent reduction in the acetylating enzyme Pka (YfiQ). *RNA* **18:** 37–41

Liang W & Deutscher MP (2012b) Transfer-messenger RNA-SmpB protein regulates ribonuclease R turnover by promoting binding of HslUV and Lon proteases. *J. Biol. Chem.* **287:** 33472–9

Liang W & Deutscher MP (2013) Ribosomes regulate the stability and action of the exoribonuclease RNase R. *J. Biol. Chem.* **288:** 34791–8

Liang W, Malhotra A & Deutscher MP (2011) Acetylation regulates the stability of a bacterial protein: growth stage-dependent modification of RNase R. *Mol. Cell* **44:** 160–6

Lin P-H & Lin-Chao S (2005) RhlB helicase rather than enolase is the beta-subunit of the *Escherichia coli* polynucleotide phosphorylase (PNPase)-exoribonucleolytic complex. *Proc. Natl. Acad. Sci. U. S. A.* **102:** 16590–5

Link TM, Valentin-Hansen P & Brennan RG (2009) Structure of *Escherichia coli* Hfq bound to polyriboadenylate RNA. *Proc. Natl. Acad. Sci. U. S. A.* **106:** 19292–7

Liou G-G, Chang H-Y, Lin C-S & Lin-Chao S (2002) DEAD box RhlB RNA helicase physically associates with exoribonuclease PNPase to degrade double-stranded RNA independent of the degradosome-assembling region of RNase E. *J. Biol. Chem.* **277:** 41157–62

Liu B, Deikus G, Bree A, Durand S, Kearns DB & Bechhofer DH (2014) Global analysis of mRNA decay intermediates in *Bacillus subtilis* wild-type and polynucleotide phosphorylase-deletion strains. *Mol. Microbiol.* **94:** 41–55

Liu B, Kearns DB & Bechhofer DH (2016) Expression of multiple *Bacillus subtilis* genes is controlled by decay of slrA mRNA from Rho-dependent 3' ends. *Nucleic Acids Res.* **44:** 3364–3372

Liu Q, Greimann JC & Lima CD (2006) Reconstitution, Activities, and structure of the eukaryotic RNA exosome. *Cell* **127:** 1223–1237

Lorentzen E, Basquin J, Tomecki R, Dziembowski A & Conti E (2008) Structure of the active subunit of the yeast exosome core, Rrp44: diverse modes of substrate recruitment in the RNase II nuclease family. *Mol. Cell* **29:** 717–28

Lu F & Taghbalout A (2013) Membrane association via an amino-terminal amphipathic helix is required for the cellular organization and function of RNase II. *J. Biol. Chem.* **288:** 7241–51

Lu F & Taghbalout A (2014) The *Escherichia coli* major exoribonuclease RNase II is a component of the RNA degradosome. *Biosci. Rep.* **34:** e00166

Mackie GA (2013) RNase E: at the interface of bacterial RNA processing and decay. *Nat. Rev. Microbiol.* **11:** 45–57

Maes A, Gracia C, Hajnsdorf E & Régnier P (2012) Search for poly(A) polymerase targets in *E. coli* reveals its implication in surveillance of Glu tRNA processing and degradation of stable RNAs. *Mol. Microbiol.* **83:** 436–451

Malabirade A, Partouche D, El Hamoui O, Turbant F, Geinguenaud F, Recouvreux P, Bizien T, Busi F, Wien F & Arluison V (2018) Revised role for Hfq bacterial regulator on DNA topology. *Sci. Rep.* **8:** 16792

Marujo PE, Hajnsdorf E, Le Derout J, Andrade R, Arraiano CM, Regnier P & Régnier P (2000) RNase II removes the oligo(A) tails that destabilize the *rpsO* mRNA of *Escherichia coli*. *RNA* **6:** 1185–1193

Mathy N, Jarrige AC, Robert-Le Meur M & Portier C (2001) Increased expression of *Escherichia coli* polynucleotide phosphorylase at low temperatures is linked to a decrease in the efficiency of autocontrol. *J. Bacteriol.* **183:** 3848–54

Matos RG, Barbas A & Arraiano CM (2009) RNase R mutants elucidate the catalysis of structured RNA: RNA-binding domains select the RNAs targeted for degradation. *Biochem. J.* **423:** 291–301

Mattaj IW (1993) RNA recognition: a family matter? *Cell* **73:** 837–40

Matus-Ortega ME, Regonesi ME, Piña-Escobedo A, Tortora P, Dehò G & García-Mena J (2007) The KH and S1 domains of *Escherichia coli* polynucleotide phosphorylase are necessary for autoregulation and growth at low temperature. *Biochim. Biophys. Acta* **1769:** 194–203

Miczak A, Kaberdin VR, Wei CL & Lin-Chao S (1996) Proteins associated with RNase E in a multicomponent ribonucleolytic complex. *Proc. Natl. Acad. Sci. U. S. A.* **93:** 3865–9

Mitchell P, Petfalski E, Shevchenko A, Mann M & Tollervey D (1997) The exosome: A conserved eukaryotic RNA processing complex containing multiple 3'-5' exoribonucleases. *Cell* **91:** 457–466

Mizushima S & Nomura M (1970) Assembly Mapping of 30S Ribosomal Proteins from *E. coli*. *Nature* **226:** 1214–1218

Mohanty BK & Kushner SR (2000) Polynucleotide phosphorylase functions both as a 3'- 5'

exonuclease and a poly(A) polymerase in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **97:** 11966–71

Mohanty BK & Kushner SR (2002) Polyadenylation of *Escherichia coli* transcripts plays an integral role in regulating intracellular levels of polynucleotide phosphorylase and RNase E. *Mol. Microbiol.* **45:** 1315–24

Mohanty BK & Kushner SR (2003) Genomic analysis in *Escherichia coli* demonstrates differential roles for polynucleotide phosphorylase and RNase II in mRNA abundance and decay. *Mol. Microbiol.* **50:** 645–658

Mohanty BK & Kushner SR (2006) The majority of *Escherichia coli* mRNAs undergo post-transcriptional modification in exponentially growing cells. *Nucleic Acids Res.* **34:** 5695–704

Mohanty BK & Kushner SR (2010) Processing of the *Escherichia coli leuX* tRNA transcript, encoding tRNA(Leu5), requires either the 3'-5' exoribonuclease polynucleotide phosphorylase or RNase P to remove the Rho-independent transcription terminator. *Nucleic Acids Res.* **38:** 597–607

Mohanty BK & Kushner SR (2016) Regulation of mRNA decay in bacteria. *Annu. Rev. Microbiol.* **70:** 25–44

Mohanty BK & Kushner SR (2018) Enzymes involved in posttranscriptional RNA metabolism in Gram-negative bacteria. *Microbiol. Spectr.* **6:** 1–16

Mohanty BK, Maples VF & Kushner SR (2004) The Sm-like protein Hfq regulates polyadenylation dependent mRNA decay in *Escherichia coli*. *Mol. Microbiol.* **54:** 905–20

Mohanty BK, Petree JR & Kushner SR (2016) Endonucleolytic cleavages by RNase E generate the mature 3' termini of the three proline tRNAs in *Escherichia coli*. *Nucleic Acids Res.* **44:** 6350–6362

Morita T & 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:** 294–8

Morita T & Aiba H (2018) Mechanism and physiological significance of autoregulation of the *Escherichia coli hfq* gene. *RNA* rna.068106.118

Morita T, Maki K & Aiba H (2005) RNase E-based ribonucleoprotein complexes:

mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs. *Genes Dev.* **19:** 2176–86

Muffler A, Fischer D & Hengge-Aronis R (1996) The RNA-binding protein HF-I, known as a host factor for phage Qbeta RNA replication, is essential for *rpoS* translation in *Escherichia coli*. *Genes Dev.* **10:** 1143–51

Neidhardt FC (Frederick FC, Ingraham JL & Schaechter M) (1990) Physiology of the bacterial cell: a molecular approach. Sinauer Associates

Newman JA, Hewitt L, Rodrigues C, Solovyova AS, Harwood CR & Lewis RJ (2012) Dissection of the network of interactions that links RNA processing with glycolysis in the *Bacillus subtilis* degradosome. *J. Mol. Biol.* **416:** 121–36

Nierhaus KH (2006) Bacterial Ribosomes: Assembly. In *eLS* Chichester, UK: John Wiley & Sons, Ltd

Nierhaus KH & Dohme F (1974) Total Reconstitution of Functionally Active 50S Ribosomal Subunits from *Escherichia coli*. *Proc. Natl. Acad. Sci.* **71:** 4713–4717

Nomura M (1999) Regulation of ribosome biosynthesis in *Escherichia coli* and *Saccharomyces cerevisiae*: diversity and common principles. *J. Bacteriol.* **181:** 6857–64

Nurmohamed S, Vaidialingam B, Callaghan AJ & Luisi BF (2009) Crystal structure of *Escherichia coli* polynucleotide phosphorylase core bound to RNase E, RNA and manganese: implications for catalytic mechanism and RNA degradosome assembly. *J. Mol. Biol.* **389:** 17–33

Nurmohamed S, Vincent HA, Titman CM, Chandran V, Pears MR, Du D, Griffin JL, Callaghan AJ & Luisi BF (2011) Polynucleotide phosphorylase activity may be modulated by metabolites in *Escherichia coli*. *J. Biol. Chem.* **286:** 14315–14323

Ogle JM & Ramakrishnan V (2005) Structural insights into translational fidelity. *Annu. Rev. Biochem.* **74:** 129–77

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

Panja S, Schu DJ & Woodson SA (2013) Conserved arginines on the rim of Hfq catalyze base pair formation and exchange. *Nucleic Acids Res.* **41:** 7536–46

Papenfort K & Vogel J (2009) Multiple target regulation by small noncoding RNAs rewires gene expression at the post-transcriptional level. *Res. Microbiol.* **160:** 278–87

Park H, Yakhnin H, Connolly M, Romeo T & Babitzke P (2015) CsrA participates in a PNPase autoregulatory mechanism by selectively repressing translation of *pnp* transcripts that have been previously processed by RNase III and PNPase. *J. Bacteriol.* **197:** 3751–3759

Peil L, Virumäe K & Remme J (2008) Ribosome assembly in *Escherichia coli* strains lacking the RNA helicase DeaD/CsdA or DbpA. *FEBS J.* **275:** 3772–82

Phadtare S (2012) *Escherichia coli* cold-shock gene profiles in response to over-expression/deletion of CsdA, RNase R and PNPase and relevance to low-temperature RNA metabolism. *Genes Cells* **17:** 850–74

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

Piazza F, Zappone M, Sana M, Briani F & Dehò G (1996) Polynucleotide phosphorylase of *Escherichia coli* is required for the establishment of bacteriophage P4 immunity. *J. Bacteriol.* **178:** 5513–21

Pobre V & Arraiano CM (2015) Next generation sequencing analysis reveals that the ribonucleases RNase II, RNase R and PNPase affect bacterial motility and biofilm formation in *E. coli*. *BMC Genomics* **16:** 72

Portier C (1975) Quaternary structure of polynucleotide phosphorylase from *Escherichia coli*: evidence of a complex between two types of polypeptide chains. *Eur. J. Biochem.* **55:** 573–82

Portier C & Regnier P (1984) Expression of the *rpsO* and *pnp* genes: structural analysis of a DNA fragment carrying their control regions. *Nucleic Acids Res.* **12:** 6091–102

Portnoy V, Evguenieva-Hackenberg E, Klein F, Walter P, Lorentzen E, Klug G & Schuster G (2005) RNA polyadenylation in Archaea: not observed in *Haloferax* while the exosome polynucleotidylates RNA in *Sulfolobus*. *EMBO Rep.* **6:** 1188–93

Prud'homme-Généreux A, Beran RK, Iost I, Ramey CS, Mackie GA & Simons RW (2004) Physical and functional interactions among RNase E, polynucleotide phosphorylase and the cold-shock protein, CsdA: evidence for a 'cold shock degradosome'. *Mol.*

Microbiol. **54:** 1409–21

Purusharth RI, Madhuri B & Ray MK (2007) Exoribonuclease R in *Pseudomonas syringae* is essential for growth at low temperature and plays a novel role in the 3' end processing of 16 and 5S ribosomal RNA. *J. Biol. Chem.* **282:** 16267–77

Py B, Higgins CF, Krisch HM & Carpousis AJ (1996) A DEAD-box RNA helicase in the *Escherichia coli* RNA degradosome. *Nature* **381:** 169–72

Rabhi M, Espéli O, Schwartz A, Cayrol B, Rahmouni AR, Arluison V & Boudvillain M (2011) The Sm-like RNA chaperone Hfq mediates transcription antitermination at Rho-dependent terminators. *EMBO J.* **30:** 2805–2816

Raj R, Mitra S & Gopal B (2018) Characterization of *Staphylococcus epidermidis* Polynucleotide phosphorylase and its interactions with ribonucleases RNase J1 and RNase J2. *Biochem. Biophys. Res. Commun.* **495:** 2078–2084

Régnier P & 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:** 602–9

Reuven NB, Zhou Z & Deutscher MP (1997) Functional overlap of tRNA nucleotidyltransferase, Poly(A) polymerase I, and Polynucleotide phosphorylase. *J. Biol. Chem.* **272:** 33255–33259

Richards J, Mehta P & Karzai AW (2006) RNase R degrades non-stop mRNAs selectively in an SmpB-tmRNA-dependent manner. *Mol. Microbiol.* **62:** 1700–12

Robert-Le Meur M & Portier C (1992) E.coli polynucleotide phosphorylase expression is autoregulated through an RNase III-dependent mechanism. *EMBO J.* **11:** 2633–41

Robert-Le Meur M & Portier C (1994) Polynucleotide phosphorylase of *Escherichia coli* induces the degradation of its RNase III processed messenger by preventing its translation. *Nucleic Acids Res.* **22:** 397–403

Rochat T, Delumeau O, Figueroa-Bossi N, Noirot P, Bossi L, Dervyn E & Bouloc P (2015) Tracking the Elusive Function of *Bacillus subtilis* Hfq. *PLoS One* **10:** e0124977

Rott R, Zipor G, Portnoy V, Liveanu V & Schuster G (2003) RNA polyadenylation and degradation in cyanobacteria are similar to the chloroplast but different from *Escherichia coli*. *J. Biol. Chem.* **278:** 15771–7

Roux CM, DeMuth JP & Dunman PM (2011) Characterization of components of the *Staphylococcus aureus* mRNA degradosome holoenzyme-like complex. *J. Bacteriol.* **193:** 5520–6

Roy-Chaudhuri B, Kirthi N & Culver GM (2010) Appropriate maturation and folding of 16S rRNA during 30S subunit biogenesis are critical for translational fidelity. *Proc. Natl. Acad. Sci. U. S. A.* **107:** 4567–72

Roy MK, Singh B, Ray BK & Apirion D (1983) Maturation of 5-S rRNA: ribonuclease E cleavages and their dependence on precursor sequences. *Eur. J. Biochem.* **131:** 119–27

Russell JH & Keiler KC (2009) Subcellular localization of a bacterial regulatory RNA. *Proc. Natl. Acad. Sci. U. S. A.* **106:** 16405–9

Salvo E, Alabi S, Liu B, Schlessinger A & Bechhofer DH (2016) Interaction of *Bacillus subtilis* Polynucleotide Phosphorylase and RNase Y: STRUCTURAL MAPPING AND EFFECT ON mRNA TURNOVER. *J. Biol. Chem.* **291:** 6655–63

Santiago-Frangos A, Jeliazkov JR, Gray JJ & Woodson SA (2017) Acidic C-terminal domains autoregulate the RNA chaperone Hfq. *Elife* **6:**

Santiago-Frangos A & Woodson SA (2018) Hfq chaperone brings speed dating to bacterial sRNA. *Wiley Interdiscip. Rev. RNA* **9:** e1475

dos Santos RF, Quendera AP, Boavida S, Seixas AF, Arraiano CM & Andrade JM (2018) Major 3'–5' Exoribonucleases in the Metabolism of Coding and Non-coding RNA. In *Progress in Molecular Biology and Translational Science* pp 101–155. Academic Press

Sarkar D, Park ES, Emdad L, Randolph A, Valerie K & Fisher PB (2005) Defining the domains of human polynucleotide phosphorylase (hPNPaseOLD-35) mediating cellular senescence. *Mol. Cell. Biol.* **25:** 7333–43

Sauer E (2013) Structure and RNA-binding properties of the bacterial LSm protein Hfq. *RNA Biol.* **10:** 610–8

Schaeffer D, Tsanova B, Barbas A, Reis FP, Dastidar EG, Sanchez-Rotunno M, Arraiano CM & van Hoof A (2009) The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities. *Nat. Struct. Mol. Biol.* **16:** 56–62

Schmeing TM & Ramakrishnan V (2009) What recent ribosome structures have revealed about the mechanism of translation. *Nature* **461:** 1234–42

Schmier BJ, Seetharaman J, Deutscher MP, Hunt JF & Malhotra A (2012) The structure and enzymatic properties of a novel RNase II family enzyme from Deinococcus radiodurans. *J. Mol. Biol.* **415:** 547–59

Schneider C, Leung E, Brown J & Tollervey D (2009) The N-terminal PIN domain of the exosome subunit Rrp44 harbors endonuclease activity and tethers Rrp44 to the yeast core exosome. *Nucleic Acids Res.* **37:** 1127–40

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

Schuwirth B, Borovinskaya M & Hau C (2005) Structures of the bacterial ribosome at 3.5 Å resolution. *Science* **310:** 827–834

Sesto N, Touchon M, Andrade JM, Kondo J, Rocha EPC, Arraiano CM, Archambaud C, Westhof É, Romby P & Cossart P (2014) A PNPase dependent CRISPR system in *Listeria*. *PLoS Genet.* **10:** e1004065

Shajani Z, Sykes MT & Williamson JR (2011) Assembly of bacterial ribosomes. *Annu. Rev. Biochem.* **80:** 501–26

Sharma IM, Korman A & Woodson SA (2018) The Hfq chaperone helps the ribosome mature. *EMBO J.* **37:** e99616

Sharma MR, Barat C, Wilson DN, Booth TM, Kawazoe M, Hori-Takemoto C, Shirouzu M, Yokoyama S, Fucini P & Agrawal RK (2005) Interaction of Era with the 30S ribosomal subunit implications for 30S subunit assembly. *Mol. Cell* **18:** 319–29

Sharpe Elles LM, Sykes MT, Williamson JR & Uhlenbeck OC (2009) A dominant negative mutant of the *E. coli* RNA helicase DbpA blocks assembly of the 50S ribosomal subunit. *Nucleic Acids Res.* **37:** 6503–14

Shi Z, Yang W-Z, Lin-Chao S, Chak K-F & Yuan HS (2008) Crystal structure of *Escherichia coli* PNPase: central channel residues are involved in processive RNA degradation. *RNA* **14:** 2361–71

Slomovic S, Portnoy V, Yehudai-Resheff S, Bronshtein E & Schuster G (2008) Polynucleotide phosphorylase and the archaeal exosome as poly(A)-polymerases.

Biochim. Biophys. Acta **1779:** 247–55

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

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

Smith BA, Gupta N, Denny K & Culver GM (2018) Characterization of 16S rRNA Processing with Pre-30S Subunit Assembly Intermediates from *E. coli*. *J. Mol. Biol.* **430:** 1745–1759

Söderberg MA & Cianciotto NP (2010) Mediators of lipid A modification, RNA degradation, and central intermediary metabolism facilitate the growth of *Legionella pneumophila* at low temperatures. *Curr. Microbiol.* **60:** 59–65

Sohlberg B, Huang J & Cohen SN (2003) The Streptomyces coelicolor polynucleotide phosphorylase homologue, and not the putative poly(A) polymerase, can polyadenylate RNA. *J. Bacteriol.* **185:** 7273–8

Song L, Wang G, Malhotra A, Deutscher MP & Liang W (2016) Reversible acetylation on Lys501 regulates the activity of RNase II. *Nucleic Acids Res.* **44:** 1979–1988

Song W-S, Lee M & Lee K (2011) RNase G participates in processing of the 5'-end of 23S ribosomal RNA. *J. Microbiol.* **49:** 508–11

Soper T, Mandin P, Majdalani N, Gottesman S & Woodson SA (2010) Positive regulation by small RNAs and the role of Hfq. *Proc. Natl. Acad. Sci.* **107:** 9602–9607

Spickler C & Mackie GA (2000) Action of RNase II and polynucleotide phosphorylase against RNAs containing stem-loops of defined structure. *J. Bacteriol.* **182:** 2422–7

Staals RHJ, Bronkhorst AW, Schilders G, Slomovic S, Schuster G, Heck AJR, Raijmakers R & Pruijn GJM (2010) Dis3-like 1: A novel exoribonuclease associated with the human exosome. *EMBO J.* **29:** 2358–2367

Steitz TA & Moore PB (2003) RNA, the first macromolecular catalyst: the ribosome is a ribozyme. *Trends Biochem. Sci.* **28:** 411–8

Stickney LM, Hankins JS, Miao X & Mackie GA (2005) Function of the conserved S1 and KH domains in Polynucleotide phosphorylase. *J. Bacteriol.* **187:** 7214–7221

Stone CM, Butt LE, Bufton JC, Lourenco DC, Gowers DM, Pickford AR, Cox PA, Vincent HA & Callaghan AJ (2017) Inhibition of homologous phosphorolytic ribonucleases by citrate may represent an evolutionarily conserved communicative link between RNA degradation and central metabolism. *Nucleic Acids Res.* **45:** 4655–4666

Storz G, Vogel J & Wassarman KM (2011) Regulation by small RNAs in bacteria: expanding frontiers. *Mol. Cell* **43:** 880–91

Strader MB, Hervey WJ, Costantino N, Fujigaki S, Chen CY, Akal-Strader A, Ihunnah CA, Makusky AJ, Court DL, Markey SP & Kowalak JA (2013) A coordinated proteomic approach for identifying proteins that interact with the *E. coli* ribosomal protein S12. *J. Proteome Res.* **12:** 1289–99

Sukhodolets M V. & Garges S (2003) Interaction of *Escherichia coli* RNA polymerase with the ribosomal protein S1 and the Sm-like ATPase Hfq. *Biochemistry* **42:** 8022–34

Sulewski M, Marchese-Ragona SP, Johnson KA & Benkovic SJ (1989) Mechanism of Polynucleotide phosphorylase. *Biochemistry* **28:** 5855–64

Sulthana S & Deutscher MP (2013) Multiple exoribonucleases catalyze maturation of the 3' terminus of 16S ribosomal RNA (rRNA). *J. Biol. Chem.* **288:** 12574–9

Sulthana S, Quesada E & Deutscher MP (2017) RNase II regulates RNase PH and is essential for cell survival during starvation and stationary phase. *Rna* **23:** rna.060558.116

Suzuki H & Tsukahara T (2014) A view of pre-mRNA splicing from RNase R resistant RNAs. *Int. J. Mol. Sci.* **15:** 9331–42

Suzuki H, Zuo Y, Wang J, Zhang MQ, Malhotra A & Mayeda A (2006) Characterization of RNase R-digested cellular RNA source that consists of lariat and circular RNAs from pre-mRNA splicing. *Nucleic Acids Res.* **34:** e63

Symmons MF, Jones GH & Luisi BF (2000) A duplicated fold is the structural basis for polynucleotide phosphorylase catalytic activity, processivity, and regulation. *Structure* **8:** 1215–26

Symmons MF, Williams MG, Luisi BF, Jones GH & Carpousis AJ (2002) Running rings around RNA: a superfamily of phosphate-dependent RNases. *Trends Biochem. Sci.* **27:** 11–8

Thurlow B, Davis JH, Leong V, F Moraes T, Williamson JR & Ortega J (2016) Binding properties of YjeQ (RsgA), RbfA, RimM and Era to assembly intermediates of the 30S

subunit. *Nucleic Acids Res.* **44:** 9918–9932

Tomecki R, Kristiansen MS, Lykke-Andersen S, Chlebowski A, Larsen KM, Szczesny RJ, Drazkowska K, Pastula A, Andersen JS, Stepien PP, Dziembowski A & Jensen TH (2010) The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L. *EMBO J.* **29:** 2342–2357

Traub P & Nomura M (1968) Structure and function of *E. coli* ribosomes. V. Reconstitution of functionally active 30S ribosomal particles from RNA and proteins. *Proc. Natl. Acad. Sci. U. S. A.* **59:** 777–84

Tseng Y-T, Chiou N-T, Gogiraju R & Lin-Chao S (2015) The protein interaction of RNA helicase B (RhlB) and Polynucleotide phosphorylase (PNPase) contributes to the homeostatic control of cysteine in *Escherichia coli*. *J. Biol. Chem.* **290:** 29953–29963

Tsui HC, Leung HC & Winkler ME (1994) Characterization of broadly pleiotropic phenotypes caused by an hfq insertion mutation in *Escherichia coli* K-12. *Mol. Microbiol.* **13:** 35–49

Tuckerman JR, Gonzalez G & Gilles-Gonzalez M-A (2011) Cyclic di-GMP activation of Polynucleotide phosphorylase signal-dependent RNA Processing. *J. Mol. Biol.* **407:** 633–639

Updegrove TB, Correia JJ, Galletto R, Bujalowski W & Wartell RM (2010) *E. coli* DNA associated with isolated Hfq interacts with Hfq's distal surface and C-terminal domain. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1799:** 588–596

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

Vakulskas CA, Potts AH, Babitzke P, Ahmer BMM & Romeo T (2015) Regulation of bacterial virulence by Csr (Rsm) systems. *Microbiol. Mol. Biol. Rev.* **79:** 193–224

Vanzo NF, Li YS, Py B, Blum E, Higgins CF, Raynal LC, Krisch HM & Carpousis AJ (1998) Ribonuclease E organizes the protein interactions in the *Escherichia coli* RNA degradosome. *Genes Dev.* **12:** 2770–81

Vecerek B, Moll I & Bläsi U (2005) Translational autocontrol of the *Escherichia coli* hfq RNA chaperone gene. *RNA* **11:** 976–84

Venkataraman K, Guja KE, Garcia-Diaz M & Karzai AW (2014) Non-stop mRNA decay: a

special attribute of trans-translation mediated ribosome rescue. *Front. Microbiol.* **5:** 93

Viegas SC, Pfeiffer V, Sittka A, Silva IJIJ, Vogel JJ & Arraiano CMCM (2007) Characterization of the role of ribonucleases in *Salmonella* small RNA decay. *Nucleic Acids Res.* **35:** 7651–64

Vincent HA & Deutscher MP (2006) Substrate recognition and catalysis by the exoribonuclease RNase R. *J. Biol. Chem.* **281:** 29769–75

Vincent HA & Deutscher MP (2009) Insights into how RNase R degrades structured RNA: analysis of the nuclease domain. *J. Mol. Biol.* **387:** 570–83

Vogel J & Luisi BF (2011) Hfq and its constellation of RNA. *Nat. Rev. Microbiol.* **9:** 578–89

Vuković L, Chipot C, Makino DL, Conti E & Schulten K (2016) Molecular mechanism of processive 3′ to 5′ RNA translocation in the active subunit of the RNA exosome complex. *J. Am. Chem. Soc.* **138:** 87–92

Wang G, Chen H-W, Oktay Y, Zhang J, Allen EL, Smith GM, Fan KC, Hong JS, French SW, McCaffery JM, Lightowlers RN, Morse HC, Koehler CM & Teitell MA (2010) PNPase regulates RNA import into mitochondria. *Cell* **142:** 456–67

Wang J, Rennie W, Liu C, Carmack CS, Prévost K, Caron M-P, Massé E, Ding Y & Wade JT (2015) Identification of bacterial sRNA regulatory targets using ribosome profiling. *Nucleic Acids Res.* **43:** 10308–20

Wang X, Wang C, Wu M, Tian T, Cheng T, Zhang X & Zang J (2017) Enolase binds to RnpA in competition with PNPase in *Staphylococcus aureus*. *FEBS Lett.* **591:** 3523–3535

Warner JR (2013) YbeY: the jealous tailor. *Mol. Cell* **49:** 422–3

Waters LS & Storz G (2009) Regulatory RNAs in bacteria. *Cell* **136:** 615–28

Wellner K, Betat H & Mörl M (2018) A tRNA's fate is decided at its 3′ end: Collaborative actions of CCA-adding enzyme and RNases involved in tRNA processing and degradation. *Biochim. Biophys. Acta* **1861:** 433–441

Williamson JR (2003) After the ribosome structures: How are the subunits assembled? *RNA* **9:** 165–167

Wilson DN & Nierhaus KH (2007) The weird and wonderful world of bacterial ribosome regulation. *Crit. Rev. Biochem. Mol. Biol.* **42:** 187–219

Wilusz CJ & Wilusz J (2013) Lsm proteins and Hfq: Life at the 3' end. *RNA Biol.* **10:** 592–601

Wimberly BT, Brodersen DE, Clemons WM, Morgan-Warren RJ, Carter AP, Vonrhein C, Hartsch T & Ramakrishnan V (2000) Structure of the 30S ribosomal subunit. *Nature* **407:** 327–39

Wong AG, McBurney KL, Thompson KJ, Stickney LM & Mackie GA (2013) S1 and KH domains of Polynucleotide phosphorylase determine the efficiency of RNA binding and autoregulation. *J. Bacteriol.* **195:** 2021–2031

Woodson SA (2008) RNA folding and ribosome assembly. *Curr. Opin. Chem. Biol.* **12:** 667–73

Woodson SA, Panja S & Santiago-Frangos A (2018) Proteins that chaperone RNA regulation. *Microbiol. Spectr.* **6:** 385–397

Wu J, Jiang Z, Liu M, Gong X, Wu S, Burns CM & Li Z (2009) Polynucleotide phosphorylase protects *Escherichia coli* against oxidative stress. *Biochemistry* **48:** 2012–20

Wurtmann EJ & Wolin SL (2010) A role for a bacterial ortholog of the Ro autoantigen in starvation-induced rRNA degradation. *Proc. Natl. Acad. Sci.* **107:** 4022–4027

Xia B, Ke H, Shinde U & Inouye M (2003) The role of RbfA in 16S rRNA processing and cell growth at low temperature in *Escherichia coli*. *J. Mol. Biol.* **332:** 575–84

Xu F & Cohen SN (1995) RNA degradation in *Escherichia coli* regulated by 3' adenylation and 5' phosphorylation. *Nature* **374:** 180–3

Yehudai-Resheff S, Hirsh M & Schuster G (2001) Polynucleotide phosphorylase functions as both an exonuclease and a poly(A) polymerase in spinach chloroplasts. *Mol. Cell. Biol.* **21:** 5408–16

Yehudai-Resheff S, Portnoy V, Yogev S, Adir N & Schuster G (2003) Domain analysis of the chloroplast polynucleotide phosphorylase reveals discrete functions in RNA degradation, polyadenylation, and sequence homology with exosome proteins. *Plant Cell* **15:** 2003–19

Yonath, A (2009) Large facilities and the evolving ribosome, the cellular machine for genetic-code translation. *J R Soc Interface* **6:** S575-S585

Yusupov MM, Yusupova GZ, Baucom A, Lieberman K, Earnest TN, Cate JH & Noller HF

(2001) Crystal structure of the ribosome at 5.5 A resolution. *Science* **292:** 883–96

Zangrossi S, Briani F, Ghisotti D, Regonesi ME, Tortora P & Dehò G (2000) Transcriptional and post-transcriptional control of polynucleotide phosphorylase during cold acclimation in *Escherichia coli*. *Mol. Microbiol.* **36:** 1470–80

Zhang A, Schu DJ, Tjaden BC, Storz G & Gottesman S (2013) Mutations in interaction surfaces differentially impact *E. coli* Hfq association with small RNAs and their mRNA targets. *J. Mol. Biol.* **425:** 3678–97

Zhang Y, Burkhardt DH, Rouskin S, Li G-W, Weissman JS & Gross CA (2018) A stress response that monitors and regulates mRNA structure is central to cold shock adaptation. *Mol. Cell* **70:** 274–286.e7

Zilhão R, Cairrão F, Régnier P, Arraiano CM, Zilhao R, Cairrao F, Regnier P & Arraiano CM (1996a) PNPase modulates RNase II expression in *Escherichia coli*: implications for mRNA decay and cell metabolism. *Mol. Microbiol.* **20:** 1033–42

Zilhão R, Plumbridge J, Hajnsdorf E, Régnier P & Arraiano CM (1996b) *Escherichia coli* RNase II: Characterization of the promoters involved in the transcription of rnb. *Microbiology* **142:** 367–375

Zilhão R, Régnier P & Maria Arraiano C (1995) The role of endonucleases in the expression of ribonuclease II in *Escherichia coli*. *FEMS Microbiol. Lett.* **130:** 237–244

Zundel MA, Basturea GN & Deutscher MP (2009) Initiation of ribosome degradation during starvation in *Escherichia coli*. *RNA* **15:** 977–83

Zuo Y & Deutscher MP (2001) Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. *Nucleic Acids Res.* **29:** 1017–26

Zuo Y, Vincent HA, Zhang J, Wang Y, Deutscher MP & Malhotra A (2006) Structural basis for processivity and single-strand specificity of RNase II. *Mol. Cell* **24:** 149–156

Chapter 2

The RNA-binding protein Hfq is important for ribosome biogenesis and affects translation fidelity

This chapter was based on:

Andrade JM*, **dos Santos RF***, Chelysheva I, Ignatova Z & Arraiano CM (2018) The RNA-binding protein Hfq is important for ribosome biogenesis and affects translation fidelity. *EMBO J.* **37:** e97631. DOI: 10.15252/embj.201797631
*Joint first authors

This article was highlighted in the News and Views section of the EMBO Journal issue: Sharma IM, Korman A & Woodson SA (2018) The Hfq chaperone helps the ribosome mature. *EMBO J.* **37:** e99616

For this chapter I performed experiments, analyzed data and revised the manuscript.

2. Chapter 2: The RNA-binding protein Hfq is important for ribosome biogenesis and affects translation fidelity

2.1. Abstract

Ribosome biogenesis is a complex process involving multiple factors. Here we show that the widely-conserved RNA chaperone Hfq, which can regulate sRNA-mRNA basepairing, plays a critical role in rRNA processing and ribosome assembly in *Escherichia coli*. Hfq binds the 17S rRNA precursor and facilitates its correct processing and folding to mature 16S rRNA. Hfq assists ribosome assembly and associates with pre-30S particles but not with mature 30S subunits. Inactivation of Hfq strikingly decreases the pool of mature 70S ribosomes. The reduction in ribosome levels depends on residues located in the distal face of Hfq but not on residues found in the proximal and rim surfaces which govern interactions with the sRNAs. Our results indicate that Hfq-mediated regulation of ribosomes is independent of its function as sRNA-regulator. Furthermore, we observed that inactivation of Hfq compromises translation efficiency and fidelity, both features of aberrantly assembled ribosomes. Our work expands the functions of the Sm-like protein Hfq beyond its function in small RNA-mediated regulation and unveils a novel role of Hfq as crucial in ribosome biogenesis and translation.

2.2. Introduction

Ribosomal RNA (rRNA) represents more than 80% of total RNA in the cell and along with a plethora of ribosomal proteins (r-proteins) constitutes the ribosome – the biosynthetic machinery of the cell. Ribosome biogenesis is a multi-step hierarchically ordered process in which processing of rRNA precursor (pre-rRNA) is a critical step. Emerging evidence suggests that pre-rRNA maturation serves as a quality control to guarantee the integrity of the functional ribosome. In *Escherichia coli*, RNase III is responsible for the initial cleavages that separate individual rRNA precursors, followed by subsequent 5' and 3' processing by multiple ribonucleases to generate the 16S, 23S and 5S rRNAs necessary to assemble the mature ribosomal subunits (Deutscher, 2009). Alterations in pre-rRNA processing cause conformational changes in the final rRNA and lead to aberrantly-assembled immature ribosomal particles with largely compromised translational accuracy (Liiv & Remme, 2004; Roy-Chaudhuri *et al*, 2010; Yang *et al*, 2014). A parallel can be drawn to eukaryotes since rRNA maturation errors lead to the production of defective ribosomal subunits (Cole *et al*, 2009; Fujii *et al*, 2012; Karbstein, 2013).

In prokaryotes, the small 30S ribosomal subunit contains 16S rRNA whereas 23S and 5S rRNA are the major components of the large 50S ribosomal subunit. The two asymmetric subunits include numerous r-proteins and associate to form the functionally active 70S ribosome (Shajani *et al*, 2011). Many auxiliary ribosome biogenesis factors, including GTPases, rRNA modification enzymes, helicases and other maturation factors, assist rRNA folding and r-protein assembly pathway (Davis & Williamson, 2017). Strikingly, mutations affecting many of these accessory proteins cause dysfunctional ribosomes. In humans, such mutations were shown to lead to severe diseases, collectively referred to as ribosomopathies (Narla & Ebert, 2010).

The bacterial RNA-binding protein Hfq is a member of the Sm/Lsm superfamily of proteins with homologues in all domains of life (Wilusz & Wilusz, 2013). Hfq is an RNA chaperone which facilitates basepairing between small regulatory RNAs (sRNAs) and their mRNA targets. Consequently, Hfq controls the expression of many mRNAs either positively or negatively (Vogel & Luisi, 2011; Hajnsdorf & Boni, 2012; Updegrove *et al*, 2016). Importantly, in many bacteria Hfq is not required for the sRNA-dependent

pathways (Christiansen *et al*, 2006; Rochat *et al*, 2015), suggesting other yet undefined function(s) of Hfq beyond regulation of sRNA activity.

Hfq interacts *in vitro* with the 16S rRNA (de Haseth & Uhlenbeck, 1980) although the functional role of this interaction has remained elusive. Furthermore, rRNA molecules are commonly found in Hfq-enriched co-immunoprecipitations, what is usually regarded as a background noise in transcriptomic studies (Zhang *et al*, 2003; Sittka *et al*, 2008; Bilusic *et al*, 2014). A crosslinking-based study in *E. coli* suggests interactions of Hfq with rRNA *in vivo* (Tree *et al*, 2014). An interaction between Hfq and S12 protein of the 30S ribosome subunit has been previously reported, yet lacking mechanistic details on its role (Strader *et al*, 2013). Clearly, Hfq interacts with rRNA but is this a functional or redundant interaction?

Here we identify a novel role of Hfq in ribosome biogenesis. Inactivation of Hfq leads to accumulation of 17S rRNA and reduced levels of 70S ribosomes in *E. coli*. Using *in vivo* and *in vitro* approaches, including ribosome profiling, we demonstrate that Hfq deletion affects the ribosome pool with direct effects on translation efficiency and fidelity. Our data propose Hfq as a novel auxiliary ribosome biogenesis factor. This expands the functional spectrum of this RNA chaperone beyond the sRNA-biology with impact on rRNA processing, ribosome biogenesis and translation fidelity.

2.3. Results

2.3.1. Hfq is required for the correct maturation and folding of the 16S rRNA

The matured 16S rRNA is carved out of a longer precursor RNA, the 17S rRNA precursor, which harbors extra nucleotides (nts) at both extremities (Figure 10A). Hfq is a pleiotropic regulator that impacts gene expression during both exponential and stationary phase (Tsui *et al*, 1994; Muffler *et al*, 1996; De Lay *et al*, 2013). Therefore, we compared the total RNA from wild-type and *hfq* cells extracted from exponential and stationary phase cells by Northern blotting using specific probes complementary to 5'- or 3'-ends of the 17S precursor rRNA. In addition, we used probes corresponding to the internal regions of 16S rRNA or 23S rRNA for control purposes (Figure 10B). Both 17S-specific probes hybridized only to 17S rRNA, whereas the 16S-probe identified both 16S and 17S rRNA.

Notably, inactivation of Hfq in both growth phases resulted in higher levels of 17S rRNA with misprocessed extremities (28% and 148% increase in exponential and stationary phase, respectively), suggesting a role for Hfq in 16S rRNA maturation. The accumulation of 17S in the Δ*hfq* mutant was additionally tested during different points of the growth curve, revealing that it occurs over time and irrespectively of the growth stage (Figure 11A).

Hfq comprises different RNA contact surfaces that have distinct RNA binding preferences. Specifically, Hfq can bind to an (ARN)x motif present in RNAs (Mikulecky *et al*, 2004; Link *et al*, 2009; Peng *et al*, 2014). Strikingly, both 5' and 3' extra nucleotide sequences of the 17S rRNA contain several of these predicted Hfq-binding sites (Figure 11B). Hence, in order to test if Hfq could in fact bind to these 17S rRNA flanking regions, we performed gel mobility shift experiments with constant amounts of the 5'-end and 3'-end sequences and increasing amounts of the Hfq protein (Figure 10C). Indeed, Hfq complexed with both flanking sequences corroborating that Hfq has the ability to bind *in vitro* to the 17S rRNA.

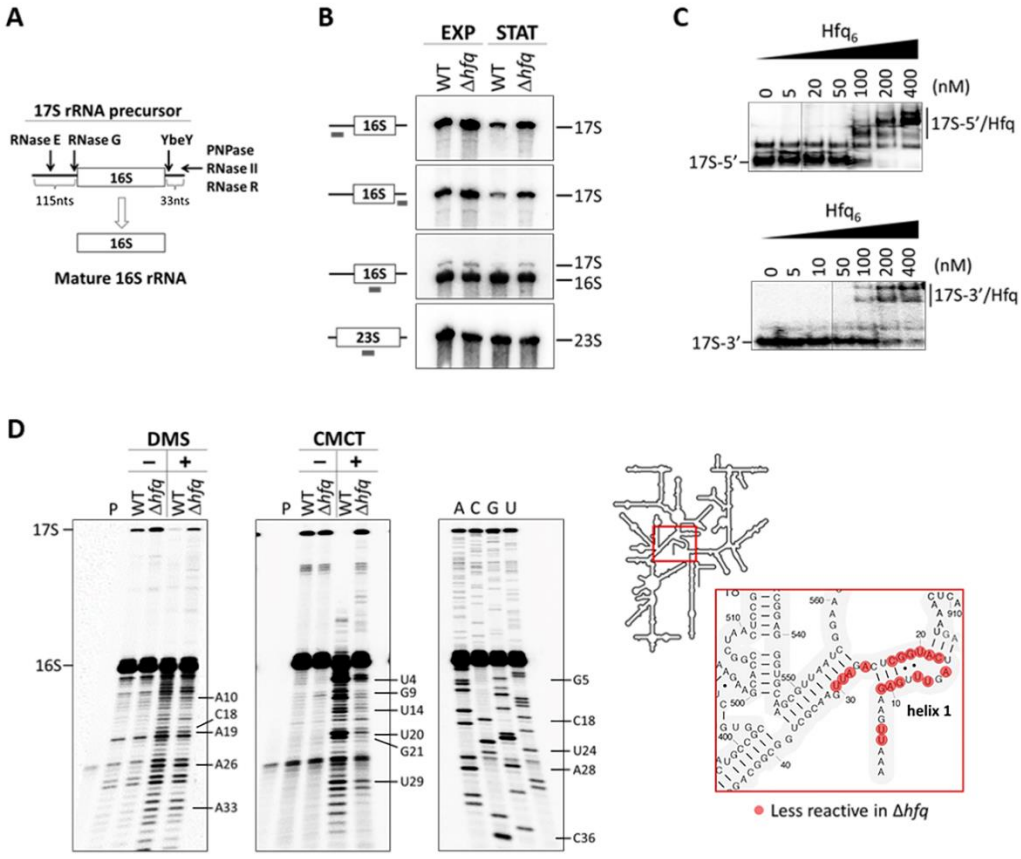


Figure 10 – Hfq is required for correct processing and folding of 16S rRNA. (A) Schematic representation of the RNase-mediated processing of the 17S rRNA precursor into mature 16S rRNA. (B) Northern blot analysis of total RNA extracted from cells in exponential (EXP) or stationary (STAT) growth phase. Samples were fractionated on a 4% polyacrylamide/7M Urea gel. A scheme of the probes binding to the rRNA sequence is displayed on the side. (C) Electrophoretic mobility shift assays of Hfq binding to the 5'- and 3'- extremities of the 17S rRNA. Increasing amounts of Hfq hexamer were mixed with a constant amount of the specific 17S-flanking sequences and resolved on a 6% (top panel) or 8% (bottom panel) native polyacrylamide gel. (D) DMS and CMCT accessibility probing of the 16S rRNA. Reverse-transcribed cDNA was fractionated on an 10% polyacrylamide/7M urea gel. Residues with altered reactivities in the Δ*hfq* mutant are indicated. The inset depicts the analyzed region of the 16S rRNA.

The additional 5'-end nucleotides present in the 17S rRNA could destabilase the folding of helix 1 and helix 2 of the mature 16S rRNA inducing alternative structures which would affect the formation of the central pseudoknot (Lodmell & Dahlberg, 1997; Roy-Chaudhuri *et al*, 2010). Furthermore, as an RNA chaperone Hfq can bind, melt and remodel RNA secondary structures (Woodson, 2008). Thus, we reasoned that the accumulation of precursor 17S rRNA observed in the absence of Hfq could affect the

correct folding of the 16S rRNA. To test this, we performed RNA mapping experiments using two distinct and complementary chemical probes: dimethyl sulfate (DMS) – that reacts with adenosines and cytidines – and N-cyclohexyl-N'-(2-morpholinoethyl)carbodiimide (CMCT) – that labels uridines and guanosines (Figure 10D). A specific antisense primer near the 5'-end of the 16S rRNA was used in the primer extension reactions which allowed good resolution of the 16S central pseudoknot, that consists of helix 1 (nucleotides 9–13/21–25) and helix 2 (nucleotides 17–19/916–918) (Brink *et al*, 1993). Several nucleotides accessible to DMS or CMCT modification in the wild-type, were less reactive to these probes in the absence of Hfq (Figure 10D). Our data imply that the folding of the 16S rRNA is altered as consequence of Hfq inactivation, resulting in the structural occlusion of those residues. Altogether, our observations indicate that Hfq interacts with 17S rRNA and is necessary for the correct processing and folding of the mature 16S rRNA, affecting the formation of the central pseudoknot.

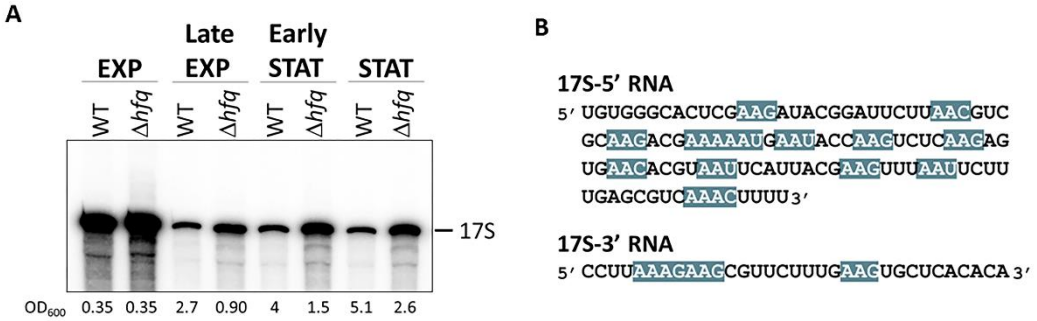


Figure 11 – Hfq regulates 17S rRNA levels. (A) 17S rRNA accumulates over time in the Δ*hfq* strain. Northern blot analysis of total RNA isolated at different timepoints following the growth curve of wild-type and Δ*hfq* cells. Samples were separated on a 4% polyacrylamide/7M Urea gel and a probe specific for the 17S 5'-end was used. (B) Predicted Hfq-binding motifs within the 17S rRNA flanking sequences. The Hfq binding motif (ARN)x is highlighted in blue. (R – purine; N – any nucleotide)

2.3.2. Hfq inactivation leads to altered ribosome sedimentation profiles

Given that misprocessing of rRNA can be consequence of defects in ribosome assembly (Liiv & Remme, 2004; Roy-Chaudhuri *et al*, 2010; Shajani *et al*, 2011), we next examined whether the defects in 16S rRNA maturation found in the Δ*hfq* strain had consequences to the pool of available ribosomes. We profiled the ribosomes from

exponential and stationary phase cultures of wild-type and mutant Δ*hfq* strains by sucrose gradient ultracentrifugation (Figure 12A and Figure 12B). The ribosome identity in the different peaks was further confirmed by analyzing their rRNAs (Figure S1). In the wild-type strain, under conditions that favor ribosome association (10 mM Mg^{2+}), the peak corresponding to the small 30S subunits was nearly absent, while the amount of the 70S ribosomes was comparable between exponential and stationary phase (Figure 12A, Figure 12B and Figure S1). In clear contrast, the levels of the mature 70S ribosomes were reduced in the Δ*hfq* mutant as compared to the wild-type, an effect particularly severe in the stationary phase. Additionally, free 30S accumulated in the Δ*hfq*, which again was more evident in stationary phase (Figure 12A and Figure 12B). The complementation of the Δ*hfq* deletion in trans with a plasmid expressing Hfq (pHfq) (Andrade *et al*, 2012), raised the amount of mature ribosomes to levels comparable to that of the wild-type strain (Figure 12A and Figure 12B). Strikingly, the plasmid expressing Hfq rescued the defects in the ribosomal amounts isolated from the Hfq-deletion strain. Note that the Δ*hfq* strain transformed with the empty vector was essentially identical to the Δ*hfq* strain suggesting no effects of the transformation itself.

Our data clearly demonstrate that the inactivation of Hfq leads to a reduction in the pool of 70S ribosomes in the cell. This could either result from imbalanced production of subunits or the occurrence of major defects in the assembly of the 70S particle upon inactivation of Hfq. To distinguish between these two possibilities, we further analyzed ribosomes under dissociative conditions (0.1 mM Mg^{2+}) to guarantee that all ribosomal subunits would be in their free state. As observed in Figure 12A and Figure 12B (right panels), both strains displayed comparable contents of 30S and 50S subunits irrespective of the growth phase. Hence, the lower levels of 70S ribosomes in the absence of Hfq (Figure 12A and Figure 12B, left panels) are a consequence of defects in the 70S assembly.

A well-known hallmark of ribosome biogenesis defects in bacteria is the cold-sensitive phenotype (Connolly & Culver, 2009). We next compared the growth of the wild-type and □hfq strains at 37ºC and 16ºC (Figure 12C). Clearly, the Δ*hfq* mutant exhibited the cold-sensitive phenotype, with severe growth defects under cold shock but not at 37ºC which correlated with the altered ribosome profile found in the absence of Hfq. This

effect is reminiscent of the cold-sensitive phenotype observed with different ribosome biogenesis factors like RbfA, KsgA, RimM and RimO (Bylund *et al*, 1998; Connolly *et al*, 2008; Leong *et al*, 2013).

rRNA synthesis feedforwards the synthesis of ribosomal proteins (Scott *et al*, 2014). Thus, to assess the expression of the r-proteins, we used ribosome profiling which captures the positions of actively translating ribosomes and the ribosome-protected fragments (RPFs) reporting on differences in gene expression at the level of translation (Ingolia *et al*, 2009; Li *et al*, 2014). This analysis was combined with RNA-Seq to determine the mRNA expression levels and the regulation of gene expression at the level of transcription. Strikingly, all ribosomal proteins were significantly translationally downregulated in the Δ*hfq* mutant strain while the levels of their transcripts remained unchanged or decreased to much lower extent (Figure 12D). Notably, among the significantly enriched gene ontology (GO) terms are genes participating in ribosome assembly. Furthermore, within the polycistronic mRNAs the translation yields of the encoded r-proteins differed implying an independent translation initiation of the r-proteins (Li *et al*, 2014). This expression pattern corroborates earlier observations for translational coupling of the expression of the ribosomal proteins and rRNA synthesis (Jinks-Robertson & Nomura, 1981; Nomura, 1999). Cumulative profiles of all expressed genes do not differ between wild-type and Δ*hfq* strain, arguing against an effect of Hfq depletion on translation initiation (Figure S2A).

Overall, our results show that the Hfq depletion leads to defects in ribosome biogenesis with consequences for the pool of mature 70S ribosomes and propose Hfq an auxiliary factor which regulates ribosome biogenesis.

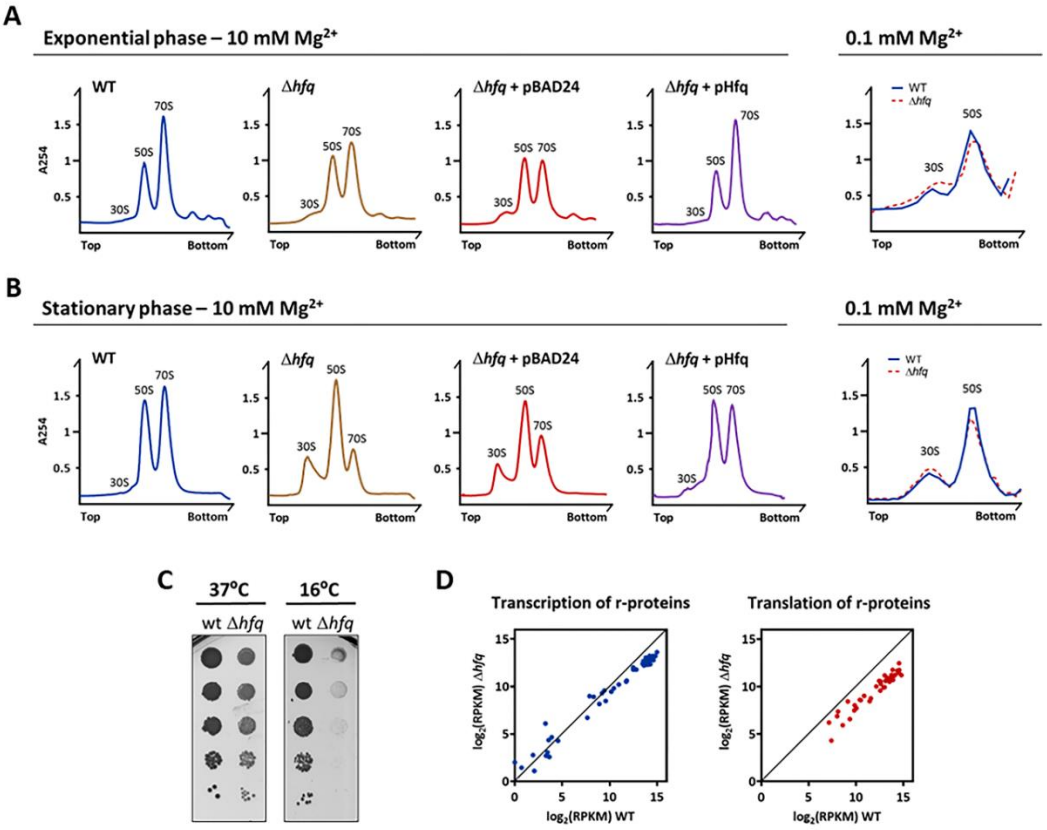


Figure 12 – Defective ribosome biogenesis in the Δ*hfq* strain. (A and B) (left panels) Ribosomes from cells in the exponential or stationary phase were fractionated on sucrose density gradients in 10 mM Mg^{2+} to stabilize 70S particles with and without trans-complementation of hfq gene using pBAD24 plasmid. Ribosome species are identified over each peak; top and bottom denote the lowest (15%) and highest (45%) sucrose concentration in the gradient, respectively. (right panels) Ribosomes purified from cells in exponential and stationary phase fractionated on sucrose density gradients at low 0.1 mM Mg^{2+} concentration to promote 70S dissociation into free 30S and 50S subunits. Top and bottom denote the lowest (10%) and highest (30%) sucrose concentration in the gradient, respectively. (C) Serial dilutions (with 1:10 steps) of wild-type and Δ*hfq* strains grown on LB-agar plates at 37ºC or 16ºC. (D) Comparison of mRNA expression (left) and protein production (right) of ribosomal proteins between wild-type and Δ*hfq* strains analyzed by RNA-Seq (left) and ribosome profiling (right), respectively.

2.3.3. Hfq copurifies with 30S immature subunits

We hypothesized Hfq would preferably bind to immature 30S subunits as these can be enriched in 17S RNA. To test this, we purified immature 30S subunits from the knockout mutant of RbfA, a late assembly factor that accumulates pre-30S particles enriched in 17S rRNA (Jones & Inouye, 1996; Bylund *et al*, 1998; Thurlow *et al*, 2016). Δ*rbf*A mutant showed a similar ribosome profile to the Δ*hfq* mutant, with increasing levels of 30S particles and lower levels of 70S ribosomes, when compared to the wild-type (Figure 13A). The peak corresponding to the 30S fraction was recovered from the sucrose gradients of the Δ*rbf*A mutant and the 30S subunits were purified in low salt conditions. In parallel, mature 30S subunits were obtained from dissociation of 70S ribosomes isolated from the wild-type, also in low salt conditions. Purified 30S samples were then analysed by mass spectrometry that identified proteins associated with 30S subunits. Most of the proteins identified corresponded to r-proteins or known factors associated to ribosomes. Strikingly, Hfq was found to copurify only with immature 30S isolated from the Δ*rbf*A but not with the mature 30S isolated from the wild-type (Figure 13B). The same 30S samples were analysed by Western blotting using an anti-Hfq antibody. Cell lysates of wild-type and Δ*hfq* strains and purified His-tagged Hfq were used as controls. Western blot confirmed the presence of Hfq in the 30S purified from the Δ*rbf*A but not from the wild-type, in total agreement with mass spectrometry data (Figure 13C). Overall, these results show that Hfq is copurifying with precursor 30S ribosomes *in vivo* and corroborates that Hfq is a novel factor that assists ribosome assembly.

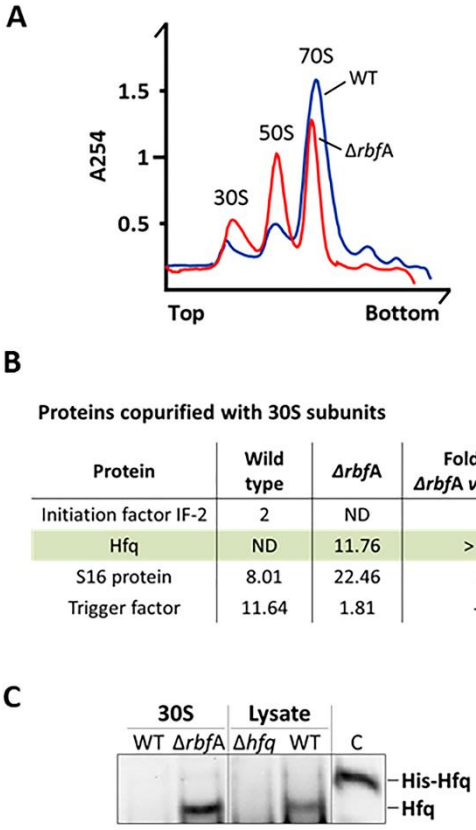


Figure 13 – Hfq copurifies with immature 30S subunits. (A) Ribosomes from wild-type and Δ*rbfA* exponential growing cells were analysed on sucrose density gradients. The Δ*rbfA* mutant displays an altered ribosome profile with an increase in 30S and 50S subunits and a reduction of 70S ribosomes compared to the wild-type. (B) Representative proteins identified by mass spectrometry of purified 30S subunits from the wild-type and Δ*rbfA* mutant. The measurement of all the peptides identified for each protein is shown as total ProtScore values calculated with the Pro GroupTM Algorithm (Sciex), with a 95% confidence. The ratio between the Δ*rbfA* mutant and wild-type are shown as normalised fold changes that are represented by positive or negative values corresponding to an increase or decrease, respectively, of the number of peptides found in the Δ*rbfA* mutant compared to the wild-type (ND, not detected). (C) Western blot analysis of purified 30S subunits using an anti-Hfq antibody. WT and Δ*hfq* cell lysates as well as purified His-Hfq protein were loaded as controls.

2.3.4. Translation efficiency is affected by Hfq inactivation

Altered ribosome biogenesis can lead to major defects in translation, thus we next assessed the translational status in the Δ*hfq* mutant. Firstly, the Δ*hfq* strain showed a reduced polysome fraction compared to that of the wild-type strain (Figure 14A). Secondly, a global measurement of protein synthesis by pulse metabolic labelling

confirmed a significant reduction of translation in Hfq-depleted background (Figure 14B). Thirdly, the global translation efficiency, which was determined by the density of ribosomes from the ribosome profiling per mRNA from the RNA-Seq dataset, was significantly reduced (Mann-Whitney U test or Wilcoxon rank-sum test, p = 0.0001996) (Figure 14C). Hence, the defects in rRNA precursor processing and ribosome biogenesis in the Δ*hfq* mutant decreased translation volume and efficiency as compared to the parental strain.

We next asked whether these changes in translation efficiency are global or a fraction of the genes escapes this trend. We performed a fold-change analysis and ranked the genes according to the fold-change in translation (i.e. only translationally up- or down-regulated in the ribosome profiling set) but with unchanged mRNA expression from the RNA-Seq experiment. Genes with changes in their RPF coverage higher than two-fold were considered. The gene ontology (GO) analysis of the down-regulated genes in Hfq-depleted background showed several pathways being affected but with a significant GO term enrichment in genes participating in ribosome biogenesis, translation and amino acid metabolism (Figure 14D). For comparison density plots of representative examples downregulated in the Δ*hfq* mutant (Figure S2B) or with unaltered translation (Figure S2C) are included. Notably, inactivation of Hfq augmented the mRNA levels of genes known to be regulated by Hfq-dependent sRNAs, while their translation was only slightly affected.

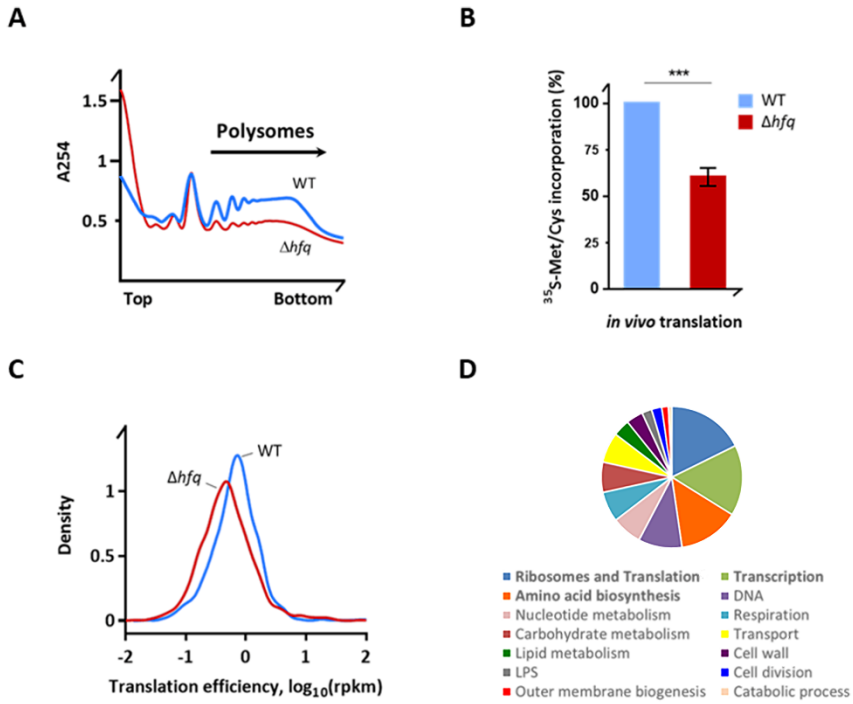


Figure 14 – The Δ*hfq* strain displays reduced translation levels. (A) Polysomal fraction is reduced in Δ*hfq* cells. Polysome profiles of the wild-type and Δ*hfq* strains were resolved on sucrose density gradient. Top and bottom denote the lowest (15%) and highest (50 %) sucrose concentration in the gradient, respectively. (B) *In vivo* incorporation of ^{35}S-Methionine/Cysteine translation assay in M9 medium. Data are normalized to the wild-type strain and are means ±SEM (n = 3). ***, $p = 0.0004$ (paired *t*-test). (C) Translation efficiency of wild-type and Hfq-depleted cells obtained by ribosome profiling. (D) GO term analysis of translationally downregulated genes in the Δ*hfq*. The top three affected categories are in bold.

2.3.5. Hfq affects translation fidelity

The ribosomal tRNA accommodation site (A-site) is formed by helix 44 of the 16S rRNA of the 30S subunit. Three aminoglycoside antibiotics, neomycin, paromomycin and kanamycin, interact with the 16S rRNA near the A-site and induce translational misreading (i.e. shift of the reading frame, stop codon readthrough) (Foster & Champney, 2008). In the presence of sub-lethal concentrations of neomycin, paromomycin or kanamycin, the Δ*hfq* mutant strain showed exacerbated growth defects relative to untreated Δ*hfq* or wild-type strains, suggesting that Hfq affects translation fidelity (Figure 15A). Additional aminoglycosides were further tested showing similar effect (Figure S3). As control, the Δ*hfq* strain did not show increased sensitivity to other classes of antibiotics, like colistin,

which targets cell membrane (Figure 15A and Figure S3). We also investigated the misreading using a collection of widely used plasmids bearing *lac*Z as reporter (O'Connor *et al*, 1997). When compared with the isogenic parent, the Δ*hfq* mutant showed a substantial increase in frameshifting, aberrant initiation from alternative start codon(s) and stop-codon readthrough (Figure 15B), indicating that the accuracy of translation in Hfq-depleted background is severely compromised. In sum, these data suggest that inactivation of Hfq decreases translation efficiency and enhances misreading of mRNA, implying a functional link between Hfq-dependent alterations in rRNA processing, ribosome biogenesis and translation fidelity.

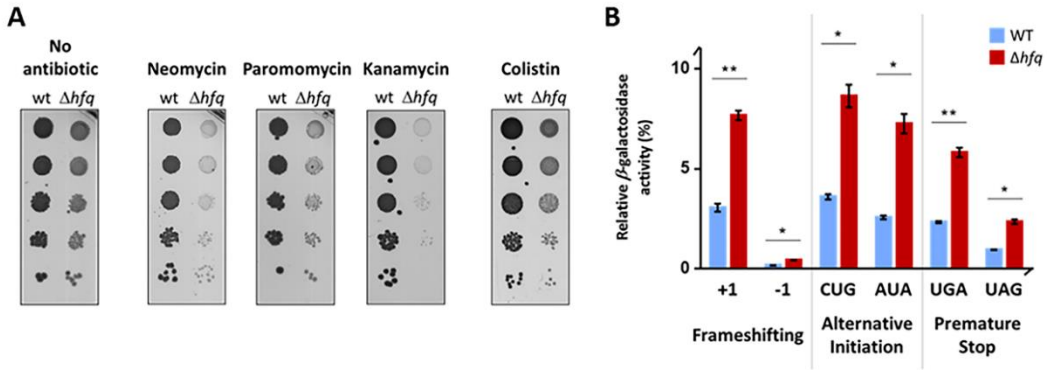


Figure 15 – Hfq-depleted cells exhibit increased codon misreading. (A) Serial dilutions (1:10) of wild-type and Δ*hfq* strains grown on LB-agar plates at 37ºC with and without sub-lethal concentrations of neomycin (1 μg/ml), paromomycin (1 μg/ml), kanamycin (1 μg/ml) or colistin (0.1 μg/ml). (B) Wild-type and Δ*hfq* strains expressing mutated *lacZ* gene (pSG plasmids) were tested for a frameshift mutation (+1 or -1), alternative initiation codons (CUG or AUA) or a non-sense stop codon mutation (UGA or UAG). For each strain, the *β*-galactosidase activity (in Miller units) was normalized to that of strain expressing the wild-type *lacZ*. Data are means ±SEM (n = 3). **, $p < 0.01$; *, $p < 0.02$ (paired *t*- test).

2.3.6. Hfq distal RNA contact surface is specifically required for ribosome biogenesis regulation

The Sm-like Hfq assembles into a hexamer with a ring-like shape that displays at least three RNA binding surfaces which confer Hfq the ability to bind simultaneously different RNA substrates. The proximal face and a charge patch in the outer rim of the hexamer bind preferably U-rich sRNAs while the distal face binds to A-rich sequences in

target mRNAs (Mikulecky *et al*, 2004; Link *et al*, 2009; Otaka *et al*, 2011; Sauer & Weichenrieder, 2011; Sauer *et al*, 2012; Panja *et al*, 2013; Zhang *et al*, 2013).

To evaluate which binding surface would be responsible for the newly identified Hfq-dependent regulation of ribosome biogenesis, representative Hfq variants with mutations in the different surfaces (Zhang *et al*, 2013) were tested. Ribosome sedimentation profiles of proximal (Q8A and F39A), rim (R16A) and distal (Y25D and K31A) mutants isolated from exponential cultures were compared to that of the wild-type strain (Figure 16A and Figure S4). In addition, rRNAs from each fraction were isolated to confirm the ribosome identity in each peak (Figure S5). Strikingly, only mutations in the distal face caused reduction in the 70S ribosome levels, which were similar to those we observed for the Hfq deletion mutant (Figure 12A and Figure 12B). The ribosome profiles of mutants in the proximal or rim surface were similar to that of the wild-type and these surfaces were shown to govern interactions with sRNAs (Sauer & Weichenrieder, 2011; Sauer *et al*, 2012; Panja *et al*, 2013; Zhang *et al*, 2013). Altogether, from these data we conclude that the distal face of Hfq is critical for the regulation of the rRNA maturation and ribosome biogenesis and propose that the novel function of Hfq in the ribosome biogenesis might be independent of sRNA binding.

The sensitivity of the Hfq variant strains against different antibiotics was also tested (Figure 16B). The proximal and rim mutants (Q8A and R16A) did not show significant growth difference to the wild-type. Only the distal Hfq-Y25D variant showed increased susceptibility to aminoglycosides, like neomycin or kanamycin, suggesting that Hfq-Y25D is impaired in translation efficiency. However, the growth defect of Hfq-Y25D strain is not as severe as the one found in the knockout Δ*hfq* mutant, which suggest that Y25D is an important residue but is not the sole responsible for the increased susceptibility to aminoglycosides. This phenotype was not observed when other classes of antibiotics were tested, such as polypeptide antimicrobials like colistin. These results from antibiotic sensitivity further support the importance of the distal face of Hfq in translation.

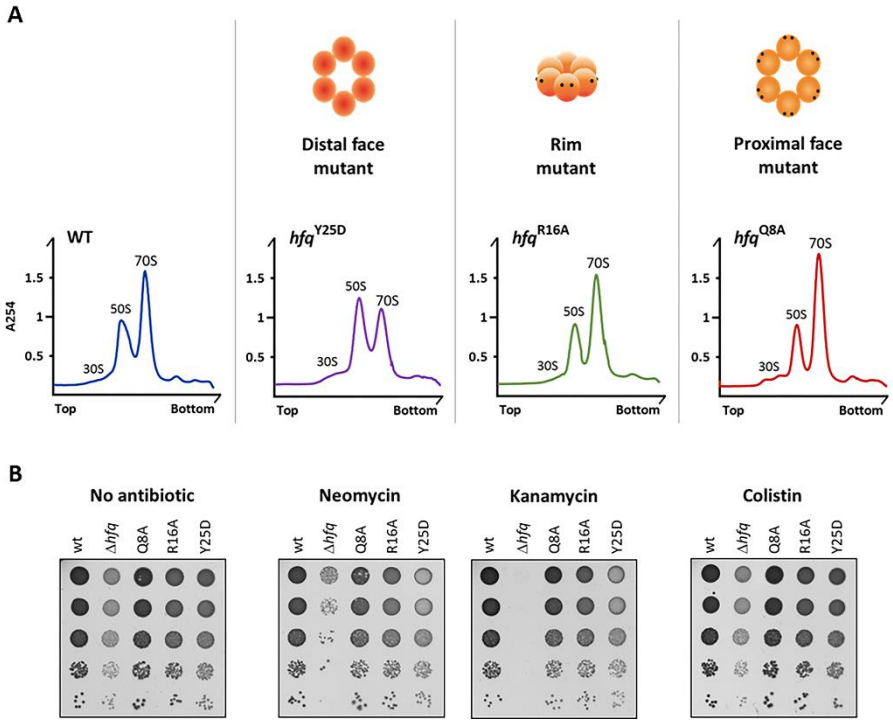


Figure 16 – The distal face of Hfq is required for correct ribosome biogenesis and translation fidelity. (A) Ribosomes purified from strains with specific point mutations in the *hfq* gene were fractionated on sucrose density gradients and compared to the wild-type strain. The binding surface affected by each mutation is schematically depicted on the top. (B) Serial dilutions (1:10) of wild-type, Δ*hfq* and Hfq variants grown on LB-agar plates at 37ºC with and without sub-lethal concentrations of neomycin (1 µg/ml), kanamycin (1 µg/ml) or colistin (0.1 µg/ml).

2.4. Discussion

Here we present results that support a novel role of Hfq in bacterial ribosome biogenesis with important consequences for translation (Figure 17). Hfq is a widely-conserved RNA-binding protein of the Sm/Lsm family of proteins (Wilusz & Wilusz, 2013) that it is mostly known for promoting sRNA basepairing with target mRNAs (Updegrove *et al*, 2016). However, a role of Hfq in regulating rRNA processing and folding has not been proposed. Our work unveils previously undescribed roles of Hfq in ribosome biogenesis and expands the repertoire of Hfq functions in the cell.

We show that Hfq is a new regulator of rRNA maturation. Hfq-depletion results in loss of normal processing of rRNA, leading to the accumulation of unprocessed 17S rRNA precursor. Earlier crosslinking studies in *E. coli* identified interactions between Hfq and

rRNA (Tree *et al*, 2014), but did not analyze it further. We find that Hfq directly interacts with the 17S rRNA and Hfq inactivation results in the misprocessing of both 17S extremities. Our data align well with observations made for Lsm proteins, the evolutionarily conserved eukaryotic counterparts of the bacterial Hfq; depletion of Lsm proteins causes defects in the processing of pre-rRNAs (Kufel *et al*, 2003; Beggs, 2005) supporting the notion for an evolutionary conserved function of the members of the Sm/Lsm protein family in rRNA processing.

Accumulation of the 17S rRNA precursor is usually linked to problems in formation of mature 30S subunit, most likely because maturation of 16S rRNA is a final step in ribosome biogenesis (Srivastava & Schlessinger, 1988; Shetty & Varshney, 2016). rRNA synthesis and maturation are tightly intertwined with the r-protein biosynthesis (Jinks-Robertson & Nomura, 1981; Nomura, 1999; Scott *et al*, 2014). In the Δ*hfq* background, accumulation of unprocessed 17S rRNA is accompanied with significant reduction of r-proteins synthesis, substantial reduction in the levels of 70S ribosomes and concomitant accumulation of immature ribosomal subunits. The initiation of translation is the most sensitive node in translation regulation and defects during this process could lead to a similar phenotype (Laursen *et al*, 2005). However, translation initiation remains unaffected upon Hfq inactivation (Figure S2). Along with the fact that *in trans* complementation with Hfq rescues the defective ribosome assembly we show that Hfq is needed for proper ribosome biogenesis but unessential for proper initiation. Hfq-depleted cells show phenotypes typically found in mutants of ribosome biogenesis factors; namely, defects in rRNA maturation, accumulation of rRNAs precursors and cold-sensitivity (Kaczanowska & Rydén-Aulin, 2007; Shajani *et al*, 2011). rRNA precursors compete with mature rRNAs for binding to r-proteins, although pre-rRNA containing ribosomes are conformationally defective (Liiv & Remme, 2004; Yang *et al*, 2014).

Several auxiliary factors associate with the ribosome during the intricate process of ribosome assembly assisting in r-protein binding and rRNA folding steps (Kaczanowska & Rydén-Aulin, 2007). Hfq is a novel assembly factor that preferentially bind immature 30S subunits, like other chaperones such as RimM or RbfA. Like these factors, Hfq probably acts to facilitate or proofread folding of the pre-rRNA and pre-30S assembly. Hfq

is a well-known RNA chaperone able to remodel RNA secondary structures (Moll *et al*, 2003; Wroblewska & Olejniczak, 2016), hence it is conceivable that it might be essential for the correct processing and folding of 16S rRNA into 30S subunits. In fact, cells lacking Hfq present an altered rRNA folding as suggested by our RNA-structure mapping. This is likely a consequence of the additional nucleotides from the 17S rRNA precursor which perturb the formation of helixes 1 and 2 of the 16S rRNA (Lodmell & Dahlberg, 1997; Roy-Chaudhuri *et al*, 2010). This in turn affects the folding of the central pseudoknot, a universally conserved structural element that establish long-range interactions within the 16S rRNA and that is critical for the overall folding of the small subunit (Brink *et al*, 1993). Moreover, alterations in the secondary structure of the pseudoknot result in error-prone ribosomes (Lodmell & Dahlberg, 1997) as we observed in the Δ*hfq* mutant. Alternatively, Hfq may promote RNA-protein interactions that are important for the correct rRNA processing. Notably, Hfq was previously shown to bind to the S12 protein of the 30S small subunit in *E. coli* (Strader *et al*, 2013). The S12 protein is a key mediator of fidelity of translation in both prokaryotes and eukaryotes and is positioned in helix 44 of the 16S rRNA that is known to form extensive contacts with the large subunit (Yusupov *et al*, 2001; Cukras *et al*, 2003). Association of Hfq with S12 is suggested to be important for the correct folding of 16S rRNA and formation of interface between ribosomal subunits and consequently the assembly of 70S ribosomes. Moreover, the sensitivity of the Δ*hfq* strain to aminoglycosides and the cumulative translation errors induced by *hfq* deletion corroborate with the observation that the conserved helix 44 of 16S rRNA maintains translation fidelity and serve as aminoglycoside target (Davis, 1987).

The role of Hfq in promoting the basepairing between regulatory small RNAs and their target mRNAs constitutes the most well-known function of this RNA-binding protein. An interesting feature of Hfq is that is possible to uncouple its multiple functions by introducing point mutations in each of its RNA-binding surfaces: the distal face of Hfq recognizes and binds to trinucleotide ARN repeats in mRNA, while the proximal and rim faces bind preferably to U-rich sequences in small RNAs (Link *et al*, 2009; Otaka *et al*, 2011; Sauer & Weichenrieder, 2011; Sauer *et al*, 2012; Panja *et al*, 2013). Strikingly, we found that the reduced levels of the 70S ribosomes in Hfq-depleted cells is dependent on

residues located at the distal face of Hfq but not on those in the proximal and rim RNA-binding faces suggesting that rRNA regulation is independent of sRNA-binding of Hfq. Despite Hfq being widely conserved, Hfq-dependent regulation of sRNAs is not a common feature; for example, this function is missing in many bacteria like *Bacillus subtilis* and *Listeria monocytogenes* (Christiansen *et al*, 2006; Rochat *et al*, 2015). Hfq is known to act independently of an sRNA as partner in a variety of cellular functions. Namely, Hfq stimulates the addition of poly(A) tails to the 3' end of mRNAs containing Rho-independent transcription terminators, promoting their degradation in *E. coli* (Le Derout *et al*, 2003; Mohanty *et al*, 2004; Folichon *et al*, 2005; Régnier & Hajnsdorf, 2013). Also, Hfq inhibits translation by binding directly to mRNAs, independent of a sRNA partner (Salvail *et al*, 2013).

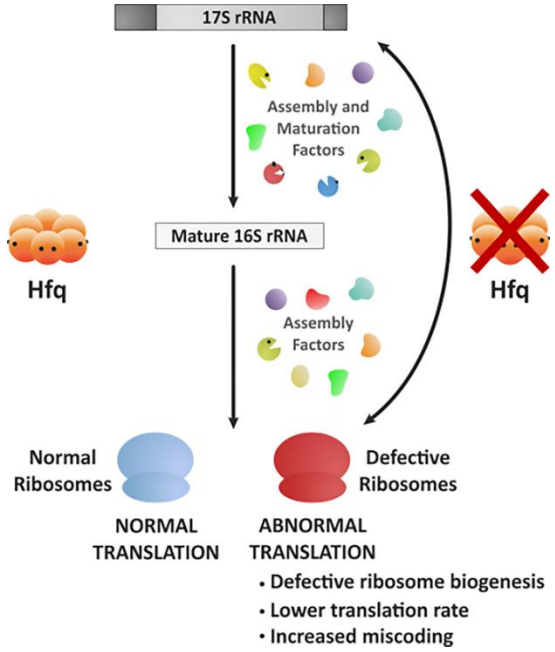


Figure 17 – Model for the Hfq regulation of ribosome biogenesis. Hfq assists ribosome biogenesis together with other ribosome assembly and maturation factors. Hfq depletion exhibits critical consequences for ribosome biogenesis and cellular translation. The Δ*hfq* mutant affects the correct maturation of 30S subunits and accumulates unprocessed 17S rRNA precursor leading to a general translation deficiency.

In summary, we have demonstrated that Hfq is a new ribosome assembly factor. Cells lacking Hfq exhibit diverse hallmarks of ribosome biogenesis defects, namely: i) misprocessing of rRNA and accumulation of 17S rRNA precursor; ii) reduced pool of 70S ribosomes; iii) an abnormal translation and compromised translation fidelity; and iv) cold-sensitive phenotype, typically associated with ribosome biogenesis factor mutants. This work expands the functions of Hfq beyond the regulation of small non-coding RNA biology and unveils unprecedented roles in ribosome biogenesis and translation.

2.5. Materials and Methods

2.5.1. Bacterial strains, plasmids and oligonucleotides

All bacterial strains, plasmids and oligonucleotides are listed in Table S1, Table S2 and Table S3, respectively. All *E. coli* K-12 strains used in this study are derivatives of strains MG1693 or MC1061. Deletion of *hfq* was obtained using the λ-Red recombination (Datsenko & Wanner, 2000). The *hfq* point mutant alleles (Zhang *et al*, 2013) were P1-transduced to our parental strain, following selection on glucose minimal plates and screening for sensitivity to chloramphenicol. The Δ*rbf*A mutant was obtained from the Keio collection (Baba *et al*, 2006). All mutations were confirmed by PCR and sequencing.

2.5.2. Bacterial growth

Strains were grown in LB medium (Difco) supplemented with thymine (50 μg/ml) at 37ºC, unless otherwise stated. Overnight cultures of single freshly grown colonies were diluted to an initial OD_{600} ~ 0.03. Cultures were collected either at exponential phase (OD_{600} ~ 0.5) or stationary phase (after ~14h growth). Antibiotics were present at the following concentrations when needed: 25 μg/ml chloramphenicol, 50 μg/ml kanamycin, 10 μg/ml tetracycline; 100 μg/ml ampicillin. For the dilution plating assays, serial dilutions were made in 10-fold increments and immediately spotted onto LB-agar plates. Sub-lethal concentrations of antibiotics were added when relevant: 1 μg/ml neomycin, 1 μg/ml paromomycin, 1 μg/ml kanamycin, 0.1 μg/ml colistin, 0.1 μg/ml gentamicin, 1 μg/ml

streptomycin, 0.01 µg/ml cefotaxime, 2 µg/ml erythromycin and 0.002 µg/ml ciprofloxacin.

2.5.3. RNA analysis

For Northern blots, total RNA was extracted as previously described (Andrade *et al*, 2012). One microgram of total RNA was resolved on 4% polyacrylamide/8M urea gels in TBE 1x buffer, transferred to a nylon membrane (GE Healthcare) and UV crosslinked. Membranes were hybridized with PerfectHyb Plus (Sigma Aldrich) and probed with ^{32}P-5'end-labeled DNA oligonucleotides. Blots were analyzed on the Fuji TLA-5100 imaging system (GE Healthcare). RNAs collected from ribosome sedimentation fractions were extracted using TRI Reagent (Sigma Aldrich) and resolved on agarose gels stained with ethidium bromide.

2.5.4. Electrophoretic mobility shift assays

Binding assays were performed essentially as previously described (Andrade *et al*, 2013). The 17S rRNA extremities were generated by *in vitro* transcription with T7 RNAP (Promega) and [α-^{32}P]-UTP (Perkin Elmer). EMSA samples were electrophoresed on native 6% or 8% polyacrylamide gels in TBE 1x buffer in a cold room. Gels were exposed to a PhosporImager screen (GE Healthcare).

2.5.5. RNA mapping

Chemical modification reactions were carried out with DMS (diluted 1:6 in ethanol) or CMCT (1 mg/mL) following protocols described in Andrade *et al*, 2013 and Caprara, 2011, respectively. Total RNA (10 µg) extracted from exponential phase cultures (OD_{600} ~ 0.35-0.40) of wild-type and Δ*hfq* strains was used. Primer extension reactions were carried out using the ^{32}P-5'-end labelled primer 46 (Clatterbuck Soper *et al*, 2013) and 100U of reverse transcriptase SuperScript III or IV (Thermo Fisher Scientific). Samples were analyzed on 10% polyacrylamide/7M urea gels run in TBE 1x buffer.

2.5.6. Ribosome extraction and sucrose sedimentation

Ribosome isolation was adapted from (Powers & Noller, 1991). Cell pellets were resuspended in ice-cold buffer A (50 mM Tris-Cl at pH 7.5, 10 mM $MgCl_2$, 100 mM NH_4Cl, 0.5 mM EDTA, and 6 mM 2-mercaptoethanol) with the addition of Complete Mini Protease Inhibitor cocktail EDTA-free (Roche) and lysed by French press. After TurboDNase (Ambion) digestion, the clarified lysate was layered over a 36% sucrose cushion composed of buffer B (50 mM Tris-Cl at pH 7.5, 10 mM $MgCl_2$, 500 mM NH_4Cl, 0.5 mM EDTA, and 6 mM 2-mercaptoethanol) and spun at 44,000 rpm for 16 h in a Beckman ultracentrifuge 90Ti rotor at 4°C. The ribosome pellets were washed once with buffer C (50 mM Tris-Cl at pH 7.5, 10 mM $MgCl_2$, 100 mM NH_4Cl, and 6 mM 2-mercaptoethanol) and then resuspended in the same buffer by gentle rocking at 4°C. Purified ribosomes were analyzed in 15%-50% (w/v) sucrose gradients prepared in buffer C with 10 mM $MgCl_2$ (associative conditions) or in 10%-30% (w/v) sucrose gradients prepared in buffer C with 0.1 mM $MgCl_2$ (dissociative conditions). Associative samples were centrifuged in a Beckman ultracentrifuge SW41 rotor for 16 h at 24,000 rpm at 4°C and analyzed by UV using the AKTA system (GE Healthcare). Dissociative samples were centrifuged in a Beckman ultracentrifuge SW28 rotor for 16 h at 24,000 rpm at 4°C and fractions collected from the top were quantified on Nanodrop.

2.5.7. Ribosome profiling, RNA-Seq and data analysis

Ribosome-protected fragments and randomly fragmented mRNA for ribosome profiling and RNA-Seq were isolated as described previously (Del Campo *et al*, 2015). Briefly, cells cultured to the exponential phase (OD_{600} 0.35-0.40) in LB medium were split into two aliquots. From one aliquot total RNA was extracted using TRI Reagent (Sigma Aldrich), enriched by depleting small RNAs with GeneJET Purification Kit (Fermentas) and rRNA with MICROBExpres Bacterial mRNA Enrichment Kit (Ambion) and fragmented in alkaline solution (2 mM EDTA and 100 mM Na_2CO_3 pH 9.2 for 40 min at 95°C) to fragments with size of 24-35 nts. The second aliquot was used to isolate mRNA-bound ribosome complexes. Cells were collected by filtration, flash-frozen without preincubation with antibiotics. Cells were lysed by freeze-rupturing (Retch Mill) and 100 A_{260} units of

ribosome-bound mRNA fraction were directly used for polysomal analysis or subjected to nucleolytic digestion with 10 units/µl micrococcal nuclease (Fermentas) for 10 min at room temperature in buffer with pH 9.2 (10 mM Tris pH 11 containing 50 mM NH_4Cl, 10 mM $MgCl_2$, 0.2% triton X-100, 100 µg/ml chloramphenicol and 20 mM $CaCl_2$) to obtain the monosomal fraction. Separation was obtained by sucrose density gradient (15-50% w/v). Subsequently, 20-35-nt RNA fragments from the monosomal fraction were size selected on a denaturing 15% polyacrylamide gel. For both ribosome-protected fragments and mRNA fragments the libraries were prepared by direct ligation of the adaptors (Del Campo *et al*, 2015) and sequenced on the Illumina GAIIx platform. Sequenced reads were quality trimmed using *fastx-toolkit* (0.0.13.2; quality threshold: 20), sequencing adapters were cut using *cutadapt* (1.8.3); minimal overlap: 1 nt) and mapped to the *E. coli* genome (strain MG1655, version U00096.3, NCBI) using Bowtie (1.1.2) allowing a maximum of two mismatches. The number of raw reads were used to generate gene read counts for each ORF, by counting the number of reads whose middle nucleotide (for even read length the nucleotide 5' of the mid-position) fell in the CDS. Gene read counts were normalized by the length of the unique CDS per kilobase (RPKM) and the total mapped reads per million (RPM) (Mortazavi *et al*, 2008). Spike-ins (ERCC, Thermo, Germany) were added to the RNA-Seq data set upon rRNA depletion with MICROBExpress kit and used to set the detection threshold in each sequencing set. The same detection threshold was used for the corresponding ribosome profiling experiment. Furthermore, to determine the reproducibility of our sequencing data sets, we used published data set serving as a truly independent biological replicate in which bacteria were grown under identical conditions (GEO accession number, GSE85540) (Hwang & Buskirk, 2017). The reproducibility is very high, R^2=0.865 and R^2=0.816 (Spearman correlation coefficient) for the RNA-Seq and ribosome profiling data sets, respectively. The correlation is even higher for the r-proteins only. For fold-change analysis we used a threshold of 2. Cumulative profiles of read density for RPFs have been computed as described (Ingolia *et al*, 2009). The overlapping genes were excluded from this analysis as initiation of the downstream gene is within the open-reading frame of the upstream gene and the RPFs in this region cannot be unambiguously assigned to either gene. Gene ontology enrichment including statistical

analysis was performed using the tools and gene lists from Gene Ontology Consortium (http://geneontology.org/).

2.5.8. Analysis of purified 30S associated proteins

For 30S purification cells were grown in 1L of LB medium at 37ºC and 160 rpm agitation to an OD_{600} ~ 0.6 for the wild-type strain and OD_{600} ~ 0.2 in the case of the Δ*rbf*A strain, as previously described with minor modifications (Thurlow *et al*, 2016). Ribosomes were isolated in a similar manner as detailed above. However, "low salt conditions" were used to allow mass spectrometry analysis, meaning that all buffers contained only 60mM of NH_4Cl. Isolated ribosomes were then quantified and separated on 15%-45% (w/v) sucrose gradients under dissociative (0.1 mM $MgCl_2$) and associative conditions (10 mM $MgCl_2$) for the wild-type and Δ*rbf*A strain, respectively. Gradients were centrifuged in a Beckman ultracentrifuge SW41 rotor at 24,000 rpm and 4ºC for 16 h and analyzed by UV using the AKTA system (GE Healthcare). Fractions corresponding to 30S peak were collected and spun in a Beckman 90Ti rotor at 44,000 rpm and 4ºC for 16h to remove the sucrose buffer from the 30S particles. The pellet was then resuspended in buffer D (10 mM Tris-Cl at pH 7.5, 10 mM $MgCl_2$, 60 mM NH_4Cl, and 3 mM 2-mercaptoethanol) and stored at -80ºC. Mature 30S subunits isolated from the wild-type strain and immature 30S particles isolated from the Δ*rbf*A strain were quantified on Nanodrop. Mass spectrometry data was obtained by the UniMS service (Mass Spectrometry Unit, ITQB/iBET, Oeiras, Portugal). Peptides were analysed using the Pro Group™ Algorithm (Sciex) and for each protein two types of scores were obtained: unused and total ProtScore. While the latter is a sum of the ion scores of all identified peptide evidence for a protein, the unused ProtScore reflects the amount of total unique peptide evidence related to the same protein. The confidence threshold was set at unused score of 2 and 1.3 with 99% and 95% confidence, respectively. A ratio from the Δ*rbf*A strain over the wild-type control was used to identify fold-change variation of proteins. Positive or negative fold-change values correspond to an increase or decrease, respectively, of the number of peptides found in the Δ*rbf*A mutant compared to the wild-type. Hfq presence in purified 30S samples (2.5

μg) was further analysed by western blot using an anti-Hfq antibody (Ziolkowska *et al*, 2006).

2.5.9. Pulse-labelling assay

Bacteria were grown in M9 medium supplemented with 0.02% casaminoacids (Difco) in an orbital shaker at 37ºC. Exponential phase cells were centrifuged, resuspended in M9 medium supplemented with 0.15 mM amino acid mix without methionine and cysteine (Promega) and incubated for 60 minutes in a water-bath at 37ºC. Labelling with ^{35}S-radiolabeled L-Met/L-Cys mix (Perkin Elmer) proceeded for 30 seconds at 37ºC. Reaction was stopped with addition of TCA to a final concentration of 5% and samples were spotted onto GF/C glass microfibers filters (Millipore). Filters were washed four times with TCA 5%, once with ethanol and then dried under vacuum. ^{35}S signal on filters was quantified by scintillation counting using the Ready Safe Liquid Scintillation cocktail (Beckman Coulter).

2.5.10. β-Galactosidase assay

Translation fidelity was analyzed by measurement of the *β*-galactosidase activity using the pSG plasmid series (O'Connor *et al*, 1997). Cells were grown to log phase (OD_{600} ~ 0.35-0.40) in LB medium at 37ºC. *β*-galactosidase activity from the plasmid encoding WT *lac*Z was used for normalization in the respective set of MC1061 strains or MC1061 Δ*hfq* mutant strains. Paired *t*-test statistical analysis performed using GraphPad Prism 6 software.

2.5.11. Statistical analysis and data deposition

The sequencing data were also submitted to GEO under the accession number **GSE100373**

2.6. References

Andrade JM, Pobre V & Arraiano CM (2013) Small RNA modules confer different stabilities and interact differently with multiple targets. *PLoS One* **8:** e52866

Andrade JM, Pobre V, Matos AM & Arraiano CM (2012) The crucial role of PNPase in the degradation of small RNAs that are not associated with Hfq. *RNA* **18:** 844–55

Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL & Mori H (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2:** 2006.0008

Beggs JD (2005) Lsm proteins and RNA processing. *Biochem. Soc. Trans.* **33:** 433–8

Bilusic I, Popitsch N, Rescheneder P, Schroeder R & Lybecker M (2014) Revisiting the coding potential of the *E. coli* genome through Hfq co-immunoprecipitation. *RNA Biol.* **11:** 641–54

Brink MF, Verbeet MP & de Boer HA (1993) Formation of the central pseudoknot in 16S rRNA is essential for initiation of translation. *EMBO J.* **12:** 3987–96

Bylund GO, Wipemo LC, Lundberg LA & Wikström PM (1998) RimM and RbfA are essential for efficient processing of 16S rRNA in *Escherichia coli*. *J. Bacteriol.* **180:** 73–82

Del Campo C, Bartholomäus A, Fedyunin I & Ignatova Z (2015) Secondary structure across the bacterial transcriptome reveals versatile roles in mRNA regulation and function. *PLoS Genet.* **11:** e1005613

Christiansen JK, Nielsen JS, Ebersbach T, Valentin-Hansen P, Søgaard-Andersen L & Kallipolitis BH (2006) Identification of small Hfq-binding RNAs in *Listeria monocytogenes*. *RNA* **12:** 1383–96

Clatterbuck Soper SF, Dator RP, Limbach PA & Woodson SA (2013) In vivo X-ray footprinting of pre-30S ribosomes reveals chaperone-dependent remodeling of late assembly intermediates. *Mol. Cell* **52:** 506–16

Cole SE, LaRiviere FJ, Merrikh CN & Moore MJ (2009) A convergence of rRNA and mRNA quality control pathways revealed by mechanistic analysis of nonfunctional rRNA decay. *Mol. Cell* **34:** 440–50

Connolly K & Culver G (2009) Deconstructing ribosome construction. *Trends Biochem. Sci.* **34:** 256–63

Connolly K, Rife JP & Culver G (2008) Mechanistic insight into the ribosome biogenesis functions of the ancient protein KsgA. *Mol. Microbiol.* **70:** 1062–75

Cukras AR, Southworth DR, Brunelle JL, Culver GM & Green R (2003) Ribosomal proteins S12 and S13 function as control elements for translocation of the mRNA:tRNA complex. *Mol. Cell* **12:** 321–8

Datsenko KA & Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* **97:** 6640–5

Davis BD (1987) Mechanism of bactericidal action of aminoglycosides. *Microbiol. Rev.* **51:** 341–50

Davis JH & Williamson JR (2017) Structure and dynamics of bacterial ribosome biogenesis. *Philos. Trans. R. Soc. B Biol. Sci.* **372:** 20160181

Le Derout J, Folichon M, Briani F, Dehò G, Régnier P & Hajnsdorf E (2003) Hfq affects the length and the frequency of short oligo(A) tails at the 3' end of *Escherichia coli rpsO* mRNAs. *Nucleic Acids Res.* **31:** 4017–23

Deutscher MP (2009) Maturation and degradation of ribosomal RNA in bacteria. *Prog. Mol. Biol. Transl. Sci.* **85:** 369–91

Folichon M, Allemand F, Régnier P & 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:** 454–63

Foster C & Champney WS (2008) Characterization of a 30S ribosomal subunit assembly intermediate found in *Escherichia coli* cells growing with neomycin or paromomycin. *Arch. Microbiol.* **189:** 441–9

Fujii K, Kitabatake M, Sakata T & Ohno M (2012) 40S subunit dissociation and proteasome-dependent RNA degradation in nonfunctional 25S rRNA decay. *EMBO J.* **31:** 2579–89

Hajnsdorf E & Boni I V (2012) Multiple activities of RNA-binding proteins S1 and Hfq. *Biochimie* **94:** 1544–53

de Haseth PL & Uhlenbeck OC (1980) Interaction of *Escherichia coli* host factor protein with Q beta ribonucleic acid. *Biochemistry* **19:** 6146–51

Hwang J-Y & Buskirk AR (2017) A ribosome profiling study of mRNA cleavage by the endonuclease RelE. *Nucleic Acids Res.* **45:** 327–336

Ingolia NT, Ghaemmaghami S, Newman JRS & Weissman JS (2009) Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* **324:** 218–23

Jinks-Robertson S & Nomura M (1981) Regulation of ribosomal protein synthesis in an *Escherichia coli* mutant missing ribosomal protein L1. *J. Bacteriol.* **145:** 1445–7

Jones PG & Inouye M (1996) RbfA, a 30S ribosomal binding factor, is a cold-shock protein whose absence triggers the cold-shock response. *Mol. Microbiol.* **21:** 1207–18

Kaczanowska M & Rydén-Aulin M (2007) Ribosome biogenesis and the translation process in *Escherichia coli*. *Microbiol. Mol. Biol. Rev.* **71:** 477–94

Karbstein K (2013) Quality control mechanisms during ribosome maturation. *Trends Cell Biol.* **23:** 242–50

Kufel J, Allmang C, Petfalski E, Beggs J & Tollervey D (2003) Lsm Proteins are required for normal processing and stability of ribosomal RNAs. *J. Biol. Chem.* **278:** 2147–56

Laursen BS, Sørensen HP, Mortensen KK & Sperling-Petersen HU (2005) Initiation of protein synthesis in bacteria. *Microbiol. Mol. Biol. Rev.* **69:** 101–23

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

Leong V, Kent M, Jomaa A & Ortega J (2013) *Escherichia coli rimM* and *yjeQ* null strains accumulate immature 30S subunits of similar structure and protein complement. *RNA* **19:** 789–802

Li GW, Burkhardt D, Gross C & Weissman JS (2014) Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell* **157:** 624–635

Liiv A & Remme J (2004) Importance of transient structures during post-transcriptional refolding of the pre-23S rRNA and ribosomal large subunit assembly. *J. Mol. Biol.* **342:** 725–41

Link TM, Valentin-Hansen P & Brennan RG (2009) Structure of *Escherichia coli* Hfq bound to polyriboadenylate RNA. *Proc. Natl. Acad. Sci. U. S. A.* **106:** 19292–7

Lodmell JS & Dahlberg AE (1997) A conformational switch in *Escherichia coli* 16S ribosomal RNA during decoding of messenger RNA. *Science* **277:** 1262–7

Mikulecky PJ, Kaw MK, Brescia CC, Takach JC, Sledjeski DD & Feig AL (2004) *Escherichia coli* Hfq has distinct interaction surfaces for DsrA, rpoS and poly(A) RNAs. *Nat. Struct. Mol. Biol.* **11:** 1206–14

Mohanty BK, Maples VF & Kushner SR (2004) The Sm-like protein Hfq regulates polyadenylation dependent mRNA decay in *Escherichia coli*. *Mol. Microbiol.* **54:** 905–20

Moll I, Leitsch D, Steinhauser T & Bläsi U (2003) RNA chaperone activity of the Sm-like Hfq protein. *EMBO Rep.* **4:** 284–9

Mortazavi A, Williams BA, McCue K, Schaeffer L & Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5:** 621–8

Muffler A, Fischer D & Hengge-Aronis R (1996) The RNA-binding protein HF-I, known as a host factor for phage Qbeta RNA replication, is essential for *rpoS* translation in *Escherichia coli*. *Genes Dev.* **10:** 1143–51

Narla A & Ebert BL (2010) Ribosomopathies: human disorders of ribosome dysfunction. *Blood* **115:** 3196–3205

Nomura M (1999) Regulation of ribosome biosynthesis in *Escherichia coli* and *Saccharomyces cerevisiae*: diversity and common principles. *J. Bacteriol.* **181:** 6857–64

O'Connor M, Thomas CL, Zimmermann RA & Dahlberg AE (1997) Decoding fidelity at the ribosomal A and P sites: influence of mutations in three different regions of the decoding domain in 16S rRNA. *Nucleic Acids Res.* **25:** 1185–93

Otaka H, Ishikawa H, Morita T & 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:** 13059–64

Panja S, Schu DJ & Woodson SA (2013) Conserved arginines on the rim of Hfq catalyze base pair formation and exchange. *Nucleic Acids Res.* **41:** 7536–46

Peng Y, Soper TJ & Woodson SA (2014) Positional effects of AAN motifs in rpoS regulation by sRNAs and Hfq. *J. Mol. Biol.* **426:** 275–85

Powers T & Noller HFF (1991) A functional pseudoknot in 16S ribosomal RNA. *EMBO J.* **10:** 2203–14

Régnier P & 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:** 602–9

Rochat T, Delumeau O, Figueroa-Bossi N, Noirot P, Bossi L, Dervyn E & Bouloc P (2015) Tracking the Elusive Function of *Bacillus subtilis* Hfq. *PLoS One* **10:** e0124977

Roy-Chaudhuri B, Kirthi N & Culver GM (2010) Appropriate maturation and folding of 16S rRNA during 30S subunit biogenesis are critical for translational fidelity. *Proc. Natl. Acad. Sci. U. S. A.* **107:** 4567–72

Salvail H, Caron M-P, Bélanger J & Massé E (2013) Antagonistic functions between the RNA chaperone Hfq and an sRNA regulate sensitivity to the antibiotic colicin. *EMBO J.* **32:** 2764–78

Sauer E, Schmidt S & Weichenrieder O (2012) Small RNA binding to the lateral surface of Hfq hexamers and structural rearrangements upon mRNA target recognition. *Proc. Natl. Acad. Sci. U. S. A.* **109:** 9396–401

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

Scott M, Klumpp S, Mateescu EM & Hwa T (2014) Emergence of robust growth laws from optimal regulation of ribosome synthesis. *Mol. Syst. Biol.* **10:** 747

Shajani Z, Sykes MT & Williamson JR (2011) Assembly of bacterial ribosomes. *Annu. Rev. Biochem.* **80:** 501–26

Shetty S & Varshney U (2016) An evolutionarily conserved element in initiator tRNAs prompts ultimate steps in ribosome maturation. *Proc. Natl. Acad. Sci.* **113:** E6126–E6134

Sittka A, Lucchini S, Papenfort K, Sharma CM, Rolle K, Binnewies TT, Hinton JCD & Vogel J (2008) Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq. *PLoS Genet.* **4:** e1000163

Srivastava AK & Schlessinger D (1988) Coregulation of processing and translation: mature 5' termini of *Escherichia coli* 23S ribosomal RNA form in polysomes. *Proc. Natl. Acad. Sci. U. S. A.* **85:** 7144–8

Strader MB, Hervey WJ, Costantino N, Fujigaki S, Chen CY, Akal-Strader A, Ihunnah CA, Makusky AJ, Court DL, Markey SP & Kowalak JA (2013) A coordinated proteomic approach for identifying proteins that interact with the *E. coli* ribosomal protein S12. *J. Proteome Res.* **12:** 1289–99

Thurlow B, Davis JH, Leong V, F Moraes T, Williamson JR & Ortega J (2016) Binding properties of YjeQ (RsgA), RbfA, RimM and Era to assembly intermediates of the 30S subunit. *Nucleic Acids Res.* **44:** 9918–9932

Tree JJ, Granneman S, McAteer SP, Tollervey D & Gally DL (2014) Identification of bacteriophage-encoded anti-sRNAs in pathogenic *Escherichia coli*. *Mol. Cell* **55:** 199–213

Tsui HC, Leung HC & Winkler ME (1994) Characterization of broadly pleiotropic phenotypes caused by an hfq insertion mutation in *Escherichia coli* K-12. *Mol. Microbiol.* **13:** 35–49

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

Vogel J & Luisi BF (2011) Hfq and its constellation of RNA. *Nat. Rev. Microbiol.* **9:** 578–89

Wilusz CJ & Wilusz J (2013) Lsm proteins and Hfq: Life at the 3' end. *RNA Biol.* **10:** 592–601

Woodson SA (2008) RNA folding and ribosome assembly. *Curr. Opin. Chem. Biol.* **12:** 667–73

Wroblewska Z & Olejniczak M (2016) Hfq assists small RNAs in binding to the coding sequence of *ompD* mRNA and in rearranging its structure. *RNA* **22:** 979–94

Yang Z, Guo Q, Goto S, Chen Y, Li N, Yan K, Zhang Y, Muto A, Deng H, Himeno H, Lei J & Gao N (2014) Structural insights into the assembly of the 30S ribosomal subunit in vivo: functional role of S5 and location of the 17S rRNA precursor sequence. *Protein Cell* **5:** 394–407

Yusupov MM, Yusupova GZ, Baucom A, Lieberman K, Earnest TN, Cate JH & Noller HF (2001) Crystal structure of the ribosome at 5.5 A resolution. *Science* **292:** 883–96

Zhang A, Schu DJ, Tjaden BC, Storz G & Gottesman S (2013) Mutations in interaction surfaces differentially impact *E. coli* Hfq association with small RNAs and their mRNA targets. *J. Mol. Biol.* **425:** 3678–97

Zhang A, Wassarman KM, Rosenow C, Tjaden BC, Storz G & Gottesman S (2003) Global analysis of small RNA and mRNA targets of Hfq. *Mol. Microbiol.* **50:** 1111–24

Ziolkowska K, Derreumaux P, Folichon M, Pellegrini O, Régnier P, Boni I V & Hajnsdorf E (2006) Hfq variant with altered RNA binding functions. *Nucleic Acids Res.* **34:** 709–20

2.7. Supplemental information

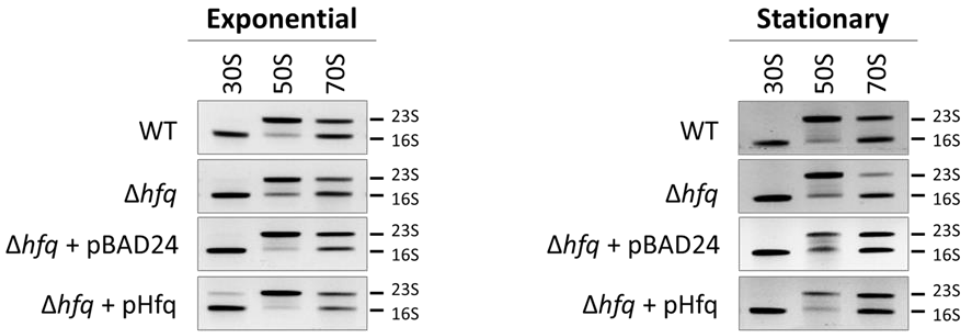


Figure S1 – Peak identification of ribosomes isolated by sucrose gradient in wild-type and Δ*hfq* strains. Fractions corresponding to 30S, 50S and 70S in Fig. 2A and 2B were collected and total RNA was isolated and separated on an agarose gel stained with ethidium bromide. The rRNA species are indicated.

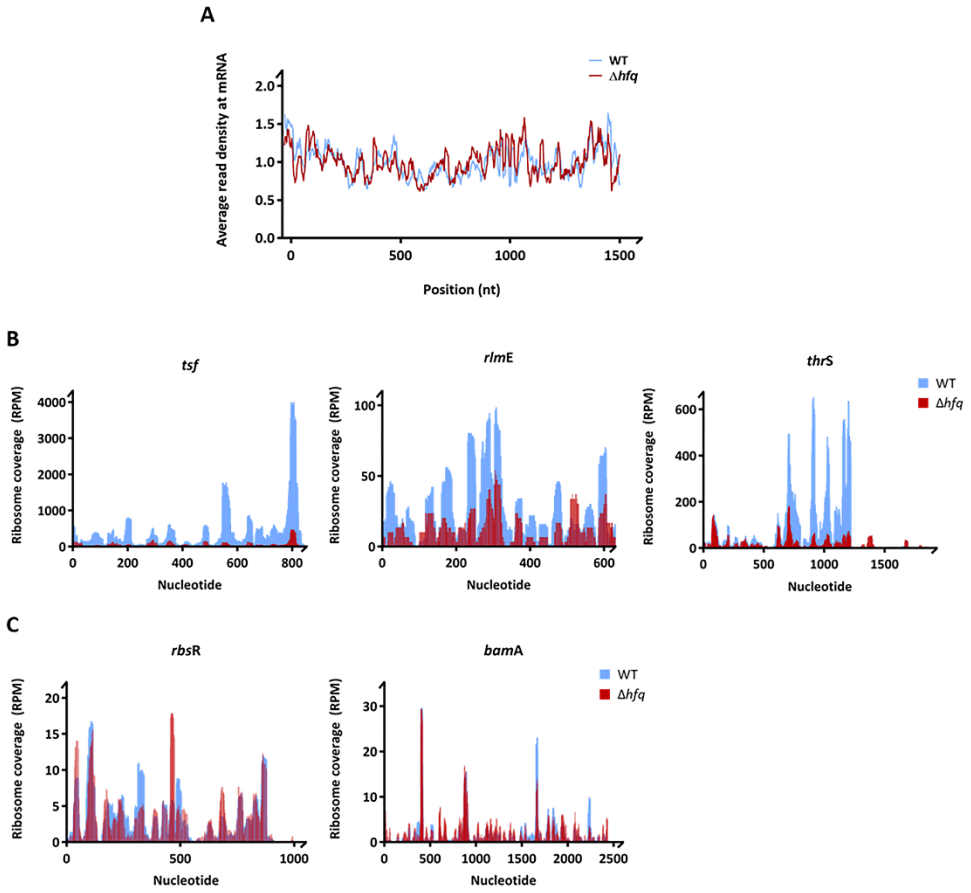


Figure S2– Cumulative metagene profile and coverage profiles of selected downregulated genes in the wild-type and Δ*hfq* strain obtained by ribosome profiling. (A) Cumulative metagene profile of the read density as a function of position for RPFs. The expressed genes were individually normalized, aligned at the start codon and averaged with equal weight. 1075 and 1231 genes from wild-type and Δ*hfq* strains, respectively, were considered in the analysis. (B) Coverage profiles of selected downregulated genes representative of gene categories affected by Hfq deletion (Fig. 4D). Elongation factor-Ts (*tsf*), 23S rRNA 2'-O-ribose U2552 methyltransferase (*rlm*E) are included in the GO term "Translation and ribosome" and threonine-tRNA ligase (*thr*S) in the GO Term "Amino acid biosynthesis". (C) Coverage profiles of exemplified genes (*bam*A and *rbs*R) whose expression remained unchanged upon *hfq* deletion. BamA is an outer membrane protein assembly factor and RbsR is a transcriptional factor of the operon involved in ribose catabolism and transport.

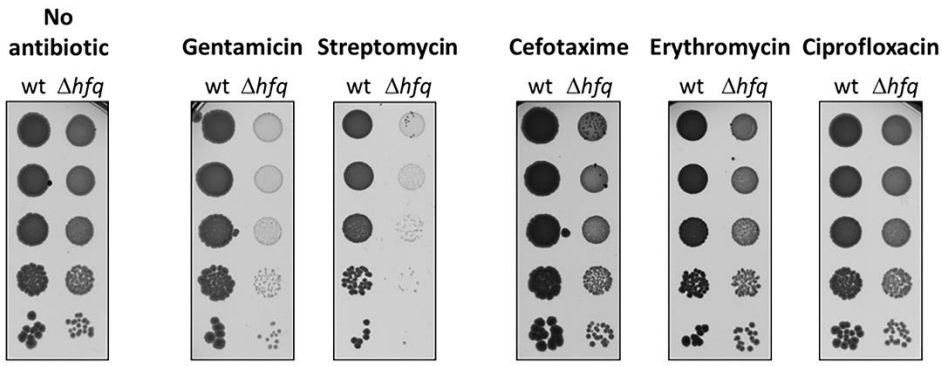


Figure S3 – Additional Δ*hfq* antibiotic sensitivity tested by serial dilution platting. Serial dilutions (1:10) of wild-type and Δ*hfq* strains grown on LB-agar plates at 37ºC with and without sub-lethal concentrations of gentamicin (0.1 µg/ml), streptomycin (1 µg/ml), cefotaxime (0.01 µg/ml), erythromycin (2 µg/ml) or ciprofloxacin (0.002 µg/ml).

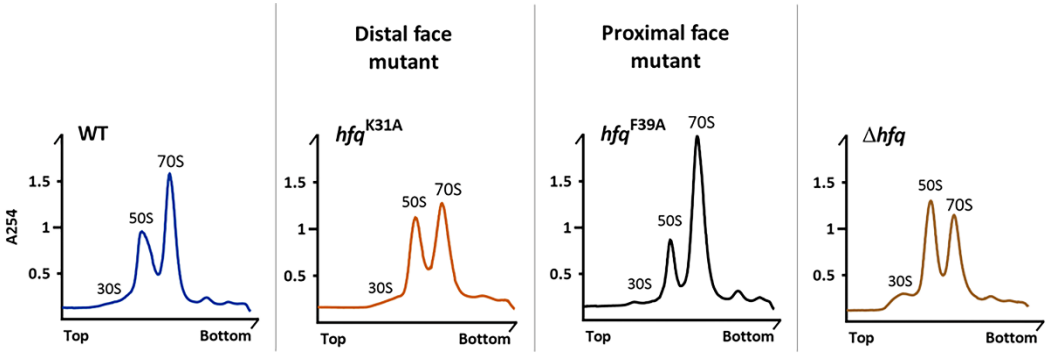


Figure S4 – Sucrose density gradients of the K31A and F39A Hfq mutants. Ribosomes purified from strains with specific point mutations in the *hfq* gene (alleles K31A and F39A) were resolved in 15-45% sucrose density gradients. Gradients from wild-type and *Δhfq* strains are included for comparison. The binding surface affected by each mutation is indicated.

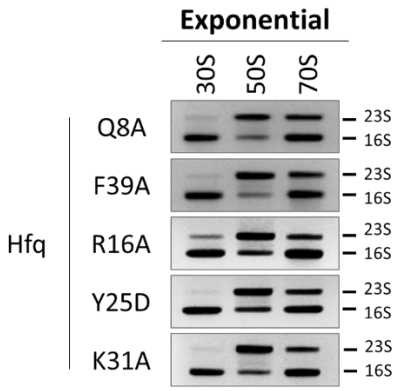


Figure S5 – Peak identification of ribosomes isolated by sucrose gradient in Hfq binding surface mutants. Fractions corresponding to 30S, 50S and 70S in Fig. 6 and S8 were collected and total RNA was isolated and separated on an agarose gel stained with ethidium bromide. The rRNA species are indicated.

Table S1 – List of strains used in this study.

Name	Relevant genotype	Reference
MG1693	*thyA rph-1*	Lab collection
MC1061	Δ(*lac*)X74	Lab collection
CMA540	MG1693 Δ*hfq::cat*	This work
CMA541	Δ(*lac*)X74 Δ*hfq::cat*	This work
CMA542	MC1061 Δ*hfq::cat-sacB* Δ*purA::kan*	This work
CMA543	MC1061 *hfq-Q8A*	This work
CMA544	MC1061 *hfq-F39A*	This work
CMA545	MC1061 *hfq-R16A*	This work
CMA546	MC1061 *hfq-Y25D*	This work
CMA547	MC1061 *hfq-K31A*	This work
JW3136-1	BW25113 Δ*rbfA::kan*	(Baba *et al*, 2006)
SG30200	PM1409 Δ*hfq::cat-sacB* Δ*purA::kan*	(Zhang *et al*, 2013)
SG30206	*lacI'::P_{BAD}-rpoS-lacZ hfq-Q8A purA+*	(Zhang *et al*, 2013)
SG30207	*lacI'::P_{BAD}-rpoS-lacZ hfq-R16A purA+*	(Zhang *et al*, 2013)
SG30209	*lacI'::P_{BAD}-rpoS-lacZ hfq-K31A purA+*	(Zhang *et al*, 2013)
SG30210	*lacI'::P_{BAD}-rpoS-lacZ hfq-F39A purA+*	(Zhang *et al*, 2013)
SG30237A	*lacI'::P_{BAD}-rpoS-lacZ hfq-Y25D purA+*	(Zhang *et al*, 2013)

Table S2 – List of plasmids used in this study.

Name	Description	Reference
pHFQ	*hfq* cloned under its own promoter in pBAD24	(Andrade *et al*, 2012)
pSG25	Wild-type copy of *lacZ*	(O'Connor & Dahlberg, 1993)
pSGlac7	+1 frameshift mutation near the 5' end of *lacZ*	(O'Connor & Dahlberg, 1993)
pSGlac10	-1 frameshift mutation near the 5' end of *lacZ*	(O'Connor & Dahlberg, 1993)
pSG163	UAG nonsense mutation near the 5' end of *lacZ*	(O'Connor & Dahlberg, 1993)
pSG3/4	UGA nonsense mutation near the 5' end of *lacZ*	(O'Connor & Dahlberg, 1993)
pSG413	AUG --> CUG start codon in *lacZ*	(O'Connor *et al*, 1997)
pSG416	AUG --> AUA start codon in *lacZ*	(O'Connor *et al*, 1997)

Table S3 – Oligonucleotide sequences used in various molecular biology techniques in this study.

Name	Oligonucleotide Sequence (5'-3')
Deletion mutant construction	
***hfq*-del-Fw**	CAGAATCGAAAGGTTCAAAGTACAAATAAGCATATAAGGAAAAGAGAGAA TG GTGTAGGCTGGAGCTGCTT
***hfq*-del-Rev**	GGAACGCAGGATCGCTGGCTCCCCGTGTAAAAAAACAGCCCGAAACCTTAG GTCCATAT GAATATCCTCCTTAG
Mutant confirmation by PCR and DNA sequencing	
***hfq*–confirm-1**	CGGTCAAACAAGCTTATAACCC
***hfq*–confirm-2**	GTGACGAAGAATTCCAGGTTGTTG
***rbf*A-confirm-1**	GGCTAACAGCCCCTTTTTTGTCAGGAG
***rbf*A-confirm-2**	GAGGACGACTCATTAGTCCTCCTTG
Northern blot probes	
16S-internal	CCCAGTAATTCCGATTAACGC
17S-5'	TTAAGAATCCGTATCTTCGAGTGCCCACA
17S-3'	TGTGTGAGCACTGCAAAGTACGCTTCTTTAAGGTAAGG
23S	CCTACACGCTTAAACCGGGAC
Primer extension	
primer 46	TCGACTTGCATGTGTTAGGC
PCR for in vitro transcription	
17S-5'-Fw	TAATACGACTCACTATAGTGTGGGCACTCGAAGATACGGATTCTTAACGTCG
17S-5'-Rev	AAAAGTTTGACGCTCAAAGAATTAAACTTCG
17S-3'-Fw	TAATACGACTCACTATAGCCTTAAAGAAGCGTTCTTTGAAGTGCTCACACA
17S-3'-Rev	TGTGTGAGCACTTCAAAGAACGCTTCTTTAAGG

Chapter 3

Hfq and RNase R interact and cooperate in a novel rRNA quality control pathway

This chapter was based on:

dos Santos, RF, Andrade JM, Pissarra, J, Deutscher MP and Arraiano; Hfq and RNase R interact and cooperate in a novel rRNA quality control-pathway. (in final preparation)

For this chapter I performed experiments and analyzed data and revised the manuscript.

3. Chapter 3: Hfq and RNase R interact and cooperate in a novel rRNA quality control pathway

3.1. Abstract

Hfq is a Sm-like RNA chaperone mostly known as regulator of small RNAs. However, we have previously showed that Hfq is also involved in the maturation of the 16S molecule and the 30S ribosomal subunit. The 3'-5' exoribonuclease RNase R is a highly processive hydrolytic enzyme that takes part in the processing of 16S rRNA as well as in various RNA quality control mechanisms. Here we unveil previously unrecognized functions of Hfq and RNase R in ribosomal RNA metabolism. We show that Hfq and RNase R can associate in a new complex and that both enzymes cooperate in a new quality control pathway that removes aberrant rRNA. Large fragments of 16S and 23S rRNA were shown to accumulate upon combined inactivation of Hfq and RNase R. Moreover, both RNA-binding proteins exhibit a synergistic effect in the maturation of 17S and pre-23S rRNA precursors, which is related with a marked decrease in the pool of 70S ribosomes. The accumulation of deleterious rRNA fragments and unprocessed rRNA precursors correlates with the defective growth phenotype in the Δ*hfq* Δ*rnr* mutant strain. The high conservation of the RNA-binding proteins Hfq and RNase R suggests a wider involvement for these proteins in rRNA metabolism and ribosome biogenesis, from prokaryotes to eukaryotes. Overall, this work highlights unprecedented roles for these RNA-binding proteins in multiple facets of ribosomal RNA biology.

3.2. Introduction

In prokaryotes, the 30S small ribosomal subunit contains an rRNA molecule (16S rRNA) and about 21 different ribosomal proteins (r-proteins), whereas the 50S large ribosomal subunit has two rRNA molecules (23S and 5S rRNA) and over 30 different r-proteins. The two asymmetric ribosomal subunits exist independently and associate to form the functionally active 70S ribosome (Shajani *et al*, 2011). The existence of quality control mechanisms that act at the level of ribosomal subunits thus preventing the assembly of defective ribosomes before they engage in active translation seems greatly advantageous for the cell. However, immature ribosomal subunits that escape such surveillance pathways may still be incorporated in 70S ribosomes. In *E. coli*, a quality control mechanism which involves the endonuclease YbeY and the 3'-5' exoribonuclease RNase R specifically recognises and degrades non-functional 70S ribosomes with 30S defective subunits (Jacob *et al*, 2013). In fact, RNase R is a unique exoribonuclease that seems to be at the centre of several RNA quality control pathways: 1) RNase R has been shown to be involved in the elimination of aberrant rRNAs fragments together with polynucleotide phosphorylase (PNPase) (Cheng & Deutscher, 2003); 2) to participate with RNase II in the extended degradation of ribosomes during starvation (Basturea *et al*, 2011); 3) to process the *tm*RNA involved in the *trans*-translation mechanism (Cairrao *et al*, 2003) and 4) to degrade non-stop mRNAs in stalled ribosomes (Richards et al., 2006; Domingues et al., 2014). Furthermore, RNase R can affect the expression levels of ribosomal proteins as it was shown to degrade the mRNA encoding ribosomal protein S15 (*rps*O gene) under conditions that favor RNA polyadenylation-dependent decay (Andrade *et al*, 2009a). The intimate relationship between RNase R and ribosomes is further underscored by the evidence that ribosomes regulate RNase R stability (Liang & Deutscher, 2013; dos Santos *et al*, 2018).

RNase R and the RNA-binding protein Hfq were found to interact with the same binding interface of the S12 protein in the 30S ribosomal subunit (Strader *et al*, 2013). Bacterial Hfq is a member of the Sm/Lsm superfamily of proteins with roles in the biology of small non-coding RNAs (Valentin-Hansen *et al*, 2004; Vogel & Luisi, 2011; Wilusz & Wilusz, 2013). It shows great affinity to 3'-end U-rich sequences of small RNAs along with

a preference for A/U-rich single-stranded RNA regions (Weichenrieder, 2014). rRNA had been previously found to coprecipitate with Hfq but this was usually regarded as a contaminant (Zhang et al., 2003; Sittka et al., 2008; Bilusic et al., 2014). However, it was recently shown that rRNA is in fact a substrate for the RNA chaperone Hfq during 16S rRNA maturation and folding, two important steps of ribosome biogenesis (Andrade *et al*, 2018).

In this work we describe a new interaction between Hfq and RNase R. Both RNA-binding proteins were found to cooperate in a previously unrecognized surveillance pathway that eliminates deleterious rRNA fragments from the cell. Additionally, double inactivation of Hfq and RNase R strongly affects 16S and 23S rRNA maturation. This correlates with an altered ribosome profile, leading to the accumulation of free subunits as well as a marked decrease of 70S ribosomes Overall, our results unveil a functional link between Hfq and RNase R in the quality control of rRNA and ribosome biogenesis.

3.3. Results

3.3.1. Hfq and RNase R engage in direct protein-protein interaction

The multifaceted RNA chaperone Hfq is known to be a ribosome biogenesis factor required for correct processing of 16S rRNA and formation of functional 70S ribosomes (Andrade *et al*, 2018). A genomic contextualization analysis of Hfq (*hfq*), performed through NCBI and STRING databases, revealed that it is located close to RNase R (*rnr*), a highly processive 3'-5' exoribonuclease (Figure 18). Both genes define a conserved genomic cluster interspaced in *E. coli* by 6 genes: *hflX* – ribosome-dissociating factor; *hflK* and *hflC* – regulators of FtsH protease; *yjeT* – uncharacterised protein; *purA* – adenylosuccinate synthetase; *nsrR* – DNA transcriptional repressor. This gene pattern is maintained among *Enterobacteriaceae* species as well as in other Gammaproteobacteria. Remarkably, the Hfq/RNase R genomic architecture is conserved in the genome of the metabolic diverse *Pseudomonas stutzeri*, of the acidophilic *Acidithiobacillus ferrooxidans*, and even in the small genome of the obligate endosymbiont *Baumania cicadellinicola*, arguing in favour of the importance of this module for bacterial homeostasis. Genes expressing proteins with related functions that are maintained in the close vicinity of each

other, like in the case of Hfq and RNase R, tend to encode interacting proteins (Snel *et al*, 2000). Consequently, we examined the possibility of a direct protein-protein interaction between Hfq and RNase R using a Far-western blot approach.

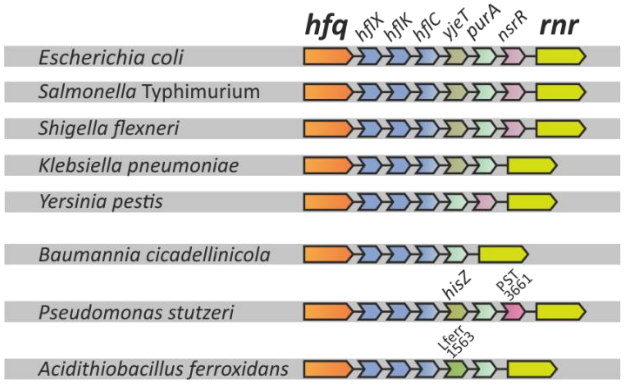


Figure 18 – Conserved genomic proximity *hfq* and *rnr*. Genomic proximity of the *hfq* and *rnr* genes in different Gammaproteobacteria. Genomes were analyzed with NCBI and STRING databases. *hfq* – RNA chaperone; *hfl*X – ribosome-dissociating factor; hflK/C – proteases; *yje*T – hypothetical protein; *pur*A – adenylosuccinate synthase; *nsr*R – transcriptional repressor; *rnr* – RNase R.

Increasing amounts of purified His_6-RNase R were separated on an SDS-PAGE gel and transferred onto a nitrocellulose membrane. Following an *in situ* renaturation of the immobilised proteins the membrane was incubated with purified His_6-Hfq protein in solution. After extensive washing steps we probed for the presence of Hfq bound to RNase R by using an anti-Hfq antiserum. Through chemiluminescence we could clearly detect Hfq in a well-defined band that corresponds to the RNase R location (~92kDa) (Figure 19A). Although increasing amounts of RNase R leaded to a stronger Hfq signal, we were still able to detect Hfq even when lower amounts (50ng) of RNase R were used, which suggests a strong interaction. Hfq and RNase R are two RNA-binding proteins with similar RNA substrates that could be mediating the interaction, even though protein purification was performed in the presence of Benzonase. To test this possibility, the purified proteins were subject to an extensive treatment with Micrococcal Nuclease (MNase) treatment in order to eliminate any potential contaminating nucleic acid. Consequently, a Far-western blot was performed showing that regardless of MNase treatment Hfq could still bind to

RNase R (Figure 19B). These findings confirm that Hfq and RNase R interact directly and not through an RNA third partner.

The Hfq/RNase R interaction was further confirmed by pulldown assays. Purified His_6-Hfq was bound to Ni-NTA beads and then incubated with total protein extracts. Hfq and its interaction partners were eluted with imidazole and separated on an SDS-PAGE. A Western blot performed with an anti-RNase R antibody revealed that Hfq was able to interact and pulldown native RNase R from the wild-type lysate (Figure 19C, left panel). A similar experiment was performed, but with the bait and prey proteins reversed, showing that purified His_6-RNase R was in its turn able to pulldown Hfq from total protein extracts (Figure 19C, right panel). Finally, the co-immunoprecipitation approach confirmed that both native proteins could interact. Accordingly, polyclonal antibodies raised against RNase R were immobilised to beads coated with Protein A/G and subsequently used to pulldown endogenous RNase R and its interacting partners from relevant different cell lysates. This enriched sample was then separated on an SDS-PAGE and Hfq presence was analysed by Western blot. Hfq co-immunoprecipitated with RNase R specifically from the wild-type proteic extract and it was absent from the control lysates of cells lacking Hfq or RNase R (Figure 19D). This is in line with a large-scale study that identified possible protein-protein interactions in *E. coli*, in which Hfq/RNase R complexes were predicted to occur (Butland *et al*, 2005).

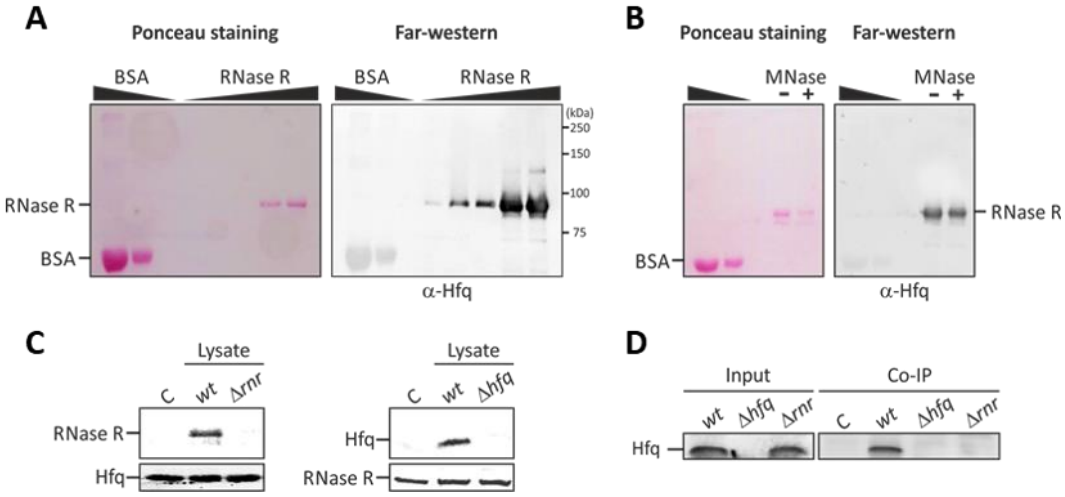


Figure 19 – Protein-protein interaction assays of Hfq/RNase R complex. (A) Far-western blot showing Hfq and RNase R interaction. Increasing amounts of purified RNase R (0.01, 0.05, 0.1, 0.25 and 0.5 µg) were resolved in a 10% SDS-PAGE and blotted to a nitrocellulose membrane. BSA (1 and 3 µg) was used as negative control. The membrane was stained with Ponceau red (left panel) prior to incubation with purified Hfq in solution (45 nM final concentration) and probing with an Hfq antibody (right panel). Ladder information and proteins are indicated on the sides. (B) Far-western blot of Hfq and RNase R samples treated with MNase. Purified RNase R either treated (+) or untreated (–) with Micrococcal Nuclease (MNase) was resolved in a 10% SDS-PAGE and blotted to a membrane. BSA (0.5 and 1 µg) was used as negative control. The membrane was stained with Ponceau red (left panel) prior to incubation with purified Hfq in solution (previously treated with MNase) and probing with an Hfq antibody (right panel). (C) Pulldowns of Hfq and RNase R. Purified His_6-RNase R or His_6-Hfq were used as the immobilized bait in Ni-NTA beads and incubated with cell lysates of *wt* or mutant strains or with binding buffer (this is indicated as C). Samples were analysed by Western blotting using Hfq (Ziolkowska *et al*, 2006) or RNase R (Cairrao *et al*, 2003) raised antibodies. (D) Co-immunoprecipitation of the Hfq/RNase R complex. A/G beads coated with anti-RNase R antibody were used to immunopurify the Hfq/RNase R complex from cell lysates of the wild-type, Δ*hfq* mutant or Δ*rnr* mutant strains. C – Beads incubated with Co-IP buffer were used as negative control

The N- and C-terminal domains of RNase R harbor the RNA-binding motifs that allow the enzyme to unwind RNA during degradation (Awano *et al*, 2010; Hossain *et al*, 2016; Chu *et al*, 2017). However, the RNB catalytic domain of RNase R alone is functional *in vitro* (Matos *et al*, 2009). In order to further understand how Hfq could bind to RNase R similar Far-western experiments were carried out using wild-type RNase R (RNR_WT) and a truncated RNase R holding only the catalytic RNB domain (RNR_RNB) (~50kDa) as the prey proteins (Figure 20A). Hfq is clearly detected in the wild-type RNase R, while no signal is detected in the RNB domain of RNase R (Figure 20B). This shows that the catalytic domain of RNase R is not sufficient for successful interaction with Hfq. Instead, contact between

the two proteins is suggested to require the flanking N- and C-terminal sequences of RNase R. However, at the moment we do not know whether both or only one of these sequences is required. Further work is necessary to assess the involvement of the CDS and S1 domains in the formation of the complex. Collectively, our results show that Hfq and RNase R directly interact with each other forming stable complexes in cellular lysates.

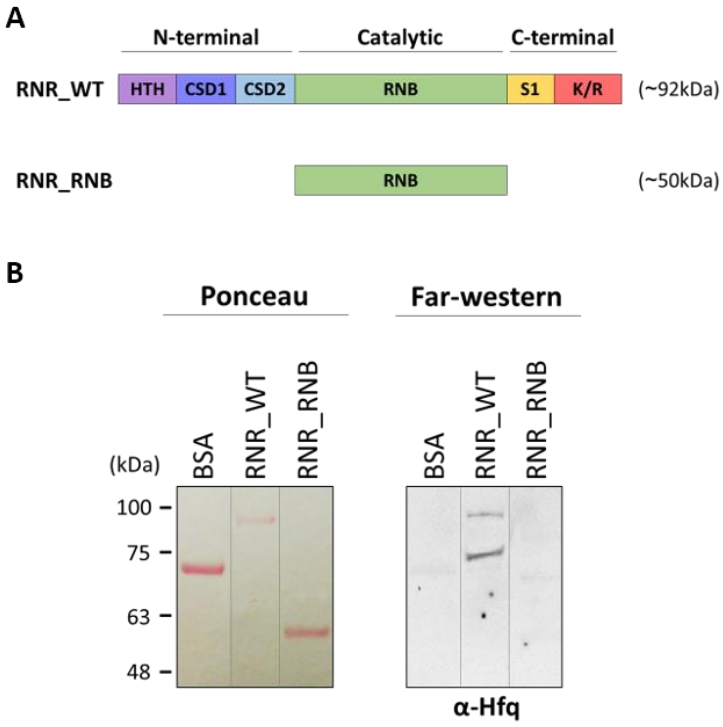


Figure 20 – Hfq interaction with wild-type RNase R and its RNB domain. Far-western blot showing Hfq and RNase R interaction. Wild-type (RNR_WT) or truncated (RNR_RNB) purified RNase R were resolved in a 4-12% SDS-PAGE and blotted to a nitrocellulose membrane. BSA was used as negative control. The membrane was stained with Ponceau red (left panel) prior to incubation with purified Hfq in solution (45 nM final concentration) and probing with an Hfq antibody (right panel). Ladder information and proteins are indicated on the side.

3.3.2. Growth defects and abnormal RNAs arise upon inactivation of Hfq and RNase R

Hfq and RNase R are two RNA-binding proteins with central roles in RNA biology, specifically in ribosomal RNA (Sulthana *et al*, 2016; Andrade *et al*, 2018). In order to investigate possible functional implications of the interaction between these enzymes, a

double deletion mutant (Δ*hfq* Δ*rnr*) was constructed. The allelic substitution of the wild-type *rnr* present in the Δ*hfq* background proved challenging due to the proximity between the genes. Regardless, an isogenic strain bearing both *hfq* and *rnr* chromosomal deletions was successfully obtained and confirmed by PCR and Western blot analysis (Figure 21).

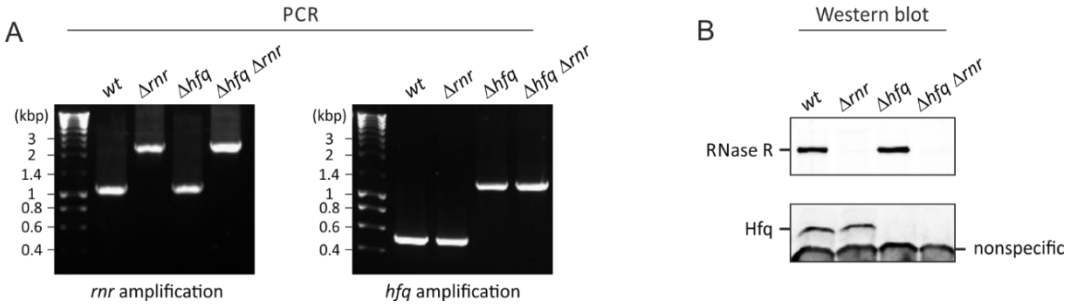


Figure 21 – Confirmation of the Δ*hfq* Δ*rnr* double mutant strain. (A) Deletion of *rnr* and *hfq* was confirmed by PCR using specific pair of primers to each gene. (B). The absence of RNase R and Hfq in the Δ*hfq* Δ*rnr* double mutant was confirmed by Western blot analysis using specific Hfq or RNase R polyclonal antibodies. Hfq antibody cross-reacts with a non-specific band that migrates below Hfq in SDS-PAGE gels (Ziolkowska *et al*, 2006).

The 24-hour growth profile in rich media of the Δ*hfq* Δ*rnr* strain along with its isogenic single mutants and parental strains was obtained by OD_{600} monitoring. As expected, the single Δ*hfq* mutant displayed a reduced growth rate and yield (Tsui *et al*, 1994) whereas the single Δ*rnr* mutant behaved similarly to the wild-type strain. Notably, the double mutant showed exacerbated growth defects with a marked increase of its doubling time – over 3 times higher than the wild-type (Figure 22A). Additionally, colony plating of these strains on LB-agar corroborated the growth difficulties of the Δ*hfq* Δ*rnr* strain, with a 2 to 3 log difference when compared to the wild-type (Figure 22B). Since both Hfq and RNase R are two important post-transcriptional regulators it was plausible that the growth defects could arise from a disturbance in RNA homeostasis. More specifically, because we previously characterised Hfq as a ribosome biogenesis factor, the concomitant inactivation of RNase R in a Δ*hfq* background could lead to defects in rRNA related processes. Therefore, total RNA was extracted from each strain and its integrity analysed on an ethidium bromide stained agarose gel (Figure 22C). The three bands corresponding to the rRNA species (23S, 16S and 5S) were identified in all strains.

Strikingly, additional bands migrating bellow the 16S rRNA were specifically present in the Δ*hfq* Δ*rnr* and not in the wild-type or single mutant strains. These RNA molecules must accumulate to high levels in order to be clearly visible under UV light in an ethidium bromide stained gel. This initial set of results clearly points towards a functional relationship for the conservation of the Hfq/RNase R genomic cluster, since disruption of both genes leads to marked growth defects and abnormal accumulation of RNA species.

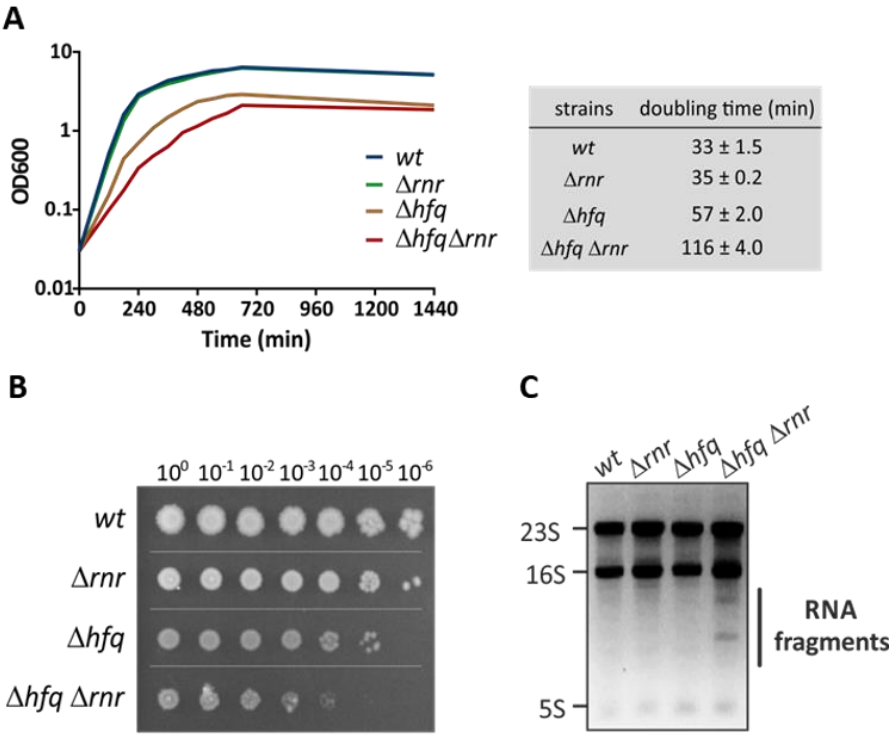


Figure 22 – Impact of the inactivation of *hfq* and *rnr* genes in the cell. (A) Growth curve of the Δ*hfq* Δ*rnr* mutant strain compared to the parental strain (wt) and single mutant strains (Δ*hfq* and Δ*rnr*). Cells were grown on LB medium at 37ºC. The doubling time of each strain is shown on the side. (B) Platting assay. 10-fold serial dilutions of cell cultures were prepared and spotted on a LB plate and grown overnight at 37ºC. (C) Total RNA was fractionated in an agarose gel and stained with ethidium bromide. The ribosomal RNAs and accumulating RNA fragments are indicated.

3.3.3. rRNA fragments accumulate in the absence of Hfq and RNase R

The double inactivation of Hfq and RNase R leads to the accumulation of RNA species of unknown origin, which suggest that both proteins cooperate in the elimination of these molecules. Since these RNAs accumulate to high levels we reasoned that they could derive from rRNA. In order to test this, Northern blot analysis was performed using radioactive oligonucleotide probes complementary to the central regions of the 16S, 23S and 5S rRNAs.

Total RNA from exponential and stationary phase cells was fractionated on a 7M urea denaturing polyacrylamide gel and transferred onto a nitrocellulose membrane to be subsequently probed for each of the three rRNAs. Immediately we could see that the probes could detect the full-length rRNAs. Additionally, lower molecular weight RNA molecules were also detected in the single and double mutants when using the 16S and 23S rRNA probes, indicating that these are rRNA-derived fragments (Figure 23). No differences in neither the single mutants nor the double mutant were found when the 5S rRNA antisense probe was used. Two different rRNA fragments, with approximately 200nts and 350nts each, were identified when using the 16S probe (Figure 23A), whereas one rRNA fragment of approximately 300nts was detected in the 23S rRNA analysis (Figure 23B). These fragments arise in the single and double mutant strains independently of the growth phase analysed and are entirely absent in the wild-type. However, they accumulate to much higher levels in the double Δ*hfq* Δ*rnr* strain – over 4 times higher when compared to the mean signal intensity of the single mutants during exponential growth.

In addition, we were able to map and delimit the origin of the fragments by using different antisense probes that would hybridise either the 16S or 23S rRNA regions flanking the central region initially tested. Thus, the 16S rRNA-derived fragments originate from the region delimited by nucleotides 559 to 969 of the full-length rRNA (Figure S6), whereas the 23S rRNA-derived fragment originates from the region between nucleotides 1290 and 1640 (Figure S7). Since the fragments come from the middle regions of the highly structured rRNAs it is plausible that they arise from endonucleolytic cleavages and require both Hfq and RNase R to be subsequently degraded. Curiously, we could detect

another rRNA fragment of approximately 450nts accumulating in the Δ*hfq* Δ*rnr* strain when a probe antisense to the 970-990nts region of the 16S rRNA was used, revealing that additional rRNA fragments may be present.

These results indicate that the absence of Hfq or RNase R triggers the appearance of rRNA fragments derived from both 16S and 23S rRNAs that are not eliminated, and that the presence of one enzyme cannot compensate the lack of the another. More importantly, inactivation of both proteins leads to an intense accumulation of rRNA fragments as observed in the Δ*hfq* Δ*rnr* strain, suggesting a strong impairment of their degradation. The fact that this accumulation pattern was observed in both exponential and stationary growth stages shows that Hfq and RNase R activity on rRNA removal is independent of the growth phase and likely to happen throughout bacterial growth. Collectively, our findings suggest that Hfq and RNase R exert a synergistic effect on the degradation of rRNA fragments and that they share the genetic pathway for their removal.

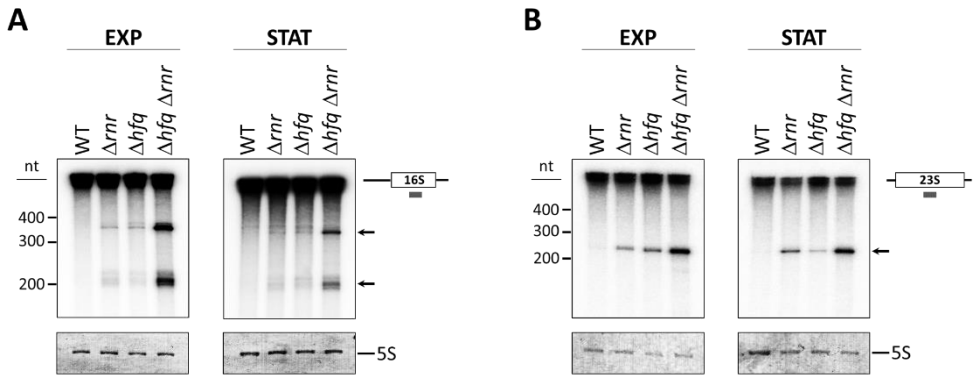


Figure 23 – RNase R and Hfq are involved in 16S and 23S rRNA fragments degradation. (A) Analysis of 16S rRNA fragments. (B) Analysis of 23S rRNA fragments. RNA extracted from exponential (left panels) or stationary (right panels) phase cultures were analyzed by Northern blotting in 8% polyacrylamide/7M urea gels. 5S rRNA levels detected with methylene blue dye are shown below each panel. Arrows indicate specific rRNA fragments.

3.3.4. RNase R is the exoribonuclease that specifically cooperates with Hfq for eliminating rRNA fragments

RNase R is one of three major 3'-5' exoribonucleases in *E. coli*, alongside RNase II and PNPase. The latter two enzymes had already been described to pair up with RNase R for the degradation of rRNA in different conditions (Cheng & Deutscher, 2003; Basturea *et al*, 2011). To study whether these two additional exoribonucleases could be participating this novel rRNA degradation pathway we examined if inactivation of both RNase II (*rnb*) or PNPase (*pnp*) alone or together with *hfq* deletion affected the accumulation pattern of the fragments.

Northern blot analysis of the single mutant strains showed that the previously detected rRNA species arise specifically upon RNase R inactivation and not when RNase II nor PNPase are missing (Figure 24). A similar analysis was made for cells lacking each one of the exoribonucleases in a Δ*hfq* background. Similarly, we detected no significant differences on the Δ*hfq* Δ*rnb* and Δ*hfq pnp-7* double mutants when comparted to the Δ*hfq* strain, except for the expected accumulation levels due to Hfq inactivation (Figure 24).

These results show that RNase R is the major 3'-5' exoribonuclease involved in the elimination of these rRNA-derived fragments and cannot be replaced by the action of RNase II or PNPase. The fact that inactivating either RNase II or PNPase in a Δ*hfq* background provoked no alteration of the Δ*hfq* rRNA accumulation pattern indicates that Hfq is cooperating specifically with RNase R for the critical elimination of these aberrant RNA molecules.

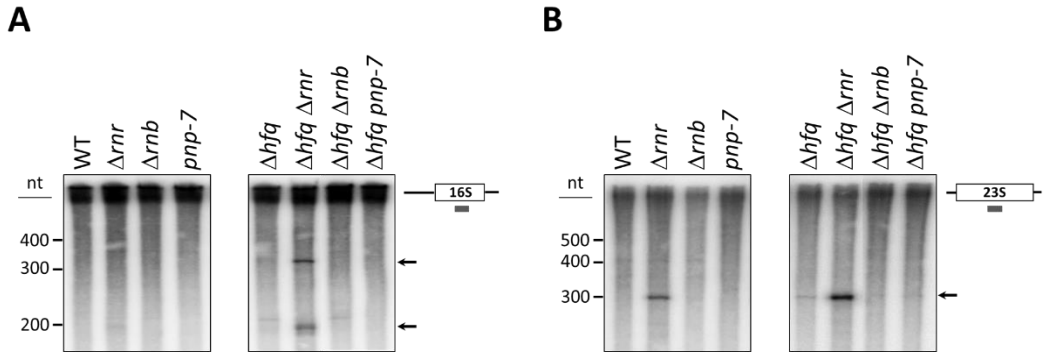


Figure 24 – RNase R and Hfq are specifically required for rRNA fragment elimination. (A) Analysis of 16S rRNA fragments. (B) Analysis of 23S rRNA fragments. RNA extracted from stationary phase cultures were analyzed by Northern blotting in 8% polyacrylamide/7M urea gels. Arrows indicate specific rRNA fragments.

3.3.5. rRNA processing is affected by the double inactivation of Hfq and RNase R

The presence of highly abundant rRNA fragments is deleterious to the cell, since their accumulation prevents ribonucleotide recycling and can lead to a detrimental competition for r-protein binding (Cheng & Deutscher, 2003). Hfq together with RNase R cooperate to eliminate such hazards and that the double inactivation of these enzymes leads to severe growth defects. Hfq is a ribosome biogenesis factor that assists in 16S rRNA processing and folding, whereas RNase R is one of the exoribonucleases responsible for 3'-end processing of 16S rRNA (Sulthana & Deutscher, 2013; Andrade *et al*, 2018). The precursor of 16S rRNA – termed 17S rRNA – has 115 extra nts at the 5' end and 33 extra nucleotides that need to be trimmed down to give raise to the mature 16S rRNA. Similarly, the pre-23S precursor has 7 nts and 9 nts extra at the 5' and 3' end, respectively, that are also eliminated during 23S maturation. In both cases the unprocessed extra sequence at one end can basepair and form a helical structure with the other at the opposite end (Figure 25A and Figure 25B). This secondary structure is believed to protect against ribonucleases obstructing the correct processing into the mature molecules.

Since both proteins were shown to be functional during rRNA maturation and share 17S rRNA as a common substrate, we reasoned that cooperation between Hfq and RNase R could extend to the maturation of rRNA molecules. To examine this point, we

used Northern blot analysis to assess the levels of these rRNA precursors in the single and double mutant strains. Total RNA extracted from each strain was fractionated on a 4% polyacrylamide denaturing gel ran overnight to allow a good separation of the 23S, 17S and 16S rRNAs. Using oligonucleotide probes that hybridize with the extra nucleotides present at the 5' and 3' ends of each rRNA precursor we were able to detect and compare the precursor levels between strains. 17S rRNA precursor accumulates in the Δ*hfq* strain as described (Andrade *et al*, 2018). However, higher levels of 17S rRNA are present in the double mutant. In sharp contrast, only low levels of 17S rRNA were detected in the wild-type and Δ*rnr* strains (Figure 25C). These results indicate that Hfq and RNase R are involved in the processing of 17S rRNA, with a predominant role of Hfq in this process. Similarly, the two probes specifically directed against the 23S rRNA precursor identified a substantial accumulation of the pre-23S rRNA in the Δ*hfq* Δ*rnr* mutant when compared to the wild-type or single mutants (Figure 25D).

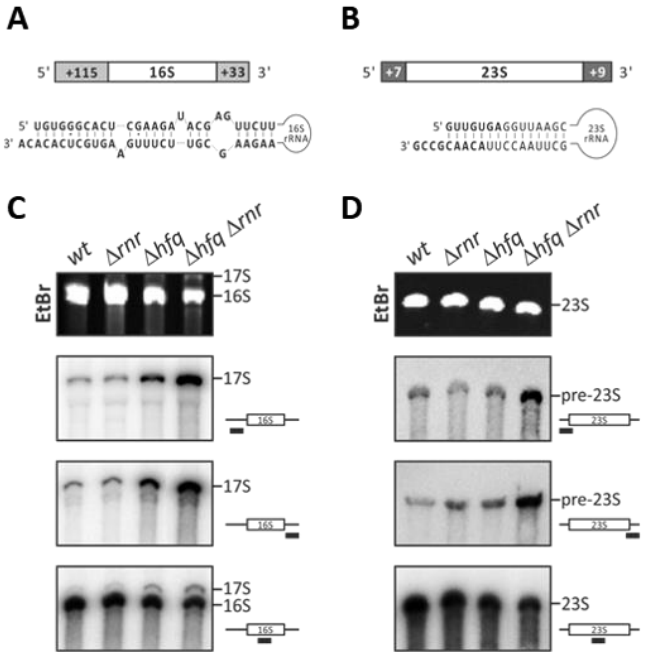


Figure 25 – 17S and pre-23S Precursors accumulate in the Δ*hfq* Δ*rnr* mutant. (A and B) Structure of the 17S and pre-23S rRNA, respectively. Mature rRNAs and extra nucleotides (in bold) are depicted. (C and D) 4% polyacrylamide gel and Northern hybridization analysis of the 17S rRNA and pre-23S rRNA, respectively. Total RNA was extracted from *wt*, Δ*rnr*, Δ*hfq* and Δ*hfq* Δ*rnr* stationary phase cultures. Prior to the membrane transfer, the gels were stained with ethidium bromide. Specific probes directed against the mature rRNA or the 3'-end or 5'-end flanking sequences were used in northern blotting hybridizations.

There are only 7 extra nucleotides present in the 5’ end of the pre-23S rRNA precursor. Therefore, we used primer extension analysis to map the 5’ end of the 23S rRNA and confirmed that high levels of immature 23S rRNA precursor were specifically found in the Δ*hfq* Δ*rnr* (Figure 26A). In addition, an overexposure of the membranes hybridized with the probes directed against the 3’ end of either the 17S rRNA or pre-23S rRNA revealed that several shorter rRNA species carrying the 3’ end precursor sequence accumulated specifically in the Δ*hfq* Δ*rnr* mutant but not on the wild-type or single mutants (Figure 26B). The data show that Hfq and RNase R are further required for the elimination of rRNA fragments that retain the 3’ end of precursor sequences. These apparently originate from endonucleolytic cleavages in the mature rRNA region, although the identity of the endoribonuclease is still unknown (Basturea et al., 2011). Such fragments were not detected when 5’ end probes were used.

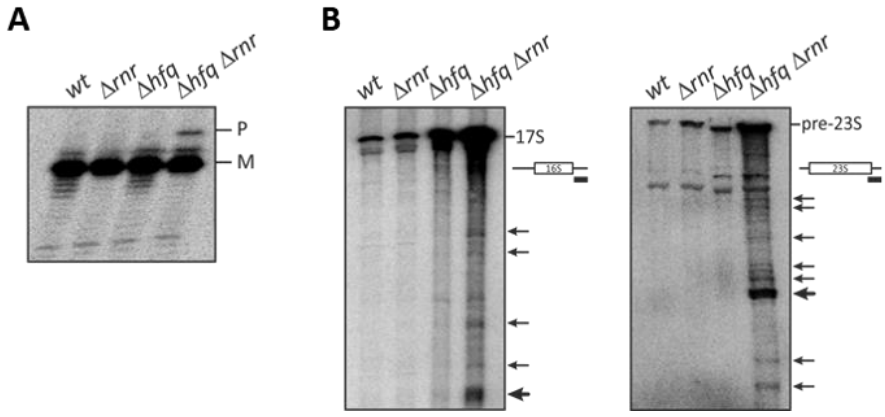


Figure 26 – Primer extension of the 23S rRNA and accumulation of precursor fragments. (A) Primer extension analysis of pre-23S 5’ end. Total RNA was extracted from stationary phase cultures. Mature (M) and precursor (P) forms are indicated. The Δ*hfq* Δ*rnr* accumulates pre-23S rRNA. (B) Northern blot analysis of rRNA precursor fragments. Specific 3’ end probes for the 17S (left) or pre-23S (right) rRNAs were used. Fragments which accumulate specifically in the double mutant are indicated by arrows on the side of the gel

Since Hfq and RNase R are involved in the processing of the 3’ end of rRNA precursors, depletion of these two proteins explains the accumulation of rRNA precursors found in the Δ*hfq* Δ*rnr* mutant strain (Sulthana & Deutscher, 2013; Andrade *et al*, 2018). However, we found that the 17S and pre-23S rRNA precursors present in the Δ*hfq* Δ*rnr* mutant are not correctly processed at the 5’ end as well. It seems rather plausible that as

result of misprocessing of the 3' end of precursor rRNAs, this sequence becomes more likely to basepair with the respective 5' end sequence. This would lead to the formation of an extended terminal stem loop that can impair the processing of the 5' end as well, as it was previously shown (Liang and Deutscher, 2013). Collectively, these results demonstrated that the absence of Hfq and RNase R leads to severe defects in the maturation of rRNAs.

3.3.6. Hfq and RNase R cooperate during ribosome biogenesis

The 17S rRNA and the pre-23S rRNA precursors retain the ability to associate with r-proteins in pre-ribosomes. However, the extra nucleotides present in rRNA precursors affect the folding of the rRNA molecule which induces conformational changes in the ribosomal subunits (Liiv & Remme, 2004; Roy-Chaudhuri *et al*, 2010). Accordingly, defects in rRNA processing can lead to defects in ribosome biogenesis (Shajani *et al*, 2011). We next analyzed if the accumulation of rRNA precursors detected in the Δ*hfq* Δ*rnr* mutant strains was correlated with changes in the ribosome profile.

Ribosomes isolated from stationary phase cultures of the wild-type and mutant strains were analyzed and compared using sucrose gradient ultracentrifugation. Under conditions that favor ribosome association (10 mM Mg^{2+}), the peak corresponding to the 30S small subunits were nearly absent and 70S ribosomes are predominant in the wild-type (Figure 27A). Identical results were obtained with the RNase R mutant. The Hfq mutant revealed an increase in the free 30S subunit population and a reduction in the amount of 70S ribosomes, as expected. rRNA species identified in the gradient fractions further confirmed the identity of the different ribosome peaks (Figure S8). Strikingly, the combined inactivation of Hfq and RNase R resulted not only in the increase of the 30S subunit peak as well as in the marked elevation of the 50S subunit peak and in the simultaneously reduction of the 70S particles. Indeed, the higher peak in the Δ*hfq* Δ*rnr* strain actually corresponds to the 50S subunit population rather than the 70S ribosomes, as confirmed by the analysis of the RNA in the gradient fractions. Furthermore, similar results were obtained when exponentially growing cells were analyzed (Figure 27C). These

results demonstrated that Hfq and RNase R are critical players in the ribosome biogenesis program regardless of the growth stage analyzed.

To better compare the total amount of 30S and 50S subunits in the different strains, we next analyzed these ribosomes under dissociative conditions (0.1 mM Mg^{2+}) in order to guarantee that all the ribosomal subunits would be in their free state. As observed in Figure 5B, the different strains in stationary phase cultures have comparable content of 30S and 50S subunits. Again, similar results were obtained when exponentially grown cells were analyzed (Figure 27B). As a result, no significant differences were found between the levels of the subunits of wild-type and Δ*hfq* Δ*rnr* strains under ribosome dissociative conditions, independent of the growth phase analysis (Figure 27B and Figure 27D). Altogether, these findings suggest that the lower amounts of 70S ribosomes that are found in the double mutant do not result from differences in the pool of available subunits but are rather consequence of defects in the assembly of 70S ribosomes.

The abnormal ribosome profile observed in the Δ*hfq* Δ*rnr* mutant suggested that inactivation of Hfq and RNase R alters the pool of functional ribosomes *in vivo*. To test this hypothesis, the polysome profiles of the parental and Δ*hfq* Δ*rnr* strains were compared. Cell lysates of exponential phase cultures were analyzed using sucrose gradient ultracentrifugation. Consistent with the ribosome profiles (Figure 27A and Figure 27C) the Δ*hfq* Δ*rnr* mutant showed increasing levels of the free subunits and a reduction in the amount of 70S ribosomes (Figure 27E). Moreover, the Δ*hfq* Δ*rnr* mutant seems to form fewer polysomes than the wild-type strain. Four polysome peaks were detected in the wild-type (polysome/70S ratio of 0.25) whereas only three polysome peaks were detected in the Δ*hfq* Δ*rnr* mutant (polysome/70S ratio of 0.17). These results support that the ribosome biogenesis defects observed in cells lacking Hfq and RNase R can cause translational defects.

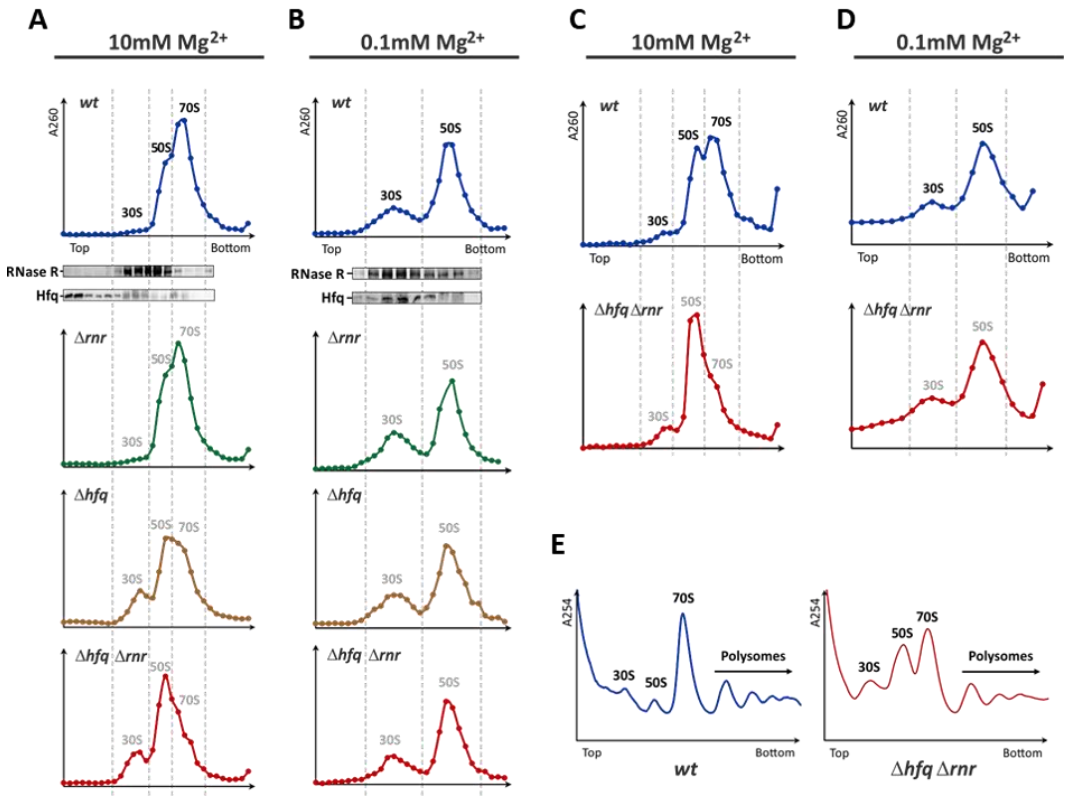


Figure 27 – Sucrose Density Gradient Analysis Reveals Defects in Ribosome Assembly. (A) 15-50% Sucrose density gradients of ribosomes extracted from stationary phase cells analyzed under associative conditions (10 mM Mg^{2+}) to promote the recovery of 70S ribosomes. The presence of RNase R and Hfq in gradient fractions of the wild-type was tested by western blotting using specific antibodies. Ribosome particles are indicated on top of each peak. (B) 10-30% Sucrose density gradients of ribosomes extracted from stationary phase cells analyzed under dissociative conditions (0.1 mM Mg^{2+}) to promote the complete dissociation of 70S ribosomes into the free 30S and 50S subunits. The presence of RNase R and Hfq in gradient fractions of the wild-type was tested by western blotting using specific antibodies. (C) 15-50% Sucrose density gradients of ribosomes extracted from exponential phase cells under associative conditions (10 mM Mg^{2+}). (D) 10-30% Sucrose density gradients of ribosomes extracted from exponential phase cells under dissociative conditions (0.1 mM Mg^{2+}). (E) Polysome analysis through a 10-40% sucrose density gradient (after chloramphenicol treatment of exponentially growing cells) performed in associative conditions (10 mM Mg^{2+}). Ribosome particles are indicated on top of each peak

Hfq and RNase R were found to interact with the S12 protein of the small ribosomal subunit in stationary phase cells (Strader *et al*, 2013). To assess the distribution of RNase R and Hfq in ribosomes, individual fractions of the associative ribosome profile of the wild-type strain were recovered and their protein content was analyzed by Western blotting using specific antibodies against either RNase R or Hfq (Figure 27A, upper panel).

This experiment showed most of RNase R to be associated with the free subunits, but not with the 70S ribosomes. This data is in agreement with work in which RNase R was shown to associate with 30S and pre-50S particles in exponentially growing *E. coli* cells (Chen & Williamson, 2013; Malecki *et al*, 2014). On the other hand, most of Hfq is free, whereas a small fraction of Hfq was bound to 30S particles and not to the 50S subunits (Figure 27A, lower panel). Analysis of fractions from the dissociative ribosome profile of the wild-type strain (Figure 27B, lower panel) confirmed that RNase R associates with both the 30S and 50S subunits while Hfq interacts mostly with 30S subunits. Overall, our results confirm that Hfq and RNase R associate with ribosome particles and can copurify with free 30S subunits. Consequently, the interaction between Hfq/RNase R/ribosome subunits is shown to be important for assembly of the ribosome.

3.4. Discussion

In this work we have shown that Hfq and RNase R are key factors in the processes of ribosome biogenesis and rRNA quality control. The 3'-5' exoribonuclease RNase R is known to modulate mRNA stability as well as the degradation and processing of rRNA (Cheng and Deutscher, 2003, 2005; Andrade et al., 2006; Andrade et al., 2009b; Sulthana and Deutscher, 2013). Hfq is an RNA-chaperone and a post-transcriptional regulator widely known for playing important roles in small RNA biology (Vogel and Luisi, 2011; Hajnsdorf and Boni, 2012). However, rRNA constitutes the large majority of RNA molecules in a cell but only recently it shown that 16S rRNA is an important substrate for Hfq (Andrade *et al*, 2018). The work presented herein, unveils previously unrecognized functions of Hfq in rRNA metabolism and provides additional insight to the recently described role as a ribosome biogenesis factor.

Processing of rRNA is an intricate process that requires the action of multiple enzymes (reviewed in Deutscher, 2009). rRNA precursors can induce conformational changes in the structure of the ribosomal subunits; these immature particles are defective and lead to reduced translational fidelity by 70S ribosomes (Liiv & Remme, 2004; Yang *et al*, 2014). In the Δ*hfq* mutant accumulation of 17S rRNA is consistent with increasing amounts of immature 30S small subunits, as previously described (Andrade *et al*, 2018).

On the other hand, inactivation of RNase R alone does not significantly affect the maturation of rRNA or the ribosome profile. This is probably because RNase R activity on the processing of rRNA is compensated by other exoribonucleases in cells (Sulthana & Deutscher, 2013). We found that Hfq and RNase R have a synergistic effect in the maturation of rRNA and ribosome biogenesis. Combined inactivation of Hfq and RNase R causes the strong accumulation of not only the 17S rRNA but also the pre-23S rRNA. Notably, 17S accumulation in the double mutant was much higher than observed in the Δ*hfq* mutant. This shows that RNase R is active in rRNA processing even in the absence of Hfq. Moreover, it indicates that impairment of Hfq and RNase R is not compensated by any of the remaining rRNA maturation factors in the cell. Such a surveillance mechanism that is active in the maturation of rRNA precursors is highly advantageous for the cell, acting as a quality control pathway on subunit production, preventing incorrect subunit assembly and consequently translational errors. Therefore, there is a strong correlation between defects in rRNA maturation and changes in ribosome profiles in the Δ*hfq* Δ*rnr* strain, which exhibits the slow-growth phenotype typically found in strains with perturbed ribosome biogenesis cascades (Connolly et al., 2008; Sharpe Elles et al., 2009; Leong et al., 2013). Comparing to wild-type or single mutant strains, it is clear that inactivation of Hfq and RNase R causes a strong reduction in the levels of 70S ribosomes along with a marked increase not only in the population of immature 30S (as observed in the Δ*hfq* mutant) but also in immature 50S ribosomal subunits. Despite an apparent normal ratio of ribosomal subunit production, these do not seem to associate into 70S ribosomes. This finding points to the existence of severe assembly defects in the Δ*hfq* Δ*rnr* double mutant strain that prevent the association of the small and large ribosomal subunits into the active ribosome particle.

Several assembly factors associate with the ribosome and participate in the intricate process of ribosome assembly *in vivo*, assisting in various r-protein binding and rRNA folding steps (Kaczanowska & Rydén-Aulin, 2007). RNase R was found to associate with the free subunits but not with the 70S ribosomes. This finding is in agreement with a quantitative mass spectrometry study that identified RNase R as a novel ribosome assembly factor in exponentially growing *E. coli* cells (Chen & Williamson, 2013).

Moreover, our data is also consistent with Hfq function as a ribosome assembly factor that associates with 30S particles (Andrade *et al*, 2018). Altogether, our results are in line with the report that Hfq and RNase R can bind to the r-protein S12 in the small subunit in stationary phase *E. coli* cells (Strader *et al*, 2013). S12 is a key mediator of fidelity of translation in both prokaryotes and eukaryotes and is positioned in helix 44 of the 16S rRNA that is known to form extensive contacts with the large subunit (Yusupov et al., 2001; Cukras et al., 2003). Consequently, association of Hfq and RNase R with S12 is suggested to be important for the correct folding of helix 44 of 16S rRNA thus affecting the formation of interface bridges between ribosomal subunits and consequently the assembly of 70S ribosomes.

We have also found that Hfq and RNase R play additional roles in an rRNA quality control mechanism. We have shown that Hfq and RNase R are partners in a hitherto unrecognized surveillance pathway that eliminates defective rRNA fragments. Removal of defective rRNA species is a crucial step in RNA quality control as their accumulation can lead to cell death (Sulthana *et al*, 2016). Single inactivation of either Hfq or RNase R resulted in a mild accumulation of such rRNA fragments. This indicates that in the absence of Hfq, RNase R cannot fully eliminate these aberrant rRNAs, despite its unique ability to degrade structured RNA molecules (Matos *et al*, 2009; Vincent & Deutscher, 2009; Chu *et al*, 2017). Conversely, rRNA fragments also accumulate in cells lacking RNase R, meaning no other ribonuclease seems able to compensate for its absence. Therefore, it is plausible that the RNA chaperone Hfq acts to remodel the highly structured rRNA molecules providing substrates that become suitable RNase R substrates. This is underscored by the fact that double deletion of Hfq and RNase R leads to a marked accumulation of the deleterious 16S and 23S rRNA fragments, regardless of the growth phase analyzed (Figure 23). Notably, we observe that the levels of 23S-derived fragments in stationary phase Δ*hfq* cells are lower (Figure 23B, right panel), arguably because RNase R levels increase during stationary growth, which could lead to a more effective degradation of these fragments. Additionally, none of the other major 3'-5' exoribonucleases seem to be active in this quality control pathway, since inactivation of RNase II and PNPase had no effect on the accumulation of these aberrant rRNAs (Figure 24A and Figure 24B, left panels). Moreover,

the double inactivation of Hfq and RNase II as well as Hfq and PNPase showed no significant effects, meaning that RNase R is able to eliminate most of the fragments with the specific cooperation of Hfq (Figure 24A and Figure 24B, right panels). The possibility of Hfq and RNase R functioning in a complex cannot be ruled out, since both proteins can directly interact with each other.

We show that Hfq and RNase R associate without any RNA mediating the interaction (Figure 19), which argues for the function of a previously unrecognized complex in a new rRNA quality control surveillance mechanism. Moreover, preliminary results point to the importance of the N- and C-terminal domains of RNase R in the formation of this complex. Far-western analysis showed that Hfq is able to bind wild-type RNase R (RNR_WT), but not to a truncated baring only the catalytic RNB domain (RNR_RNB) (Figure 20). Both N- and C-terminal regions contain RNA-binding domains (CSD1+2 and S1) that were recently shown to cap the top of the RNB catalytic domain (Chu *et al*, 2017). The importance of these domains not only to the unwinding activity but also to the specificity of substrate recognition by RNase R had been previously reported (Vincent & Deutscher, 2006; Matos *et al*, 2009; Hossain *et al*, 2016; Chu *et al*, 2017; dos Santos *et al*, 2018). Hence, these RNA-binding motifs do not allow the substrate to directly contact with the catalytic center, unless a single stranded 3' stretch with at least 7 nucleotides is present in the RNA substrate. Hence, elimination of highly structured RNA fragments derived from the 16S and 23S rRNAs, may prove difficult even for RNase R. These rRNA fragments may still retain extensive secondary structures that occlude the 3' end requirement. Association of Hfq with the RNA-binding domains of RNase R may allow the RNA-chaperone to act on these fragments and provide a 3' linear stretch to which RNase R can bind and degrade.

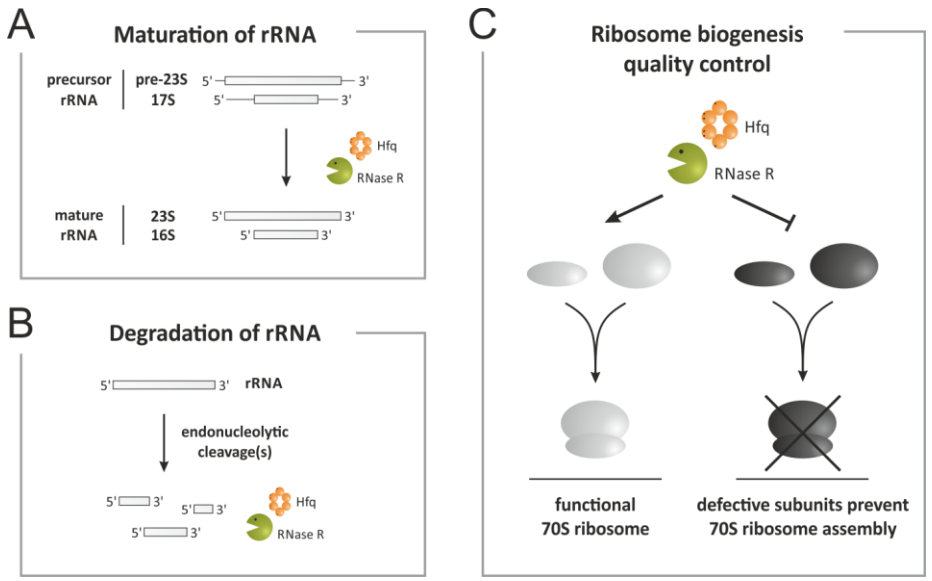


Figure 28 – Model for the Intricate Roles of Hfq and RNase R on Ribosome Biogenesis and rRNA Degradation. (A) Hfq and RNase R are required for the correct processing of the 16S and 23S rRNAs. In the absence of Hfq and RNase R, the rRNA precursors are not properly processed and accumulate to high levels in the cell. (B) Hfq and RNase R are key factors in a novel rRNA quality control mechanism that eliminates abnormal rRNA fragments from the cell. These defective rRNA fragments probably arise from endonucleolytic cleavages. No other major exoribonuclease seems to be involved in this RNA degradation pathway. (C) Hfq and RNase R are required for accurate ribosome assembly. Upon Hfq and RNase R inactivation, the levels of intact 70S ribosomes are strongly reduced which correlate with a marked increase in the population of free 30S and 50S subunits that are unable to associate. This finding strongly suggests the existence of severe defects in the ribosome assembly in the absence of Hfq and RNase R.

Overall, this work expanded the known functions of Hfq and RNase R and revealed a new quality control mechanism in which both proteins are crucial (Figure 28). We have unveiled a synergistic role between these two RNA-binding proteins during rRNA maturation and ribosome biogenesis. The basic program of ribosome biogenesis and much of the machinery involved is conserved among bacteria and eukaryotes (Hage and Tollervey, 2004; Dinman, 2009). The high conservation of the RNA-binding proteins Hfq and RNase R suggests thus a wider involvement for these proteins in rRNA metabolism throughout evolution.

3.5. Materials and Methods

3.5.1. Bacterial strains

All experiments use derivatives of *E. coli* K-12 strain MG1693 (Andrade *et al*, 2012). Deletion mutants for *hfq* and *rnr* were constructed by the one step inactivation of chromosomal genes method (Datsenko & Wanner, 2000) with details described elsewhere (Andrade *et al*, 2018). A linear PCR product amplified from plasmid pDK3 with primers hfq-del-Fw/hfq-del-Rev was transformed into Δ*rnr* strains carrying the helper plasmid pDK46. The resulting Δ*hfq* Δ*rnr* deletion mutant was selected in the presence of chloramphenicol and tested by PCR and Western blotting. P1-mediated transduction was used to transfer mutation to a fresh *E. coli* wild-type background.

3.5.2. Bacterial growth

Bacteria were grown at 37ºC in Luria-Bertani (LB) medium supplemented with thymine (50 μg/ml). Antibiotics were present at the following concentrations when needed: 25 μg/ml for chloramphenicol, 25 μg/ml for kanamycin, 10 μg/ml for tetracycline and 100 μg/ml for ampicillin. In the dilution plating assays, serial dilutions were made in 10-fold increments and immediately spotted onto LB agar plates using a replica plater. The plates were incubated at 37ºC for ~36h. Strains were inoculated in 50 mL from overnight cultures of isolated fresh grown colonies to an initial OD_{600} ~ 0.03. Cells were incubated with orbital shaking and culture density was monitored by measuring the optical density. Cultures were collected either at exponential phase (OD_{600} ~ 0.5) or stationary phase (OD_{600} ~ 5.5 for *hfq*$^+$ strains or OD_{600} ~ 2.5 for Δ*hfq* mutants). In the growth curves, cell doubling times were calculated as DT = (t2-t1) x [log2/ (log OD_{600}(t2) – logOD_{600}(t1)], using the period of exponential growth.

3.5.3. Protein purification

Purification of His_6-RNase and His_6-Hfq was performed by histidine-affinity chromatography using HiTrap Chelating HP columns (GE Healthcare) and AKTA FLPC system (GE Healthcare) following the protocols described previously (Matos *et al*, 2009;

Folichon *et al*, 2003). Plasmids pABA-RNR (for overexpressing His_6-RNase R) and pTE607 (for overexpressing His_6-Hfq) were transformed into BL21(DE3) *E. coli* strain. Cells were grown at 37°C in LB medium supplemented with 100 µg/ml ampicillin to an OD600 of 0.5 and then protein overexpression was induced by addition of 1 mM IPTG (isopropyl β-D-thiogalactoside) for 3h. Cell cultures were pelleted by centrifugation at 8,500*g* for 15 min and stored at −80 C until use.

For RNase R purification, the cell pellet was resuspended in lysis buffer (50mM Tris at pH 7.5, 100 mM NaCl, 1mM EDTA, 5% glycerol and 1mM DTT) in the presence of Complete Mini Protease Inhibitor Cocktail EDTA-free (Roche) and disrupted by French Press. The crude extracts were treated with Benzonase (Sigma) to degrade the nucleic acids and clarified by a 30 min of centrifugation at 10,000*g*. The clarified extracts were then added to a 1 ml HiTrap Chelating Sepharose column equilibrated in buffer RNR-A (20 mM Tris at pH 8 and 500 mM NaCl) plus 20 mM imidazole and 2 mM 2-mercaptoethanol). Protein elution was achieved by a continuous imidazole gradient (from 20 to 500 mM) in buffer RNR-A. The fractions containing the purified protein were pooled and buffer-exchanged to buffer RNR-B (20 mM Tris at pH 8, 300 mM KCl and 2 mM 2-mercaptoethanol) using a 5 ml desalting column (GE Healthcare). Eluted proteins were concentrated by centrifugation at 7,000*g* for 15 min at 15°C with Amicon Ultra Centrifugal Filter Devices of 30 kDa molecular-mass cut-off (Millipore). RNase R concentration was determined by spectrophotometry measurement at A280 and 50% (v/v) glycerol was added to the final fractions before storage at −20°C.

For Hfq purification, the cell pellet was resuspended in lysis buffer (20mM Tris at pH 7.8, 500mM NaCl, 10% glycerol and 0.1% Triton X-100) in the presence of Complete Mini Protease Inhibitor Cocktail EDTA-free (Roche) and disrupted by French Press. The crude extracts were treated with Benzonase (Sigma) to degrade the nucleic acids and clarified by a 30 min of centrifugation at 10,000*g*. After a clarification step, imidazole–HCl (pH 7.8) was then added to the supernatant to a final concentration of 1 mM and the suspension was applied to a 1 ml HiTrap Chelating Sepharose column. The resin was then sequentially washed with 15 ml of Buffer Hfq-A (20 mM Tris at pH 7.8, 300 mM NaCl and 20 mM imidazole) and 15 mL of buffer Hfq-B (50 mM sodium phosphate at pH 6.0, 300

mM NaCl). Hfq was eluted with buffer Hfq-B with 250 mM imidazole. Fractions containing Hfq were determined by SDS–PAGE analysis, pooled and heated to 80°C for 15 min. Insoluble material was removed by centrifugation and the supernatant was buffer-exchanged to buffer Hfq-C (50 mM Tris–HCl at pH 7.5, 1 mM EDTA, 50 mM NH4Cl, 5% glycerol and 0.1% Triton X-100) using a 5 ml desalting column (GE Healthcare). Eluted Hfq was concentrated by centrifugation at 7,000*g* for 15 min at 15°C with Amicon Ultra Centrifugal Filter Devices of 3 kDa molecular-mass cut-off (Millipore) and its concentration was determined by spectrophotometric measurement at A280. The protein was kept at 4°C.

3.5.4. Pull down assay

Purified His_6-RNase R or His_6-Hfq were incubated with Ni-NTA beads (Qiagen) in 1 mL of binding buffer (50 mM Tris at pH 7.6, 100 mM NaCl, 10 mM imidazole) for 60 min at 4^{o}C. Stationary phase cultures resuspended in lysis buffer (50 mM Tris at pH 8, 125 mM NaCl, 10% glycerol, 0.1% TritonX-100, 1 mM PMSF) were disrupted by French Press and 1 mg of cell lysates was added to the beads. Incubation proceeded overnight at 4^{o}C with gentle rocking. The Ni-NTA resin was recovered by centrifugation and washed five times with binding buffer. Bound proteins were then eluted with elution buffer (50 mM Tris at pH 7.6, 100 mM NaCl, 300 mM imidazole). Eluted proteins were separated by SDS-PAGE and probed with Hfq or RNase R antibodies.

3.5.5. Co-immunoprecipitation RNase R and Hfq

Cell lysates were incubated with RNase R antibody bound to Protein A/G agarose beads, based on the instructions of the Pierce Crosslink Immunoprecipitation Kit (Thermo Scientific). Stationary phase cells were disrupted in ice cold IP Lysis (50 mM Tris at pH 8, 500 mM NaCl, 0.1% TritonX-100, 10% glycerol) in the presence of Complete Mini Protease Inhibitor Cocktail EDTA-free (Roche) using the French Press. Protein concentration in cell extracts was determined by Bradford assay. RNase R antibody was bound to Protein A/G Plus beads. 20µl of Protein A/G Plus resin was prepared and washed with Coupling Buffer (10 mM sodium phosphate at pH 7.2, 150 mM NaCl). RNase R antibody was then added

to the resin and incubated in a mixer for 2 h at 4°C. Subsequently, the resin was washed two times with Coupling Buffer. RNase R antibody was further crosslinked to Protein A/G Plus beads through the action of DSS (disuccinimidyl suberate) for 1h at room temperature. The resin was washed twice with Elution Buffer provided with the kit (pH 2.8) and washed two more times with Crosslink Washing Buffer (25 mM Tris at pH 7.4, 15 mM NaCl, 1 mM EDTA, 1% NP-40, 5% glycerol) to eliminate non-crosslinked antibody. 3 mg of cell extracts were added to the antibody-crosslinked resin and incubated overnight at 4°C with gentle mixing. After centrifugation of the flowthrough, the column was washed three times with TBS (0.02 M Tris, 0.137 M NaCl) and once with Conditioning Buffer provided with the IP kit. Antigens were eluted with sample loading buffer (50 mM Tris at pH 6.8, 2% SDS, 20 mM DTT, 10% glycerol, 0.04% BB) following a 10 min incubation at 100°C. Samples were collected by centrifugation and eluted Hfq in complex with RNase R was identified by Western blot analysis

3.5.6. Far-western blot analysis

Far-Western blotting was performed as described previously (Wu *et al*, 2007). RNase R and Hfq were purified as described above. Benzonase (Sigma-Aldrich) was used for elimination of contaminating nucleic acid during the protein extraction. When required, the purified proteins were further subjected to a Micrococcal Nuclease (Thermo Scientific) treatment for a more stringent elimination of contaminating nucleic acids. Increasing amounts of purified RNase R were loaded on a SDS-PAGE gel and transferred to a nitrocellulose membrane (Hybond ECL, GE Healthcare). BSA was used as a control. The membrane was stained with Ponceau S to verify the presence of immobilized proteins. After destaining, the proteins in the membrane were subject to a denaturation process in freshly prepared AC buffer (100 mM NaCl, 20 mM Tris pH 7.5, 0.5 mM EDTA, 10% glycerol, 0.1% Tween-20, 2% skim milk powder and 1 mM DTT) containing decreasing amounts of guanidine-HCl from 6 to 0.1 M. Renaturation of RNase R was done overnight at 4°C in AC buffer without guanidine-HCl. Membranes were blocked with 5% milk in TBST (20mM Tris at pH 7.6, 137 mM NaCl, 0.1% Tween-20) for 1 h at room temperature. Purified Hfq was added to AC buffer at a final concentration of 1 µg/ml and membranes

were incubated overnight at 4°C. Membranes were washed three times in TBST and incubated with primary (anti-Hfq) and secondary (horseradish-conjugated) antibodies in TBST with 5% milk. After incubation in Western Lightning Plus-ECL (Perkin-Elmer), the chemiluminescent signal was detected by the Chemidoc XRS$^+$ system (Bio-Rad).

3.5.7. RNA extraction, Northern blot and Primer extension

Total RNA was prepared from *E. coli* cells by the phenol:chloroform method. To examine 16S and 23S rRNAs, 1 µg of total RNA was resolved either in denaturing agarose or polyacrylamide gels following the protocols described in Andrade et al., 2013. For identification of the precursor rRNAs, 4% polyacrylamide/8M urea gels in TBE buffer were used as described (Leong *et al*, 2013). In Northern blots, all the lanes being compared in one panel correspond to the exact same gel but the lane order may differ from the original loading for presentation purposes. If required, digital contrast of the image was applied to the entire gel. DNA oligonucleotides complementary for 16S rRNA or 23S rRNA sequences were labeled at the 5' end by T4 polynucleotide kinase (Fermentas) and [γ-^{32}P]-ATP (Perkin-Elmer). Primers used for identification of pre-23S rRNA (Charollais *et al*, 2003) or pre-16S rRNA (Leong *et al*, 2013) were previously reported. The oligonucleotides used in this work are listed in Table S1 and were purchased from Stab Vida. Membranes were hybridized at 42^{o}C and analyzed with a PhosphorImager (Molecular Dynamics). The 5' end sequence of 23S rRNA was analyzed by primer extension using the 23S-PE primer (Table S1) labeled at the 5' end with [γ-^{32}P]-ATP. Primer extension analysis was carried as described (Andrade *et al*, 2012).

3.5.8. Ribosome extraction and Ribosome profile analysis

Ribosome isolation was adapted from Zundel et al., 2009. Cell pellets were resuspended in ice-cold buffer A (50 mM Tris-Cl at pH 7.5, 10 mM $MgCl_2$, 100 mM NH_4Cl, 0.5 mM EDTA, and 6 mM 2-mercaptoethanol) with the addition of Complete Mini Protease Inhibitor Cocktail EDTA-free (Roche) and lysed by four passes in a French Press. TurboDNase (Ambion) was added to the lysate. The cell lysate was centrifuged twice at 14,000 rpm for 10 min at 4^{o}C. Clarified lysate was layered over a 36% sucrose cushion

composed of buffer B (50 mM Tris-Cl at pH 7.5, 10 mM $MgCl_2$, 500 mM NH_4Cl, 0.5 mM EDTA, and 6 mM 2-mercaptoethanol) and spun at 44,000 rpm for 16 h in a Beckman ultracentrifuge 90Ti rotor at 4°C. The ribosome pellets were washed once with buffer C (50 mM Tris-Cl at pH 7.5, 10 mM $MgCl_2$, 100 mM NH_4Cl, and 6 mM 2-mercaptoethanol) and then resuspended in the same buffer by gentle rocking at 4°C.

Purified ribosomes were analyzed in 15%-50% (w/v) sucrose gradients in buffer C (50 mM Tris-Cl at pH 7.5, 10 mM $MgCl_2$, 100 mM NH_4Cl, and 6 mM 2-mercaptoethanol). This amount of $MgCl_2$ favors ribosomal subunits association. In contrast, to completely dissociate the 70S ribosomes into the individual 30S and 50S subunits, ribosomes were loaded onto 10%-30% (w/v) sucrose gradients in buffer C containing only 0.1 mM $MgCl_2$. In both conditions, samples were centrifuged in a Beckman ultracentrifuge SW28 rotor for 16 h at 24,000 rpm at 4°C. Fractions (1mL) were collected from the top and quantified by A260 measurement on a Nanodrop machine.

For identification of the rRNA species present in the ribosomal fractions, RNA was extracted by phenol:chloroform, precipitated in ethanol and 300 mM sodium acetate in the presence of glycogen as carrier and analyzed on agarose gels stained with ethidium bromide. Proteins in the ribosomal fractions were precipitated by addition of trichloroacetic acid to a final concentration of 10% and washed with cold acetone. Proteins were quantified by the Lowry assay. The localization of RNase R and Hfq was analyzed by Western blot using either RNase R or Hfq polyclonal antibodies.

3.5.9. Polysome profile analysis

Polysomes were prepared and resolved as described previously (Awano *et al*, 2010)with minor modifications. Briefly, *E. coli* cells were grown in LB at 37°C to exponential phase. At an OD600 of 0.5, polysomes were stabilized by addition of chloramphenicol to the culture to a final concentration of 0.1 mg/ml. After 5 minutes incubation with shaking, cells were harvested by centrifugation for 10 min at 5,000*g* at 4°C. The cell pellet was resuspended in cold buffer BP (20 mM Tris-Cl at pH 7.5, 10 mM $MgCl_2$, 100 mM NH_4Cl, and 5 mM 2-mercaptoethanol). Lysozyme was added to a final concentration of 1 mg/ml and the samples were incubated on ice for 2 min. The

suspension was frozen in liquid nitrogen and then thawed slowly in a water bath at 30ºC until no traces of ice remained. After three cycles of freezing and thawing, cell lysis was completed with addition of sodium deoxycholate to a final concentration of 0.05% and a 5 min incubation on ice. The lysate was spun at 14,000rpm for 10 min at 4ºC. Supernatants were loaded onto the top of 10%-40% (w/v) sucrose gradients prepared in buffer BP using the Gradient Master system (Biocomp) following the manufacturer's settings. Gradients were ultracentrifuged at 36,000 rpm for 3 h in a Beckman SW41 rotor at 4ºC. Polysome profiles were analyzed at 254 nm using the AKTA FPLC system (GE Healthcare) and the peak area was quantified using the Unicorn 5.11 software (GE Healthcare).

3.6. References

Andrade JM, Cairrão F & Arraiano CM (2006) RNase R affects gene expression in stationary phase: regulation of ompA. *Mol. Microbiol.* **60:** 219–28

Andrade JM, Hajnsdorf E, Régnier P & Arraiano CM (2009a) The poly(A)-dependent degradation pathway of *rpsO* mRNA is primarily mediated by RNase R. *RNA* **15:** 316–26

Andrade JM, Pobre V & Arraiano CM (2013) Small RNA modules confer different stabilities and interact differently with multiple targets. *PLoS One* **8:** e52866

Andrade JM, Pobre V, Matos AM & Arraiano CM (2012) The crucial role of PNPase in the degradation of small RNAs that are not associated with Hfq. *RNA* **18:** 844–55

Andrade JM, Pobre V, Silva IJ, Domingues S & Arraiano CM (2009b) The role of 3'-5' exoribonucleases in RNA degradation. *Prog. Mol. Biol. Transl. Sci.* **85:** 187–229

Andrade JM, dos Santos RF, Chelysheva I, Ignatova Z & Arraiano CM (2018) The RNA-binding protein Hfq is important for ribosome biogenesis and affects translation fidelity. *EMBO J.* **37:** e97631

Awano N, Rajagopal V, Arbing M, Patel S, Hunt J, Inouye M & Phadtare S (2010) *Escherichia coli* RNase R has dual activities, helicase and RNase. *J. Bacteriol.* **192:** 1344–52

Basturea GN, Zundel MA & Deutscher MP (2011) Degradation of ribosomal RNA during starvation: comparison to quality control during steady-state growth and a role for RNase PH. *RNA* **17:** 338–45

Bilusic I, Popitsch N, Rescheneder P, Schroeder R & Lybecker M (2014) Revisiting the coding potential of the *E. coli* genome through Hfq co-immunoprecipitation. *RNA Biol.* **11:** 641–54

Butland G, Peregrín-Alvarez JM, Li J, Yang W, Yang X, Canadien V, Starostine A, Richards D, Beattie B, Krogan N, Davey M, Parkinson J, Greenblatt J & Emili A (2005) Interaction network containing conserved and essential protein complexes in *Escherichia coli*. *Nature* **433:** 531–7

Cairrao F, Cruz A, Mori H, Arraiano CM, Cairrão F, Cruz A, Mori H, Arraiano CM, Cairrao F, Cruz A, Mori H & Arraiano CM (2003) Cold shock induction of RNase R and its role in the maturation of the quality control mediator SsrA/tmRNA. *Mol. Microbiol.* **50:** 1349–60

Charollais J, Pflieger D, Vinh J, Dreyfus M & Iost I (2003) The DEAD-box RNA helicase SrmB is involved in the assembly of 50S ribosomal subunits in *Escherichia coli*. *Mol. Microbiol.* **48:** 1253–65

Chen SS & Williamson JR (2013) Characterization of the ribosome biogenesis landscape in *E. coli* using quantitative mass spectrometry. *J. Mol. Biol.* **425:** 767–79

Cheng Z-F & Deutscher MP (2003) Quality control of ribosomal RNA mediated by Polynucleotide phosphorylase and RNase R. *Proc. Natl. Acad. Sci. U. S. A.* **100:** 6388–93

Cheng Z-F & Deutscher MP (2005) An important role for RNase R in mRNA decay. *Mol. Cell* **17:** 313–8

Chu L-Y, Hsieh T-J, Golzarroshan B, Chen Y-P, Agrawal S & Yuan HS (2017) Structural insights into RNA unwinding and degradation by RNase R. *Nucleic Acids Res.* **45:** 12015–12024

Connolly K, Rife JP & Culver G (2008) Mechanistic insight into the ribosome biogenesis functions of the ancient protein KsgA. *Mol. Microbiol.* **70:** 1062–75

Cukras AR, Southworth DR, Brunelle JL, Culver GM & Green R (2003) Ribosomal proteins S12 and S13 function as control elements for translocation of the mRNA:tRNA complex. *Mol. Cell* **12:** 321–8

Datsenko KA & Wanner BL (2000) One-step inactivation of chromosomal genes in

Escherichia coli K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* **97:** 6640–5

Deutscher MP (2009) Maturation and degradation of ribosomal RNA in bacteria. *Prog. Mol. Biol. Transl. Sci.* **85:** 369–91

Dinman JD (2009) The eukaryotic ribosome: current status and challenges. *J. Biol. Chem.* **284:** 11761–5

Domingues S, Moreira RN, Andrade JM, dos Santos RF, Bária C, Viegas SC & Arraiano CM (2015) The role of RNase R in trans-translation and ribosomal quality control. *Biochimie* **114:** 113–8

Folichon M, Arluison V, Pellegrini O, Huntzinger E, Régnier P & Hajnsdorf E (2003) The poly(A) binding protein Hfq protects RNA from RNase E and exoribonucleolytic degradation. *Nucleic Acids Res.* **31:** 7302–10

Hage A El & Tollervey D (2004) A surfeit of factors: why is ribosome assembly so much more complicated in eukaryotes than bacteria? *RNA Biol.* **1:** 10–5

Hajnsdorf E & Boni I V (2012) Multiple activities of RNA-binding proteins S1 and Hfq. *Biochimie* **94:** 1544–53

Hossain ST, Malhotra A & Deutscher MP (2016) How RNase R degrades structured RNA: ROLE OF THE HELICASE ACTIVITY AND THE S1 DOMAIN. *J. Biol. Chem.* **291:** 7877–87

Jacob AI, Köhrer C, Davies BW, RajBhandary UL & Walker GC (2013) Conserved bacterial RNase YbeY plays key roles in 70S ribosome quality control and 16S rRNA maturation. *Mol. Cell* **49:** 427–38

Kaczanowska M & Rydén-Aulin M (2007) Ribosome biogenesis and the translation process in *Escherichia coli*. *Microbiol. Mol. Biol. Rev.* **71:** 477–94

Leong V, Kent M, Jomaa A & Ortega J (2013) *Escherichia coli rimM* and *yjeQ* null strains accumulate immature 30S subunits of similar structure and protein complement. *RNA* **19:** 789–802

Liang W & Deutscher MP (2013) Ribosomes regulate the stability and action of RNase R. *J. Biol. Chem.* **288:** 34791–8

Liiv A & Remme J (2004) Importance of transient structures during post-transcriptional refolding of the pre-23S rRNA and ribosomal large subunit assembly. *J. Mol. Biol.* **342:** 725–41

Malecki M, Bárria C & Arraiano CM (2014) Characterization of the RNase R association with ribosomes. *BMC Microbiol.* **14:** 34

Matos RG, Barbas A & Arraiano CM (2009) RNase R mutants elucidate the catalysis of structured RNA: RNA-binding domains select the RNAs targeted for degradation. *Biochem. J.* **423:** 291–301

Richards J, Mehta P & Karzai AW (2006) RNase R degrades non-stop mRNAs selectively in an SmpB-tmRNA-dependent manner. *Mol. Microbiol.* **62:** 1700–12

Roy-Chaudhuri B, Kirthi N & Culver GM (2010) Appropriate maturation and folding of 16S rRNA during 30S subunit biogenesis are critical for translational fidelity. *Proc. Natl. Acad. Sci. U. S. A.* **107:** 4567–72

dos Santos RF, Quendera AP, Boavida S, Seixas AF, Arraiano CM & Andrade JM (2018) Major 3'–5' exoribonucleases in the metabolism of coding and non-coding RNA. In *Progress in Molecular Biology and Translational Science* pp 101–155. Academic Press

Shajani Z, Sykes MT & Williamson JR (2011) Assembly of bacterial ribosomes. *Annu. Rev. Biochem.* **80:** 501–26

Sharpe Elles LM, Sykes MT, Williamson JR & Uhlenbeck OC (2009) A dominant negative mutant of the *E. coli* RNA helicase DbpA blocks assembly of the 50S ribosomal subunit. *Nucleic Acids Res.* **37:** 6503–14

Sittka A, Lucchini S, Papenfort K, Sharma CM, Rolle K, Binnewies TT, Hinton JCD & Vogel J (2008) Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq. *PLoS Genet.* **4:** e1000163

Snel B, Lehmann G, Bork P & Huynen MA (2000) STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene. *Nucleic Acids Res.* **28:** 3442–4

Strader MB, Hervey WJ, Costantino N, Fujigaki S, Chen CY, Akal-Strader A, Ihunnah CA, Makusky AJ, Court DL, Markey SP & Kowalak JA (2013) A coordinated proteomic approach for identifying proteins that interact with the *E. coli* ribosomal protein S12. *J. Proteome Res.* **12:** 1289–99

Sulthana S, Basturea GN & Deutscher MP (2016) Elucidation of pathways of ribosomal RNA degradation: an essential role for RNase E. *RNA* **22:** 1163–71

Sulthana S & Deutscher MP (2013) Multiple exoribonucleases catalyze maturation of the 3' terminus of 16S ribosomal RNA (rRNA). *J. Biol. Chem.* **288:** 12574–9

Tsui HC, Leung HC & Winkler ME (1994) Characterization of broadly pleiotropic phenotypes caused by an hfq insertion mutation in *Escherichia coli* K-12. *Mol. Microbiol.* **13:** 35–49

Valentin-Hansen P, Eriksen M & Udesen C (2004) The bacterial Sm-like protein Hfq: a key player in RNA transactions. *Mol. Microbiol.* **51:** 1525–33

Vincent HA & Deutscher MP (2006) Substrate recognition and catalysis by the exoribonuclease RNase R. *J. Biol. Chem.* **281:** 29769–75

Vincent HA & Deutscher MP (2009) Insights into how RNase R degrades structured RNA: analysis of the nuclease domain. *J. Mol. Biol.* **387:** 570–83

Vogel J & Luisi BF (2011) Hfq and its constellation of RNA. *Nat. Rev. Microbiol.* **9:** 578–89

Weichenrieder O (2014) RNA binding by Hfq and ring-forming (L)Sm proteins: a trade-off between optimal sequence readout and RNA backbone conformation. *RNA Biol.* **11:** 537–49

Wilusz CJ & Wilusz J (2013) Lsm proteins and Hfq: Life at the 3' end. *RNA Biol.* **10:** 592–601

Wu Y, Li Q & Chen X-Z (2007) Detecting protein-protein interactions by Far western blotting. *Nat. Protoc.* **2:** 3278–3284

Yang Z, Guo Q, Goto S, Chen Y, Li N, Yan K, Zhang Y, Muto A, Deng H, Himeno H, Lei J & Gao N (2014) Structural insights into the assembly of the 30S ribosomal subunit in vivo: functional role of S5 and location of the 17S rRNA precursor sequence. *Protein Cell* **5:** 394–407

Yusupov MM, Yusupova GZ, Baucom A, Lieberman K, Earnest TN, Cate JH & Noller HF (2001) Crystal structure of the ribosome at 5.5 A resolution. *Science* **292:** 883–96

Zhang A, Wassarman KM, Rosenow C, Tjaden BC, Storz G & Gottesman S (2003) Global analysis of small RNA and mRNA targets of Hfq. *Mol. Microbiol.* **50:** 1111–24

Ziolkowska K, Derreumaux P, Folichon M, Pellegrini O, Régnier P, Boni I V & Hajnsdorf E (2006) Hfq variant with altered RNA binding functions. *Nucleic Acids Res.* **34:** 709–20

Zundel MA, Basturea GN & Deutscher MP (2009) Initiation of ribosome degradation during starvation in *Escherichia coli*. *RNA* **15:** 977–83

3.7. Supplemental information

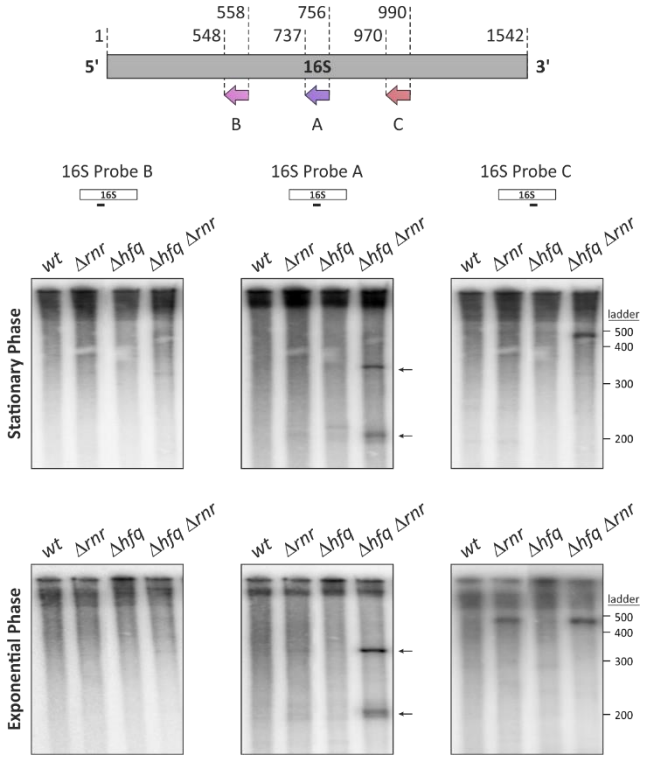


Figure S6 – Mapping of the 16S rRNA breakdown species. Total RNA resolved in 8% polyacrylamide gels and analyzed by Northern blotting using three different antisense probes of the mature 16S rRNA sequence: primer 16S-internal-A (complementary to nucleotides 737 to 756); primer 16S-internal-B (complementary to nucleotides 548 to 558); primer 16S-internal-C (complementary to nucleotides 970 to 990).

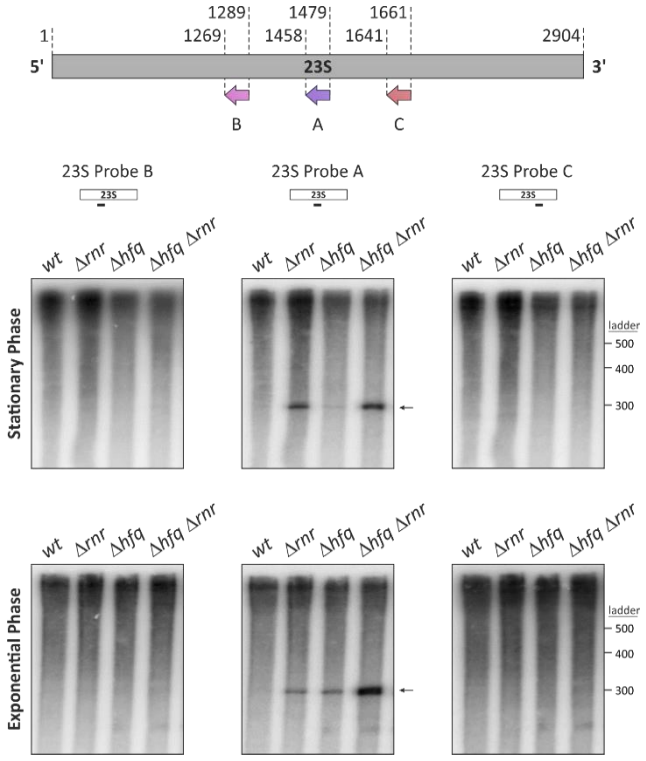


Figure S7 – Mapping of the 23S rRNA breakdown species. Total RNA resolved in 8% polyacrylamide gels and analyzed by Northern blotting using three different antisense probes of the mature 23S rRNA sequence: primer 23S-internal-A (complementary to nucleotides 1458 to 1479); primer 23S-internal-B (complementary to nucleotides 1269 to 1289); primer 23S-internal-C (complementary to nucleotides 1641 to 1661).

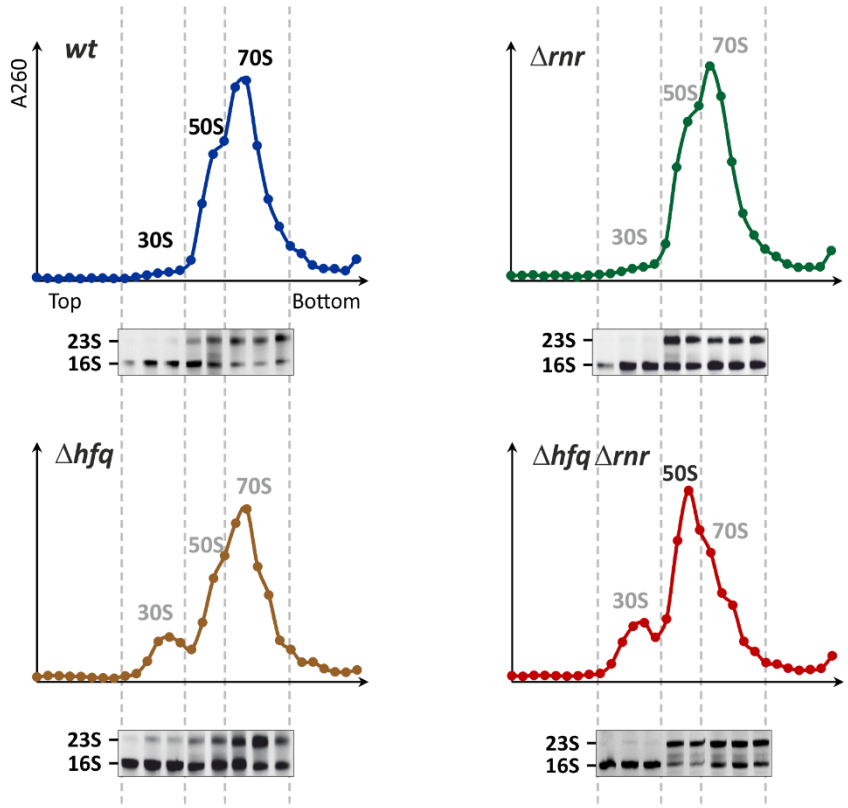


Figure S8 – Identification of rRNA isolated from ribosomal fractions. Ribosomes were analyzed by ultracentrifugation of 15%-50% (w/v) sucrose gradients under associative conditions (in the presence of 10 mM Mg^{2+}). Gradient fractions (1ml) were collected from the top of the centrifuge tubes and analyzed for A260 measurements on a Nanodrop spectrophotometer. RNA from each fraction was phenol-chloroform extracted and analyzed on 1.5% agarose gels stained with ethidium bromide. rRNA species confirm the identity of the ribosomal particle through the sucrose gradient.

Chapter 4

rRNA as source of sRNAs: Ribosomal RNA-derived fragments (rRF)

For this chapter I carried out the experiments validating the predicted rRNA-derived fragments and performed various computational analysis to assess possible functions. These functions are not yet experimentally validated and therefore rely on bioinformatic analysis for their discussion.

4. Chapter 4: Ribosomal RNA-derived RNA fragments (rRF)

4.1. Abstract

Ribosomal RNA is the central component of the ribosome, the macromolecular machine responsible for protein synthesis. rRNA provides the overall structure and is involved in all catalytic steps of translation reactions. It is transcribed as one long primary transcript that is subsequently processed into the mature rRNA molecules (16S, 23S and 5S) in bacteria. Particularly important is the processing of the 16S rRNA, serving as a late assembly checkpoint of ribosome biogenesis. Its precursor form (the 17S rRNA) contains 115 extra nucleotides at the 5' end and 33 nucleotides at the 3' that are trimmed out. Here we report that these extra nucleotide sequences at both ends of the 17S rRNA precursor are not promptly eliminated as previously though. We were able to detect two precursor rRNA-derived fragments (rRFs) of 115 and 33 nucleotides in length – rRF-115 and rRF-33, respectively. Independent bioinformatical predictions were carried out revealing that both rRFs have the potential to interact with transcripts involved in DNA replication, transcription, protein quality control and lipid biosynthesis. Additionally, interactions between these rRFs and other known sRNAs were computationally assessed, providing a subset of possible interacting partners involved in the sessile/motile lifestyle decision of the cell. Moreover, both rRFs were predicted to bind to multiple sites of the CsrB and CsrC sRNAs, which are the major repressors of the carbon storage regulator protein (CsrA). Altogether, our results show that rRF-115 and rRF-33 are present during exponential growth and that they may have a much more generalised role in the control of gene expression than previously envisioned. We propose that rRFs might act as a cellular signal that would link ribosome biogenesis to other vital cellular functions creating an intricate circuitry to supervise bacterial growth.

4.2. Introduction

Ribosomes are macromolecular machines that host the vital biological process of translation. These are large ribonucleoprotein (RNP) complexes essential for life of all organisms. Ribosomes are mainly constituted by ribosomal RNA (rRNA) plus a dedicated set of over 50 proteins (r-proteins). Specifically, the bacterial ribosome comprises two unidentical subunits – the 30S and 50S subunits. The first harbors the 16S rRNA molecule plus 21 r-proteins while the latter contains the 23S rRNA, the 5S rRNA and a set of 33 r-proteins. (Kaczanowska & Rydén-Aulin, 2007; Shajani *et al*, 2011). The RNA that constitutes these RNP complexes is not only responsible for the ribosome structural features but also for its function during all steps of the translation cycle. Ribosomes are therefore ribozymes that rely on the correct formation and maturation of its rRNAs (Nissen *et al*, 2000; Steitz & Moore, 2003). Hence, rRNA processing is one of the most important steps during ribosome biogenesis.

Over the years, ribosome biogenesis studies have lagged behind their structural characterization. As a result, the various steps of ribosome production still hold unidentified players and pathways. Regardless, rRNA processing has been fairly characterized in *Escherichia coli* and the many ribonucleases responsible for the maturation steps identified (Connolly & Culver, 2009; Shajani *et al*, 2011; Jacob *et al*, 2013; Sulthana & Deutscher, 2013). A polycistronic RNA of the entire rDNA operon is transcribed with all the three rRNAs embedded (and, depending on the locus, one or two tRNAs additionally). Processing of the primary transcript starts as soon as the sequences recognized by RNase III arise from the RNA polymerase complex. These endonucleolytic cleavages release the 17S, pre-23S and 9S rRNAs, which are the precursor forms of the 16S, 23S and 5S rRNAs, respectively. The precursors then undergo further processing by different sets of ribonucleases. The processing of the 16S rRNA from its precursor – 17S rRNA – is particularly important, since it is believed to serve as the final checkpoint for the release of functional matured 30S subunits. Accordingly, the 17S rRNA harbors 115 extra nucleotides at the 5’ end and 33 extra nucleotides at the 3’ end that need to be trimmed out. RNase E and RNase G participate in the 5’ end processing, the first cuts at the -49 position, leaving 66 nts that are subsequently processed by the latter. At the 3’ end, four

exoribonucleases – RNase R, RNase II, PNPase and RNase PH – redundantly participate in the elimination of the extra nucleotides, while YbeY is believed to endonucleolytically cut out the entire 33 nucleotide flanking sequence. Nevertheless, many non-ribonucleolytic biogenesis factors are known to affect the maturation of the 17S rRNA, although mechanistic insights are still lacking (Leong *et al*, 2013; Thurlow *et al*, 2016; Andrade *et al*, 2018). Moreover, there is no reported timeline of the processing events and additional RNases might also be involved. Consequently, there are certainly concurrent pathways for the correct processing of rRNAs, which would create a heterogeneity of byproducts. Drawing a parallel with rRNA processing, tRNAs also undergo RNase-mediated processing steps of their precursors to yield the matured functional forms (Lalaouna *et al*, 2015). Strikingly, a 3' end flanking sequence of the pre-tRNA^{leuZ} was shown to modulate the action of two sRNAs, RyhB and RybB. This pre-tRNA sequence can serve as a sponge, basepairing with the targeted sRNAs in order to avoid transcriptional noise (Lalaouna *et al*, 2015). The functional tRNA fragment is therefore termed tRNA-derived fragment (tRFs). With the recent discovery of tRFs, it is foreseeable that a similar process can occur during rRNA processing which would originate functional rRNA-derived fragments (rRFs). In fact, fragments of rRNA precursor sequences were found during a study that aimed to detect sRNAs with lengths between 30 to 65 nts in *E. coli* (Kawano *et al*, 2005). However, they were disregarded as products of rRNA processing and no biological functions were considered. Recently, rRFs were detected on both animal and plant eukaryotic cells, mainly through high-throughput approaches (Zhang *et al*, 2013; Asha & Soniya, 2017). Possible roles were predicted, although experimental validation and functional characterization is still lacking.

Here we identify two rRFs that that originate from precursor 17S rRNA flanking sequences – rRF-115 and rRF-33. These rRFs biogenesis is presumed to arise after rRNA processing occurs. Using bioinformatic approaches we predict possible mRNA targets of these fragments, which collectively seem to have a large impact in cell phsysiology, affecting different processes, ranging from replication, transcription and translation. Additionally, an interaction analysis with both rRFs ran independently against established

and predicted *E. coli* sRNAs revealed enriched interactions with sRNAs implicated in the sessile/motile decision.

4.3. Results

4.3.1. Two rRNA-derived fragments arise from 16S rRNA processing

The 16S rRNA is processed from a precursor 17S rRNA molecules bearing 115 and 33 extra nucleotides at the 5' and 3' end, respectively. Various maturation pathways are in place to ensure the processing of the 16S rRNA, however producing different rRNA byproducts. By analyzing the known enzymatic map of 16S rRNA maturation we reasoned that the entire 5' end sequence of the 17S rRNA could be released by the endonucleolytic action of RNase G (Li *et al*, 1999). Similarly, the recently discovered endoribonuclease YbeY could process out the entire 3' end flanking sequence (Davies *et al*, 2010; Jacob *et al*, 2013) (Figure 29A). To test this, total RNA from exponential phase cultures was extracted and analyzed through Northern blot using radiolabeled oligonucleotide probes that would hybridize specifically to the 17S rRNA flanking sequences. Both the 5' and the 3' end sequences were detected in range of their expected size (Figure 29B). We identified them as rRNA-derived fragments of 115 nts (rRF-115) and of 33 nts (rRF-33), accordingly. We could detect both rRFs by Northern blot using small amounts of total RNA which argues in favor for their high levels and stability. In fact, secondary structure prediction of both sequences revealed closed stem-loop structures with high basepairing probabilities (Figure 29C). These structures could certainly help stabilize the rRFs upon their release from the 17S precursor and protect them from degradation. Additionally, folding predictions were performed using multiple software packages, and all support the same secondary structure (Figure S9 and. Figure S10)

Several ribosome biogenesis factors are known to affect the processing of 16S rRNA. This is the case for the RNA chaperone Hfq and the 3'-5' exoribonuclease RNase R. Hfq (*hfq*) and RNase R (*rnr*) were shown to affect the maturation status of the 16S rRNA (Sulthana & Deutscher, 2013; Andrade *et al*, 2018). Strains lacking Hfq exhibit an accumulation of unprocessed 17S rRNA at both ends (Chapter 2), and this effect is steeper with the concomitant depletion of RNase R (Chapter 3). We have performed Northern

blot analysis on total RNAs extracted from single and double *hfq* and *rnr* deletion strains, using the same oligonucleotide probes for detecting the rRFs. Curiously, rRF-33 levels decrease in Δ*hfq* mutant and are barely detected in the Δ*rnr* Δ*hfq* mutant, which correlate with the 17S rRNA 3' end maturation status of these strains. This does not appear to be the case for rRF-115 (Figure S11).

Our results show that rRFs can be detected in exponential growing cells and that these are relatively stable RNAs, according to their predicted secondary structure. But do these rRFs comprise any functional potential?

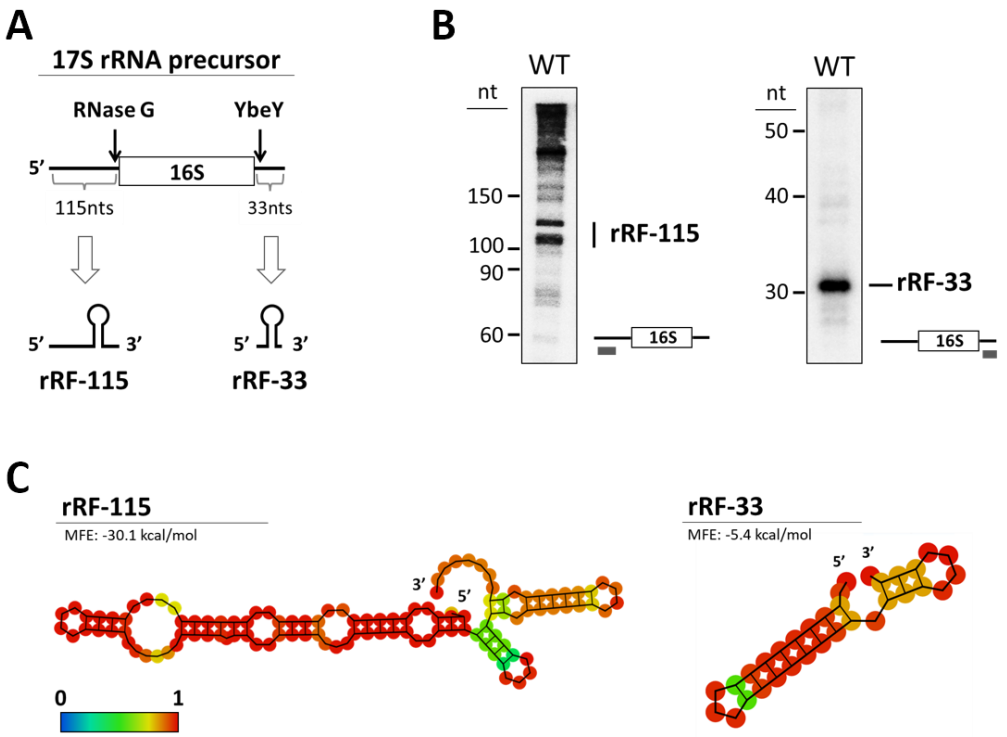


Figure 29 – rRF-115 and rRF-33 detection and secondary structure prediction. (A) Possible biogenesis pathway for rRF-115 and rRF-33 mediated by endoribonucleases. (B) Northern blot analysis of total RNA extracted from cells in exponential growth phase. Samples were fractionated on an 8% polyacrylamide/7M Urea gel. A scheme of the probes binding to the rRNA sequence is displayed on the side, as well as the relevant ladder bands of the Decade™ Marker (Ambion). (C) RNAfold webserver predicted secondary structures for rRF-115 and rRF-33. Minimum free energies (MFE) are provided on the top and coloring represents basepair probabilities, as indicated by the gradient below.

4.3.2. rRF-115 and rRF-33 mRNA target prediction

The size of the two rRFs, identified by Northern blotting (Figure 29), falls within range of functional small RNAs (Wang *et al*, 2016). Moreover, they seem to fold into stem-loop structures, also a feature of bacterial sRNAs. We hypothesized that these rRFs may be functional molecules, acting as sRNAs and therefore regulating mRNA translation and stability. To identify possible mRNA targets of these rRFs we used the sRNA target prediction tool CopraRNA (http://rna.informatik.uni-freiburg.de/CopraRNA/Input.jsp). Initially a blast sequence analysis was performed with both rRFs, revealing that these were highly conserved among closely related bacteria of the Enterobacteriaceae family. Accordingly, 5 to 7 species were selected for CopraRNA target analysis (see Supplementary information). Sequences comprising -200 to 100 nucleotides from the start codon of all identified ORFs were used and subject to hybridization analysis against rRF-115 and rRF-33 sequences. The top 25 hits are summarized for both rRFs (Table 2 and Table 3), but whole analysis is provided as supplementary material (Table S4 and Table S5).

Two ribosomal proteins L24 (*rplX*) and S15 (*rpsO*) appear to be plausible targets of rRF-115. Moreover, transcription modulation could be achieved through regulation of the RNA polymerase β' subunit (*rpoC*) and its sigma factor E (*rpoE*). In the case of rRF-33 important targets involved in DNA replication and repair through regulation of DNA polymerase III theta subunit (*holE*) and the DNA polymerase V protein UmuD (*umuD*) were identified. The levels of different alarmones through the regulation of ppGpp synthetase I (*relA*) and cyclic-di-GMP phosphodiesterase (*yhjH*) could also pose interesting regulatory implications.

Table 2 – List of rRF-115 targets predicted by the CopraRNA software. Targets are presented according to CopraRNA rank (first column) that balances interaction conservation among the species analyzed (see supplemental information)

Rank	Gene Name	Annotation	Position mRNA	Energy kcal/mol
1	***yrdA***	bacterial transferase hexapeptide domain protein	[87,95]	**-5.97**
2	***rnd***	ribonuclease D	[-182,-146]	**-12.67**
3	***ptsP***	PEP-protein phosphotransferase enzyme I; GAF domain containing protein	[88,99]	**-10.31**
4	***yrfF***	putative RcsCDB-response attenuator inner membrane protein	[-105,-88]	**-12.33**
5	***gstB***	glutathione S-transferase	[53,66]	**-13.19**
6	***bamA***	BamABCDE complex OM biogenesis outer membrane pore-forming assembly factor	[66,80]	**-8.02**
7	***nrdB***	ribonucleoside-diphosphate reductase 1 beta subunit ferritin-like protein	[39,48]	**-9.72**
8	***rpoE***	RNA polymerase sigma E factor	[-186,-158]	**-11.01**
9	***basS***	sensory histidine kinase in two-component regulatory system with BasR	[-194,-185]	**-7.88**
10	***yrdD***	ssDNA-binding protein function unknown	[-71,-46]	**-10.40**
11	***treR***	trehalose 6-phosphate-inducible trehalose regulon transcriptional repressor	[-36,-21]	**-10.54**
12	***rpoC***	RNA polymerase beta prime subunit	[44,56]	**-8.61**
13	***glnD***	uridylyltransferase	[82,93]	**-10.86**
14	***hslU***	molecular chaperone and ATPase component of HslUV protease	[-38,-20]	**-10.66**
15	***yieH***	phosphoenolpyruvate and 6-phosphogluconate phosphatase	[84,98]	**-7.97**
16	***ilvN***	acetolactate synthase 1 small subunit	[-73,-60]	**-10.45**
17	***rplX***	50S ribosomal subunit protein L24	[-71,-53]	**-10.33**
18	***mltB***	membrane-bound lytic murein transglycosylase B	[-139,-102]	**-12.04**
19	***obgE***	GTPase involved in cell partioning and DNA repair	[-17,18]	**-9.32**
20	***mutH***	methyl-directed mismatch repair protein	[-89,-76]	**-10.28**
21	***znuC***	zinc ABC transporter ATPase	[-63,-53]	**-8.77**
22	***rpsO***	30S ribosomal subunit protein S15	[-140,-130]	**-7.63**
23	***tmk***	thymidylate kinase	[75,84]	**-8.28**
24	***tehB***	tellurite selenium methyltransferase SAM-dependent (…)	[-8,18]	**-9.54**
25	***ftsI***	transpeptidase involved in septal peptidoglycan synthesis (…)	[-158,-136]	**-9.01**

Table 3 – List of rRF-33 targets predicted by the CopraRNA software. Targets are presented according to CopraRNA rank (first column) that balances interaction conservation among the species analyzed (see supplemental information)

Rank	Gene Name	Annotation	Position mRNA	Energy kcal/mol
1	***acpP***	acyl carrier protein (ACP)	[-1,21]	**-20.78**
2	***oxyR***	oxidative and nitrosative stress transcriptional regulator	[-94,-61]	**-14.59**
3	***ycjQ***	putative Zn-dependent NAD(P)-binding oxidoreductase	[-91,-63]	**-16.82**
4	***hemX***	putative uroporphyrinogen III methyltransferase	[-157,-149]	**-11.32**
5	***gmk***	guanylate kinase	[-89,-81]	**-11.38**
6	***yhjH***	cyclic-di-GMP phosphodiesterase FlhDC-regulated	[81,88]	**-13.02**
7	***holE***	DNA polymerase III theta subunit	[73,91]	**-12.48**
8	***umuD***	translesion error-prone DNA polymerase V subunit; RecA-activated auto-protease	[-132,-111]	**-14.45**
9	***msrB***	methionine sulfoxide reductase B	[-30,-22]	**-11.73**
10	***aceF***	pyruvate dehydrogenase dihydrolipoyltransacetylase component E2	[-157,-136]	**-11.48**
11	***mntR***	$Mn^{(2+)}$-responsive manganese regulon transcriptional regulator	[-52,-30]	**-12.69**
12	***grpE***	heat shock protein	[-86,-75]	**-9.35**
13	***yfgG***	uncharacterized protein	[-27,5]	**-13.57**
14	***relA***	(p)ppGpp synthetase I/GTP pyrophosphokinase	[93,100]	**-11.11**
15	***eutS***	putative ethanol utilization carboxysome structural protein	[6,15]	**-7.48**
16	***allD***	ureidoglycolate dehydrogenase	[68,75]	**-11.03**
17	***metH***	homocysteine-N5-methyltetrahydrofolate transmethylase B12-dependent	[-1,7]	**-11.45**
18	***nrdD***	anaerobic ribonucleoside-triphosphate reductase	[52,68]	**-11.60**
19	***adhE***	fused acetaldehyde-CoA dehydrogenase (…)	[-119,-110]	**-12.67**
20	***queE***	7-carboxy-7-deazaguanine synthase; queosine biosynthesis	[-17,0]	**-11.76**
21	***yqcC***	DUF446 family protein	[-190,-182]	**-12.52**
22	***yfcF***	glutathione S-transferase	[-199,-189]	**-6.78**
23	***iscX***	$Fe^{(2+)}$ donor and activity modulator for cysteine desulfurase	[-21,-14]	**-8.48**
24	***yceG***	septation protein ampicillin sensitivity	[-19,-11]	**-9.76**
25	***fucI***	L-fucose isomerase	[68,85]	**-11.61**

Additional bioinformatic prediction tools were used with similar parameters in order to refine the extensive list of possible targets. Namely, RNA Predator (http://rna.tbi.univie.ac.at/cgi-bin/RNApredator/target_search.cgi) and TargetRNA2 (http://cs.wellesley.edu/~btjaden/TargetRNA2/). Both rRF sequences were ran independently against the genome of *E. coli*. RNA Predator yielded a large set of targets for each rRFs but only the top 200 were considered, whereas the TargetRNA2 outputted smaller lists. Targets predicted by all three algorithms were summarized (Table 4). For rRF-115, although 28 targets were simultaneously detected by both CopraRNA and RNA Predator, only one target was suggested by all three tools, the HslU (*hslU*) protein. HslU is the ATPase component of the HslUV protease but also exhibits a protein chaperone function, being actively involved in post-translational quality control. In the case of rRF-33, 41 mRNA targets were common between CopraRNA and RNA Predator, yet TargetRNA2 predicted 5 interactions present in all three analysis. This restricted list of targets comprises: *acpP* – acyl carrier protein; *dnaN* – DNA polymerase III β subunit; *umuD* – DNA polymerase V associated protein; *relA* – ppGpp synthase I; *yhjH* (*pgeH*) - cyclic-di-GMP phosphodiesterase; and *yqcC* – a biofilm related protein.

The refined targets provide a starting point to understand the possible regulatory functions of rRF-115 and rRF-33. It is conceivable for this subset of biological processes that include protein quality control, replication, intracellular signaling and biofilm formation, maintain a regulatory circuit with ribosome biogenesis in order to optimize cellular resources and growth rate.

Table 4 – Refined list of targets for both rRF-115 and rRF-33 that were independently predicted by three different softwares (CopraRNA, RNA Predator and TargetRNA2). Interaction energies from CopraRNA are shown in kcal/mol.

rRF	Gene	Energy (kcal/mol)	Annotation	Process
rRF-115	*hslU*	-10.66	molecular chaperone and ATPase component of HslUV protease	**Protein quality control**
rRF-33	*acpP*	-20.78	acyl carrier protein (ACP)	**Lipid synthesis**
	dnaN	-9.59	DNA polymerase III beta subunit	**DNA replication & repair**
	umuD	-14.45	translesion error-prone DNA polymerase V subunit	
	relA	-11.11	ppGpp synthetase I	**Intracellular signaling**
	yhjH	-13.02	cyclic-di-GMP phosphodiesterase FlhDC-regulated	
	yqcC	-12.52	DUF446 family protein	**Biofilm**

4.3.3. rRFs as sRNA modulators

Besides interacting with mRNA transcripts, sRNAs have also been shown to bind and modulate the action of other sRNAs (Figueroa-Bossi *et al*, 2009; Miyakoshi *et al*, 2015). These were shown to arise from processing of not only mRNA transcripts but also of a tRNA molecule. The latter example is a tRNA-derived fragment (tRF) corresponding to the 3' external transcribed sequence of the $tRNA^{LeuZ}$ that was shown to function as a sponge for the RyhB and RybB sRNAs. Therefore, we reasoned that both rRF-115 and rRF-33 could also be targeting sRNAs either to degradation or to titer them. To test this, we used the IntaRNA software (http://rna.informatik.uni-freiburg.de/IntaRNA/Input.jsp) to predict RNA-RNA interactions between the rRFs and a complete list of known and predicted *E. coli* sRNAs.

The top 10 targets are shown in independent tables for rRF-115 (Table 5) and rRF-33 (Table 6). In the case of rRF-115 three sRNAs (isrB, IS128 and C0362) are classified as phantom genes (ORFs that were once thought to occur but are currently disregarded as such). nc8 and nc1 are predicted sRNAs identified in a pulldown assay of the nucleoid-associated protein HU, but only nc1 was experimentally validated. Notably, three of the predicted sRNAs interactions – DicF, rseX and CsrB – have described functions in the

regulation of motility *vs* biofilm lifestyles. For the rRF-33/sRNA interactions, C0299 is also a phantom gene and nc7 was identified in the same study as nc1 and nc8, with no additional characterization. Interestingly, three predicted sRNA interactors – CsrB, CsrC and SgrS – are also implicated in the motility/biofilm processes. This raises the question if both rRFs are cooperating in the regulation of motility and biofilm formation. Additionally, OmrB and CsrB were predicted to hybridize with both rRFs, which could suggest an intricate regulatory network.

Table 5 – List of rRF-115 sRNA targets predicted by the IntaRNA software. Targets are ordered by hybridization energy and the position of the nucleotides involved in the interaction is also provided.

sRNA	sRNA size (nt)	Position sRNA	Position rRF-115	Energy (kcal/mol)
IsrB	160	90-96	98-104	**-9.2**
DicF	53	41852	28-34	**-7.09**
SibE	150	43314	102-108	**-6.64**
IS128	209	96-102	98-104	**-6.56**
RseX	91	47-53	34-40	**-6.46**
nc8	256	96-102	96-102	**-5.79**
OmrB	76	45-51	74-80	**-5.65**
C0362	386	374-380	101-107	**-5.65**
CsrB	360	284-290	95-101	**-5.56**
nc1	168	129-135	92-98	**-5.49**

Table 6 – List of rRF-33 sRNA targets predicted by the IntaRNA software. Targets are ordered by hybridization energy and the position of the nucleotides involved in the interaction is also provided.

sRNA	sRNA size (nt)	Position sRNA	Position rRF-33	Energy (kcal/mol)
C0299	79	21-27	12-18	**-14.84**
CsrB	369	212-218	26-32	**-10.96**
OmrB	76	36-42	23-29	**-9.16**
SsrS/6S	183	81-87	23-29	**-7.85**
nc7	191	45-51	24-30	**-7.59**
CsrC	245	10-16	10-16	**-7.11**
SibC	151	31-37	24-30	**-6.54**
SgrS	227	175-181	23-29	**-5.59**
SsrA/tmRNA	363	316-322	11-17	**-5.01**
RyeA	249	17-23	8-14	**-4.85**

Remarkably, rRF-33 seems to interact with highly structured sRNAs with important functions. For instance, SsrS/6S RNA regulates transcription by directly binding the RNA polymerase complex (Wassarman & Storz, 2000), whilst SsrA/tmRNA is involved in the rescue of stalled ribosomes (Janssen & Hayes, 2012; Domingues *et al*, 2015). CsrB and CsrC are also structured sRNAs, both modulators of the carbon storage regulator A (CsrA) – a post-transcriptional pleiotropic regulator that typically binds and represses mRNA transcripts. Nevertheless, other regulatory mechanisms have been described where CsrA can also transcription or RNA decay (Yakhnin *et al*, 2013; Figueroa-Bossi *et al*, 2014). The Csr system is highly conserved being found in almost all bacteria phyla. CsrB/C sRNAs inhibit the activity of CsrA by directly binding and sequestering the protein (Figure 30) (Romeo, 1998; Babitzke & Romeo, 2007).

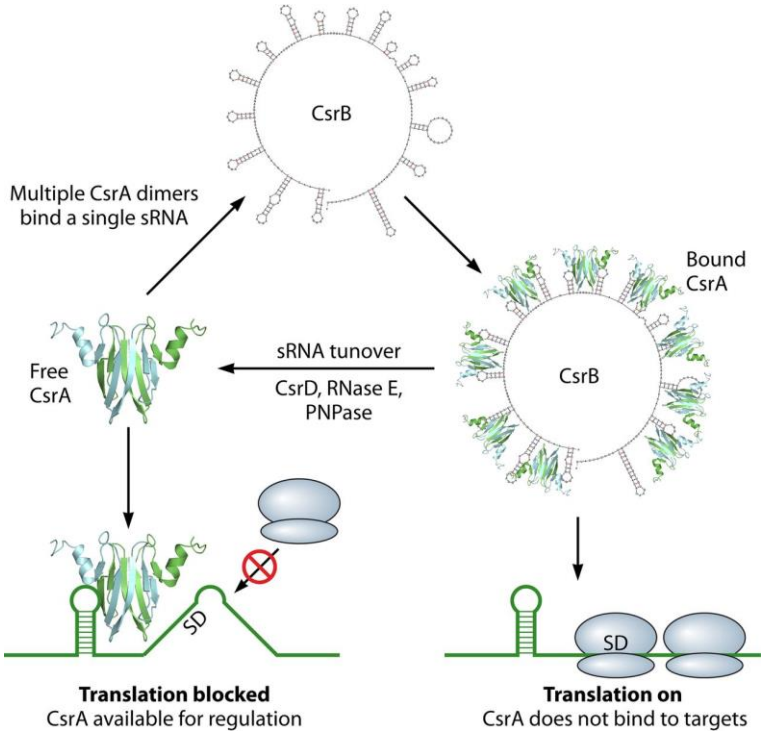
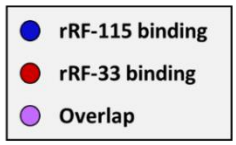


Figure 30 – CsrB action on the RNA-binding protein CsrA (adapted from Vakulskas *et al*, 2015). CsrB and CsrC both have the ability to sequester CsrA which often leads to a translational activation of the mRNA transcript targeted by this pleiotropic repressor.

Since both CsrB and CsrC have multiple sites of CsrA binding sites we further scanned both sRNAs for additional binding to the rRFs by relaxing the constraints of the IntaRNA input parameters. This allows the software to predic sub-optimal interactions in different sites of the queried CsrB and CsrC sequences. Strikingly, we found that both rRFs could collectively bind to 19 sites on CsrB (Figure 31) and 6 sites on CsrC (Figure 32), with variable stability energies. Four rRF-115 and rRF-33 interactions seem to overlap (violet colored nucleotides) namely, interactions number 4/5, 9/10 and 17/18 in CsrB plus the 4/5 in CsrC. The data seems to indicate that rRF-115 and rRF-33 cooperate to disrupt the characteristic stem-loops to which CsrA binds, hindering both sRNAs.



Interaction number	Nucleotides in hybridization	rRF	Hybrid energy (kcal/mol)
1	6-18	rRF-115	-10.0
2	36-43	rRF-115	-2.4
3	38-57	rRF-33	-14.4
4	73-89	rRF-33	-16.9
5	74-81	rRF-115	-7.0
6	91-104	rRF-115	-7.5
7	102-119	rRF-33	-17.2
8	115-128	rRF-115	-13.7
9	141-147	rRF-115	-4.6
10	142-148	rRF-33	-6.2
11	170-177	rRF-33	-6.7
12	181-190	rRF-115	-8.2
13	187-202	rRF-33	-15.6
14	221-229	rRF-33	-15.7
15	240-246	rRF-115	-5.7
16	266-274	rRF-115	-9.5
17	287-294	rRF-33	-7.7
18	287-299	rRF-115	-10.9
19	336-342	rRF-33	-10.7

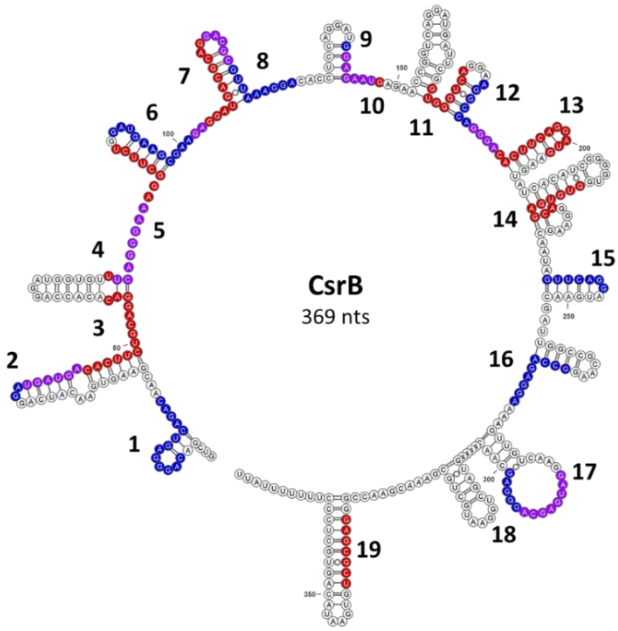
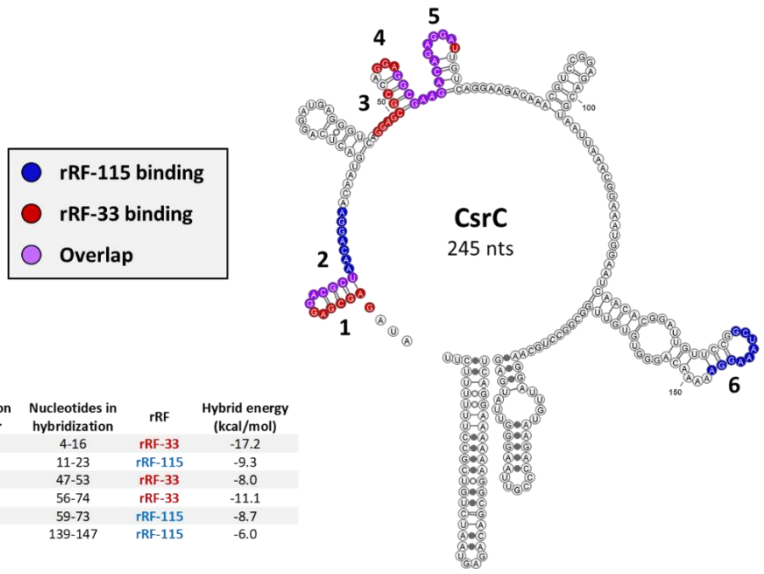


Figure 31 – Predicted multiple binding sites of CsrB targeted by rRF-115 and rRF-33. Secondary structure visualization was created using the VARNA applet. A table with all the suboptimal interactions between each rRF and the target CsrB sRNA is displayed in the bottom left with the corresponding interaction number (1-19) depicted on the RNA molecule. CsrB nucleotides in blue, represent rRF-115 binding; in red rRF-33 binding and in violet nucleotides that are predicted to interact with both rRFs.



Interaction number	Nucleotides in hybridization	rRF	Hybrid energy (kcal/mol)
1	4-16	rRF-33	-17.2
2	11-23	rRF-115	-9.3
3	47-53	rRF-33	-8.0
4	56-74	rRF-33	-11.1
5	59-73	rRF-115	-8.7
6	139-147	rRF-115	-6.0

Figure 32 – Predicted multiple binding sites of CsrC targeted by rRF-115 and rRF-33. Secondary structure visualization was created using the VARNA applet. A table with all the suboptimal interactions between each rRF and the target CsrC sRNA is displayed in the bottom left with the corresponding interaction number (1-6) depicted on the RNA molecule. CsrC nucleotides in blue, represent rRF-115 binding; in red rRF-33 binding and in violet nucleotides that are predicted to interact with both rRFs.

Collectively our results suggest that rRF-115 and rRF-33 could be involved in the modulation of motility and biofilm formation by regulating the sRNAs that impact these processes, possibly working as sponges to avoid transcriptional noise. Moreover, both rRFs could bind multiple sites of the CsrB and CsrC RNAs directly competing with CsrA or even promoting CsrB/C degradation. These observations nicely fit in a putative regulatory network that couples the highly energetic process of ribosome biogenesis to other crucial biological functions.

4.4. Discussion

RNA transcripts are important regulators whose roles span through virtually all cellular processes. In bacteria, a large number of these RNA regulators exist as relatively short transcripts that act on independently expressed targets. Regulatory small RNAs have been known since the late 1970's but only the advent of genome-wide methods like RNA-sequencing, allowed for their large-scale identification and led to an explosion of the number of known sRNA. Initial approaches for sRNA identification focused on intergenic regions (i.e. harboring no ORFs) where promoter and terminator elements control the sRNA expression (Vogel & Papenfort, 2006). However, alternative sRNA biogenesis pathways have been described. Small RNAs were found to derive from 5' UTRs and 3' UTRs of mRNAs (Kawano *et al*, 2005; Chao & Vogel, 2016; Miyakoshi *et al*, 2015) and from deep within the coding sequence of some transcripts (decRNAs) (Dar & Sorek, 2018). Additionally, sRNAs that arise from tRNA processing have also been described. These tRNA-derived fragments (tRFs) were shown to be functional RNAs in the regulation of gene expression (Kim *et al*, 2017).

Here we describe two novel small non-coding RNAs that derive from rRNA precursor sequences (rRF). rRF-115 corresponds to the entire 5' end flanking sequence of the 17S rRNA precursor, while rRF-33 corresponds to the 3' end flanking sequence. This indicates that parallel rRNA processing pathways are in place, much like expected due to the high importance of rRNA maturation. As a consequence, different byproducts arise from alternative maturation pathways, which could represent the genesis of the two rRFs detected in exponentially growing cells. Since both rRFs correspond to the entire flanking sequences of 17S rRNA we postulated that they arise from endoribonucleolytic cleavages. Two endoribonucleases are responsible for such processing, RNase G in the case of rRF-115 and YbeY in the case of rRF-33 (Shajani *et al*, 2011; Jacob *et al*, 2013; Li *et al*, 1999) (Figure 29A).

Our data suggests that both rRFs can interact with other small non-coding RNAs, which is reminiscent of the mode of action of the $tRNA^{LeuZ}$ tRF (Lalaouna *et al*, 2015). Notably, the list of promising regulators appears to be enriched in sRNAs that modulate bacterial motility and biofilm formation, namely, DicF, RseX, SgrS, OmrB, CsrB and CsrC

(Andreassen *et al*, 2018). This would imply rRF-115 and rRF-33 in the regulation of the sessile/motile decision, cooperating to repress motility during nutrient availability conditions. The latter two sRNAs, CsrB and CsrC, are critical in the regulation of CsrA, a widely conserved pleiotropic regulator involved in many biological processes. Glycogen synthesis and catabolism, gluconeogenesis, glycolysis, cell surface properties, adherence and motility are modulated by CsrA in *E. coli*. Moreover, numerous stationary phase genes are repressed by CsrA, whilst certain exponential-phase metabolic pathways are activated by this RNA-binding protein (Romeo, 1998; Vakulskas *et al*, 2015). Our results show both rRFs could bind to multiple sites of the CsrB and CsrC RNA sequence that would lead to the disruption of the secondary stem-loop structures responsible for sequestering CsrA. This would lead to an ineffective action of CsrB/C against CsrA, releasing the protein to repress its targets. In this way, rRF-115 and rRF-33 would act to prevent transcriptional noise from CsrB and CsrC during exponential growth, acting as either sRNA sponges or even guiding RNase-mediated degradation of CsrB and CsrC.

Besides acting as sRNA sponges rRF-115 and rRF-33 could also regulate mRNA transcripts. Our bioinformatic analysis generated a list of possible targets predicted to be regulated by these rRFs. Namely the HslU protein that could be regulated by rRF-115 (Table 4). HslU is the ATPase component of the HslUV proteolytic system, acting as protein chaperone that unfolds and translocates targeted proteins into the catalytic site of HslV (*hslV*) protease. In bacteria, proteases recognize their substrates by the presence of specific unstructured peptide signals. Different proteolytic complex have distinct affinities for binding a recognition tag (Sauer *et al*, 2004). During heat stress, the HslUV complex is more robust in degradation while other proteases are only slightly active (Kwon *et al*, 2004; Burton *et al*, 2005), which raises its harmful potential if not properly regulated. HslV is encoded immediately upstream of the HslU belonging to the same operon that yields a bicistronic transcript. The identified interaction between rRF-115 and HslU transcript (between 20-38 nucleotides upstream of the start codon) actually falls within the last nucleotides of HslV coding sequence. This raises the possibility of rRF-115 being regulating the expression of both components of the HslUV protease. The regulation of this heat-shock protease by rRF-115 would add another layer of control to the HslUV complex.

Other interesting rRF-115 targets are the L24 (*rplX*) and S15 (*rpsO*) ribosomal proteins (Table 2) whose regulation may impact a larger subset of 10 r-proteins present in the same operon of L24. In fact, the predicted interaction with the *rplX* transcript is 53 to 71 nucleotides upstream of translation start, which is located within the L14 (*rplN*) coding sequence, the first r-protein of the operon. This would allow for the close regulation of r-protein levels to maintain the correct stoichiometry for ribosome production, adding a feedback loop between ribosome biogenesis and r-protein synthesis. Moreover, rRF-115 was found to interact with components of the RNA polymerase machinery. Namely, the β' subunit (*rpoC*) that can also be co-transcribed with other r-proteins and the stress induced RNA polymerase sigma factor E (*rpoE*) that is also repressed by CsrA. Modulation of RNA polymerase components expression would serve as a feedback regulation to rRNA synthesis. Our knowledge of the molecular signals that sense the presence of functional ribosomes, and how it becomes transferred to the initiating RNA polymerase, is scarce (Kaczanowska & Rydén-Aulin, 2007). The implication of rRFs in this regulation would help explain how the cell maintains the correct amount of its rRNA components.

In the case of rRF-33 targetome, modulation of DNA polymerase activities through its theta (*holE*) and beta (*dnaN*) subunits (Table 3 and Table 4, respectively) nicely fits the observation that ribosome biogenesis is closely related to the bacterial growth rate (Nomura, 1999). Moreover, rRF-33 could also interact with UmuD (*umuD*), a component of the DNA polymerase V involved in the SOS response. Interestingly, the top target for rRF-33 in both IntaRNA and RNA Predator predictions is the acyl-carrier protein (*acpP*) which were also among the restricted list of TargetRNA2 predictions. The acyl-carrier protein (ACP) is a universal and highly conserved carrier of acyl intermediates during fatty acid synthesis and one of the most abundant proteins in *E. coli* cytosol. Fatty acid biosynthesis is required for the survival and proliferation of all organisms by generating the molecules that form the barrier that separates the intracellular space from the outside environment (Marcella & Barb, 2018). Consequently, we can envision how the cell would benefit from regulation networks that maintain a cross-talk between ACP levels and

translation capacity. rRF-33 would hence serve as a possible activator of the *acpP* transcript feedforwarding fatty acid biosynthesis after successful ribosome biogenesis.

With this work we were able to detect the rRFs through Northern blotting with minimal amounts of total RNA arguing that these RNAs linger in the cell avoiding degradation. Their predicted secondary structure seems to help in stabilization, since both sequences form stem-loop structures that could inhibit certain RNases. In fact, the folding energy of the rRF-115 (-0.26 kcal/mol/nt) falls within range of the ones observed for other small non-coding RNAs (Dar & Sorek, 2018). Although in the case of rRF-33 this value seems to be higher (-0.16 kcal/mol/nt). Moreover, this rRFs were shown to bind the RNA chaperone Hfq *in vitro* (Andrade *et al*, 2018) which can regulate their stability.

Collectively, our data points to a wide impact in cell physiology by rRF-115 and rRF-33. The observations that both rRFs are easily detected in small amounts of total RNA and that they seem to fold into structurally stable RNAs suggests that these are steadily maintain in fast growing cells. With a wide range of plausible targets to be assessed, these small rRNA-derived molecules would possibly serve as sensors of successful ribosome biogenesis, allowing a cross-talk between many indispensable cellular functions and protein synthesis capacity. Hence, this regulatory network would allow for the optimization of cellular resources. Although further elucidation is still required, the fitness advantage gained from functionalizing such highly abundant sequences seem evident. Similar rRF sequences arising from within the rRNA were previously detected in eukaryotic cells and have been suggested to be involved in various signaling pathways and biological processes. Albeit, this is the first functional analysis of sRNAs arising from the correct processing of the highly abundant 16S rRNA, unraveling new potential regulatory circuits able to sense successful ribosome biogenesis.

4.5. Materials and Methods

4.5.1. Bacterial strains and oligonucleotides

All *E. coli* K-12 strains used in this study are derivatives of the MG1693 strain, that was used as the wild-type. The single deletion mutants for *hfq* and *rnr* as well as the double Δ*hfq* Δ*rnr* were already part of the lab's collection. The oligonucleotides that hybridize with 17S rRNA precursor 5' end (TTAAGAATCCGTATCTTCGAGTGCCCACA) and 3' end (TGTGTGAGCACTGCAAAGTACGCTTCTTTAAGGTAAGG) sequences were used for the detection of rRF-115 and rRF-33, respectively.

4.5.2. Bacterial growth

Strains were grown in LB medium (Difco) supplemented with thymine (50 µg/ml) at 37ºC. Overnight cultures of single freshly grown colonies were diluted to an initial OD_{600} ~ 0.03. Cultures were collected at exponential phase (OD_{600} ~ 0.5). Antibiotics were present at the following concentrations when needed: 25 µg/ml chloramphenicol, 25 µg/ml kanamycin.

4.5.3. RNA analysis

For Northern blots, total RNA was extracted as previously described (Andrade *et al*, 2012). One microgram of total RNA was resolved on 8% polyacrylamide/7M urea gels in TBE 1x buffer, transferred to a nylon membrane (GE Healthcare) and UV crosslinked. Membranes were hybridized with PerfectHyb Plus (Sigma Aldrich) and probed with ^{32}P-5'end-labeled DNA oligonucleotides. Blots were analyzed on the Fuji TLA-5100 imaging system (GE Healthcare). The RNA ladder Decade™ Marker (Ambion) was used as standard.

4.5.4. Secondary structure prediction

rRF-115 and rRF-33 secondary structures were predicted using the RNAfold software using standard parameters (Hofacker, 2003). In order to strengthen the confidence of predicted structure each sequences was then further inputted into

additionally RNA secondary structure prediction algorithms, namely RNAstructure (Reuter & Mathews, 2010) and mFold (Zuker, 2003). The optimal structures predicted by each software were further subjected to alignment through the LocaRNA software (Will *et al*, 2012) to assess the nucleotide interaction conservation between the different software outputs. Treatment of the CsrB and CsrC secondary structures was performed using the VARNA applet (Darty *et al*, 2009).

4.5.5. Computational predictions of mRNA targets

The rRF-115 and rRF-33 *E. coli* sequences were independently analysed across multiple softwares and are given in the supplemental materials sections. In all predictions, default CopraRNA (Wright *et al*, 2014) UTR regions were used, comprising the -200/+100 nucleotides around each start codon. This parameter was adjusted when RNA Predator (Eggenhofer *et al*, 2011) and TargetRNA2 (Kery *et al*, 2014) softwares were used for refinement. 7 nucleotides were always required for seed region computations with no mismatches allowed. CopraRNA top 200 hits were summarized and are provided in the supplementary material section.

4.5.6. Computational predictions of rRF/sRNA interactions

The rRF-115 and rRF-33 *E. coli* sequences were analysed using the IntaRNA software (Wright *et al*, 2014). A list of known and predicted sRNAs in *E. coli* was extracted from the Bacterial Small Regulatory RNA Database (BSRD - http://www.bac-srna.org/BSRD/index.jsp) (Li *et al*, 2013) and submitted as the target RNA input of IntaRNA. Folding parameters were maintained with the default values, using the Turner model (Mathews *et al*, 2004). Number of (sub)optimal interactions was set to 1 to allow only one optimal hit per sRNA. The seed region of the predicted interactions was maintained with 7 nucleotides allowing no mismatches. These parameters were "relaxed" to allow one mismatch in the seed and multiple suboptimal interactions only when analysing multiple binding sites on the CsrB and CsrC small RNAs. Top hits are presented and ranked according to their interaction energy, unless stated otherwise.

4.6. References

Andrade JM, Pobre V, Matos AM & Arraiano CM (2012) The crucial role of PNPase in the degradation of small RNAs that are not associated with Hfq. *RNA* **18:** 844–55

Andrade JM, Dos Santos RF, Chelysheva I, Ignatova Z & Arraiano CM (2018) The RNA-binding protein Hfq is important for ribosome biogenesis and affects translation fidelity. *EMBO J.* **37:** e97631

Andreassen PR, Pettersen JS, Szczerba M, Valentin-Hansen P, Møller-Jensen J & Jørgensen MG (2018) sRNA-dependent control of curli biosynthesis in *Escherichia coli*: McaS directs endonucleolytic cleavage of *csgD* mRNA. *Nucleic Acids Res.* **46:** 6746–6760

Asha S & Soniya E V. (2017) The sRNAome mining revealed existence of unique signature small RNAs derived from 5.8SrRNA from *Piper nigrum* and other plant lineages. *Sci. Rep.* **7:** 41052

Babitzke P & Romeo T (2007) CsrB sRNA family: sequestration of RNA-binding regulatory proteins. *Curr. Opin. Microbiol.* **10:** 156–63

Burton RE, Baker TA & Sauer RT (2005) Nucleotide-dependent substrate recognition by the AAA+ HslUV protease. *Nat. Struct. Mol. Biol.* **12:** 245–251

Chao Y & Vogel J (2016) A 3' UTR-derived small RNA provides the regulatory noncoding arm of the inner membrane stress response. *Mol. Cell* **61:** 352–363

Connolly K & Culver G (2009) Deconstructing ribosome construction. *Trends Biochem. Sci.* **34:** 256–63

Dar D & Sorek R (2018) Bacterial noncoding RNAs excised from within protein-coding transcripts. *MBio* **9:** e01730-18

Darty K, Denise A & Ponty Y (2009) VARNA: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics* **25:** 1974–5

Davies BW, Köhrer C, Jacob AI, Simmons LA, Zhu J, Aleman LM, Rajbhandary UL & Walker GC (2010) Role of *Escherichia coli* YbeY, a highly conserved protein, in rRNA processing. *Mol. Microbiol.* **78:** 506–18

Domingues S, Moreira RN, Andrade JM, dos Santos RF, Bárria C, Viegas SC & Arraiano CM (2015) The role of RNase R in trans-translation and ribosomal quality control. *Biochimie* **114:** 113–8

Eggenhofer F, Tafer H, Stadler PF & Hofacker IL (2011) RNApredator: fast accessibility-based prediction of sRNA targets. *Nucleic Acids Res.* **39:** W149-54

Figueroa-Bossi N, Schwartz A, Guillemardet B, D'Heygère F, Bossi L & Boudvillain M (2014) RNA remodeling by bacterial global regulator CsrA promotes Rho-dependent transcription termination. *Genes Dev.* **28:** 1239–51

Figueroa-Bossi N, Valentini M, Malleret L, Fiorini F & Bossi L (2009) Caught at its own game: regulatory small RNA inactivated by an inducible transcript mimicking its target. *Genes Dev.* **23:** 2004–15

Hofacker IL (2003) Vienna RNA secondary structure server. *Nucleic Acids Res.* **31:** 3429–3431

Jacob AI, Köhrer C, Davies BW, RajBhandary UL & Walker GC (2013) Conserved bacterial RNase YbeY plays key roles in 70S ribosome quality control and 16S rRNA maturation. *Mol. Cell* **49:** 427–38

Janssen BD & Hayes CS (2012) The tmRNA ribosome rescue system. *Adv. Protein Chem. Struct. Biol.* **86:** 151–91

Kaczanowska M & Rydén-Aulin M (2007) Ribosome biogenesis and the translation process in *Escherichia coli*. *Microbiol. Mol. Biol. Rev.* **71:** 477–94

Kawano M, Reynolds AA, Miranda-Rios J & Storz G (2005) Detection of 5'- and 3'-UTR-derived small RNAs and cis-encoded antisense RNAs in *Escherichia coli*. *Nucleic Acids Res.* **33:** 1040–1050

Kery MB, Feldman M, Livny J & Tjaden B (2014) TargetRNA2: identifying targets of small regulatory RNAs in bacteria. *Nucleic Acids Res.* **42:** W124–W129

Kim HK, Fuchs G, Wang S, Wei W, Zhang Y, Park H, Roy-Chaudhuri B, Li P, Xu J, Chu K, Zhang F, Chua M-S, So S, Zhang QC, Sarnow P & Kay MA (2017) A transfer-RNA-derived small RNA regulates ribosome biogenesis. *Nature* **552:** 57–62

Kwon A-R, Trame CB & McKay DB (2004) Kinetics of protein substrate degradation by HslUV. *J. Struct. Biol.* **146:** 141–147

Lalaouna D, Carrier M-C, Semsey S, Brouard J-S, Wang J, Wade JTT & Massé E (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

Leong V, Kent M, Jomaa A & Ortega J (2013) *Escherichia coli rimM* and *yjeQ* null strains accumulate immature 30S subunits of similar structure and protein complement. *RNA* **19:** 789–802

Li L, Huang D, Cheung MK, Nong W, Huang Q & Kwan HS (2013) BSRD: a repository for bacterial small regulatory RNA. *Nucleic Acids Res.* **41:** D233–D238

Li Z, Pandit S & Deutscher MP (1999) RNase G (CafA protein) and RNase E are both required for the 5' maturation of 16S ribosomal RNA. *EMBO J.* **18:** 2878–85

Marcella AM & Barb AW (2018) Acyl-coenzyme A:(holo-acyl carrier protein) transacylase enzymes as templates for engineering. *Appl. Microbiol. Biotechnol.* **102:** 6333–6341

Mathews DH, Disney MD, Childs JL, Schroeder SJ, Zuker M & Turner DH (2004) Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc. Natl. Acad. Sci.* **101:** 7287–7292

Miyakoshi M, Chao Y & Vogel J (2015) Cross talk between ABC transporter mRNAs via a target mRNA-derived sponge of the GcvB small RNA. *EMBO J.* **34:** 1478–1492

Nissen P, Hansen J, Ban N, Moore PB & Steitz T a (2000) The structural basis of ribosome activity in peptide bond synthesis. *Science* **289:** 920–30

Nomura M (1999) Regulation of ribosome biosynthesis in *Escherichia coli* and *Saccharomyces cerevisiae*: diversity and common principles. *J. Bacteriol.* **181:** 6857–64

Reuter JS & Mathews DH (2010) RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics* **11:** 129

Romeo T (1998) Global regulation by the small RNA-binding protein CsrA and the non-coding RNA molecule CsrB. *Mol. Microbiol.* **29:** 1321–30

Sauer RT, Bolon DN, Burton BM, Burton RE, Flynn JM, Grant RA, Hersch GL, Joshi SA, Kenniston JA, Levchenko I, Neher SB, Oakes ESC, Siddiqui SM, Wah DA & Baker TA (2004) Sculpting the Proteome with AAA+ Proteases and Disassembly Machines. *Cell* **119:** 9–18

Shajani Z, Sykes MT & Williamson JR (2011) Assembly of bacterial ribosomes. *Annu. Rev. Biochem.* **80:** 501–26

Steitz TA & Moore PB (2003) RNA, the first macromolecular catalyst: the ribosome is a ribozyme. *Trends Biochem. Sci.* **28:** 411–8

Sulthana S & Deutscher MP (2013) Multiple exoribonucleases catalyze maturation of the 3' terminus of 16S ribosomal RNA (rRNA). *J. Biol. Chem.* **288:** 12574–9

Thurlow B, Davis JH, Leong V, F Moraes T, Williamson JR & Ortega J (2016) Binding properties of YjeQ (RsgA), RbfA, RimM and Era to assembly intermediates of the 30S subunit. *Nucleic Acids Res.* **44:** 9918–9932

Vakulskas CA, Potts AH, Babitzke P, Ahmer BMM & Romeo T (2015) Regulation of bacterial virulence by Csr (Rsm) systems. *Microbiol. Mol. Biol. Rev.* **79:** 193–224

Vogel J & Papenfort K (2006) Small non-coding RNAs and the bacterial outer membrane. *Curr. Opin. Microbiol.* **9:** 605–11

Wang Y, Li H, Sun Q & Yao Y (2016) Characterization of Small RNAs Derived from tRNAs, rRNAs and snoRNAs and Their Response to Heat Stress in Wheat Seedlings. *PLoS One* **11:** e0150933

Wassarman KM & Storz G (2000) 6S RNA regulates *E. coli* RNA polymerase activity. *Cell* **101:** 613–23

Will S, Joshi T, Hofacker IL, Stadler PF & Backofen R (2012) LocARNA-P: Accurate boundary prediction and improved detection of structural RNAs. *RNA* **18:** 900–914

Wright PR, Georg J, Mann M, Sorescu DA, Richter AS, Lott S, Kleinkauf R, Hess WR & Backofen R (2014) CopraRNA and IntaRNA: predicting small RNA targets, networks and interaction domains. *Nucleic Acids Res.* **42:** W119–W123

Yakhnin AV, Baker CS, Vakulskas CA, Yakhnin H, Berezin I, Romeo T & Babitze P (2013) CsrA activates *flhDC* expression by protecting flhDC mRNA from RNase E-mediated cleavage. *Mol Microbiol.* **87:** 851–866

Zhang K, Zheng S, Yang JS, Chen Y & Cheng Z (2013) Comprehensive profiling of protein lysine acetylation in *Escherichia coli*. *J. Proteome Res.* **12:** 844–851

Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31:** 3406–15

4.7. Supplementary information

Secondary structure prediction across multiple software packages

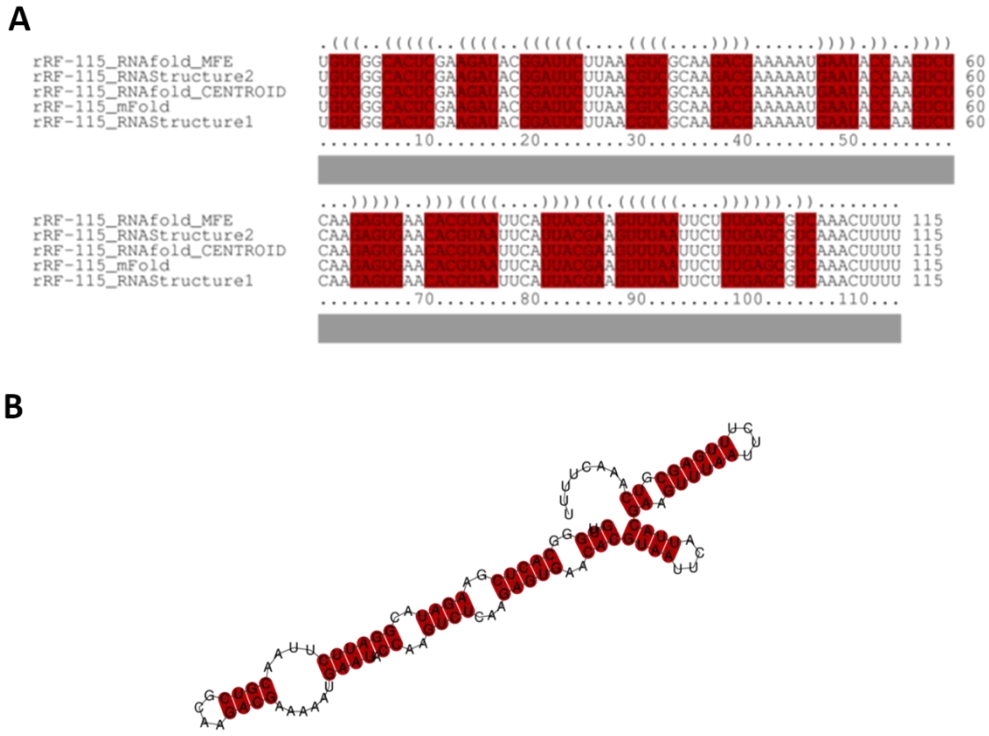


Figure S9 – rRF-115 secondary structure prediction results from multiple software packages (RNAfold, mFold and RNA Structure). (A) Alignment of rRF-115 sequence with the most conserved structured in dot-bracket format on top. (B) Conserved secondary structure of the analyzed sequence. Nucleotides in red show basepair conservation across multiple software analysis.

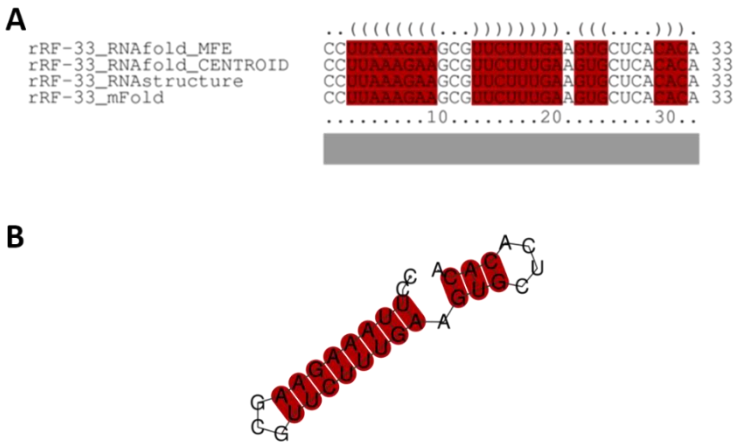


Figure S10 – rRF-33 secondary structure prediction results from multiple software packages (RNAfold, mFold and RNA Structure). (A) Alignment of rRF-33 sequence with the most conserved structured in dot-bracket format on top. (B) Conserved secondary structure of the analyzed sequence. Nucleotides in red show basepair conservation across multiple software analysis.

Northern blot detection of rRF-115 and rRF-33 in Δ*rnr*, Δ*hfq* and Δ*hfq* Δ*rnr* strains

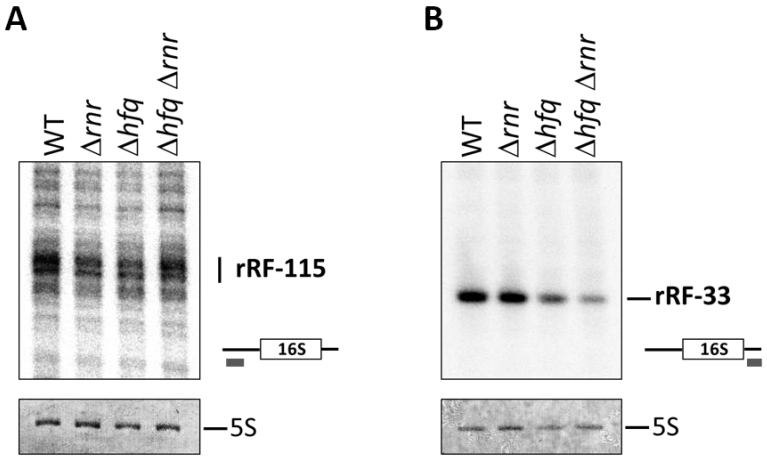


Figure S11 – Northern blot analysis of rRF-115 and rRF-33 in Δ*rnr*, Δ*hfq* and Δ*rnr* Δ*hfq* strains. (A) rRF-115 and (B) rRF-33 Northern blot analysis of total RNA extracted from cells in exponential growth phase. Samples were fractionated on an 8% polyacrylamide/7M Urea gel. A scheme of the probes binding to the rRNA sequence is displayed on the side. Methylene blue staining of 5S rRNA is provided in bottom of each panel.

FASTA input sequences for CopraRNA/IntaRNA analysis

rRF-115

>NC_000913_Escherichia_coli_str._K-12_substr._MG1655
TGTGGGCACTCGAAGATACGGATTCTTAACGTCGCAAGACGAAAAATGAATACCAAGTCTCAAGA
GTGAACACGTAATTCATTACGAAGTTTAATTCTTTGAGCGTCAAACTTTT
>NC_004337_Shigella_flexneri_2a_str._301
TGTGGGCACTCGAAGATACGGATTCTTAACGTCGCAAGACGAAAAATGAATACCAAGTCTCAAGA
GTGAACACGTAATTCATTACGAAGTTTAATTCTTTGAGCATCAAACTTTT
>NC_009344_Shigella_dysenteriae_Sd197
TGTGGGCACTCGAAGATACGGATTCTTAACGTCGCAAGACGAAAAATGAATACCAAGTCTCAAGA
GTGAACACGTAATTCATTACGAAGTTTAATTCTTTGAGCATCAAACTTTT
>NZ_CP024288_Escherichia_albertii
TGTGGGCACTCGAAGATACGGATTCTTAACGTCGCAAGACGAAAAATGAATACCAAGTCTCAAGA
GTGAACACGTAATTCATTACGAAGTTTAATTCTTTGAGCATCAAACTTTT
>NC_003197_Salmonella_enterica_subsp._enterica_serovar_Typhimuriu
m
TGTGGGCACTCGAAGATACGGATTCTTAACGTCGCAAGACGAAAAATGAATACCAAGTCTCAAGA
GTGAACACGTAATTCATTACGAAGTTTAATTCTTTGAGCATCAAACTTTT

rRF-33

>NC_000913 _Escherichia_coli_str._K-12_substr._MG1655
CCTTAAAGAAGCGTTCTTTGAAGTGCTCACACA
>NZ_CP025981 _Escherichia_marmotae
CCTTAAAGAAGCGTTCTTTGAAGTGCTCACACA
>NC_011743 _Escherichia_fergusonii_ATCC_35469
CCTTAAAGAAGCGTTCTTTGAAGTGCTCACACA
>NZ_CP019986 _Citrobacter_werkmanii
CCTTAAAGAAGCGTTCTTTGAAGTGCTCACACA
>NC_004337 _Shigella_flexneri_2a_str._301
CCTTAAAGAAGCGTTCTTTGTAGTGCTCACACA
>NC_009344 _Shigella_dysenteriae_Sd197
TTAAAGAAGCGTTCTTTGAAGTGCTCACACA
>NC_003197
_Salmonella_enterica_subsp._enterica_serovar_Typhimurium
CCTTAAAGAAGCGTACTTTGAAGTGCTCACACA

CopraRNA summarized result table for rRF-115 mRNA targets

Table S4 – Full set of rRF-115 targets predicted by CopraRNA

Rank	CopraRNA p-value	CopraRNA fdr	Locus Tag	Gene Name	Energy kcal/mol	IntaRNA p-value	Position mRNA	Annotation
1	0	0	b3279	*yrdA*	-5.97	0.202153	287 -- 295	bacterial transferase hexapeptide domain protein
2	9.139e-07	0.001553	b1804	*rnd*	-12.67	0.000589	18 -- 54	ribonuclease D
3	3.565e-05	0.04038	b2829	*ptsP*	-10.31	0.007441	288 -- 299	PEP-protein phosphotransferase enzyme I; GAF domain containing protein
4	6.219e-05	0.05283	b3398	*yrfF*	-12.33	0.000888	95 -- 112	putative RcsCDB-response attenuator inner membrane protein
5	9.292e-05	0.06315	b0838	*gstB*	-13.19	0.000300	253 -- 266	glutathione S-transferase
6	0.0001255	0.07108	b0177	*bamA*	-8.02	0.051022	266 -- 280	BamABCDE complex OM biogenesis outer membrane pore-forming assembly factor
7	0.0002394	0.107	b2235	*nrdB*	-9.72	0.012787	239 -- 248	ribonucleoside-diphosphate reductase 1 beta subunit ferritin-like protein
8	0.0002519	0.107	b2573	*rpoE*	-11.01	0.003753	14 -- 42	RNA polymerase sigma E factor
9	0.0002987	0.1078	b4112	*basS*	-7.88	0.056757	6 -- 15	sensory histidine kinase in two-component regulatory system with BasR
10	0.0003892	0.1078	b3283	*yrdD*	-10.40	0.006814	129 -- 154	ssDNA-binding protein function unknown
11	0.0004333	0.1078	b4241	*treR*	-10.54	0.005989	164 -- 179	trehalose 6-phosphate-inducible trehalose regulon transcriptional repressor
12	0.0004386	0.1078	b3988	*rpoC*	-8.61	0.032424	244 -- 256	RNA polymerase beta prime subunit
13	0.0004434	0.1078	b0167	*glnD*	-10.86	0.004359	282 -- 293	uridylyltransferase
14	0.0004443	0.1078	b3931	*hslU*	-10.66	0.005305	162 -- 180	molecular chaperone and ATPase component of HslUV protease
15	0.0005114	0.1113	b3715	*yieH*	-7.97	0.052825	284 -- 298	phosphoenolpyruvate and 6-phosphogluconate phosphatase
16	0.0005241	0.1113	b3670	*ilvN*	-10.45	0.006535	127 -- 140	acetolactate synthase 1 small subunit
17	0.0006078	0.1215	b3309	*rplX*	-10.33	0.007261	129 -- 147	50S ribosomal subunit protein L24
18	0.0006842	0.1281	b2701	*mltB*	-12.04	0.001245	61 -- 98	membrane-bound lytic murein transglycosylase B
19	0.0007233	0.1281	b3183	*obgE*	-9.32	0.018065	183 -- 218	GTPase involved in cell partioning and DNA repair

20	0.0007538	0.1281	b2831	*mutH*	-10.28	0.007633	111 -- 124	methyl-directed mismatch repair protein
21	0.0009393	0.1455	b1858	*znuC*	-8.77	0.028440	137 -- 147	zinc ABC transporter ATPase
22	0.0009418	0.1455	b3165	*rpsO*	-7.63	0.067945	60 -- 70	30S ribosomal subunit protein S15
23	0.001093	0.1603	b1098	*tmk*	-8.28	0.041813	275 -- 284	thymidylate kinase
24	0.001132	0.1603	b1430	*tehB*	-9.54	0.014865	192 -- 218	tellurite selenium methyltransferase SAM-dependent; tellurite selenium resistance protein
25	0.001246	0.1677	b0084	*ftsI*	-9.01	0.023408	42 -- 64	transpeptidase involved in septal peptidoglycan synthesis; penicillin-binding protein 3
26	0.001283	0.1677	b2184	*yejH*	-9.10	0.021781	255 -- 298	putative ATP-dependent DNA or RNA helicase
27	0.001386	0.172	b2374	*frc*	-10.08	0.009194	212 -- 231	formyl-CoA transferase NAD(P)-binding
28	0.001422	0.172	b2201	*ccmA*	-9.06	0.022354	211 -- 242	heme export ABC transporter ATPase
29	0.001468	0.172	b0893	*serS*	-8.04	0.050127	201 -- 230	seryl-tRNA synthetase
30	0.00177	0.1954	b3196	*yrbG*	-9.94	0.010418	1 -- 11	putative calcium/sodium:proton antiporter
31	0.00183	0.1954	b3770	*ilvE*	-8.99	0.023857	11 -- 22	branched-chain amino acid aminotransferase
32	0.001874	0.1954	b2126	*yehU*	-8.53	0.034480	135 -- 144	inner membrane putative sensory kinase in two-component system with YehT
33	0.001898	0.1954	b3159	*yhbV*	-12.10	0.001153	245 -- 259	U32 peptidase family protein
34	0.002028	0.2026	b3492	*yhiN*	-8.02	0.051224	205 -- 219	putative oxidoreductase
35	0.002161	0.2046	b0405	*queA*	-7.86	0.057629	232 -- 251	S-adenosylmethionine:tRNA ribosyltransferase-isomerase
36	0.002285	0.2046	b0493	*ybbO*	-9.04	0.022787	51 -- 65	putative oxidoreductase
37	0.002288	0.2046	b2994	*hybC*	-8.22	0.043843	196 -- 216	hydrogenase 2 large subunit
38	0.00233	0.2046	b2947	*gshB*	-8.29	0.041507	202 -- 226	glutathione synthetase
39	0.002362	0.2046	b1205	*ychH*	-8.87	0.026347	235 -- 264	DUF2583 family putative inner membrane protein
40	0.002408	0.2046	b3950	*frwB*	-8.01	0.051605	109 -- 147	putative enzyme IIB component of PTS
41	0.002777	0.2155	b1127	*pepT*	-10.66	0.005327	39 -- 70	peptidase T
42	0.002778	0.2155	b3452	*ugpA*	-9.07	0.022287	183 -- 203	sn-glycerol-3-phosphate ABC transporter permease
43	0.002784	0.2155	b2705	*srlD*	-8.06	0.049536	173 -- 195	sorbitol-6-phosphate dehydrogenase

44	0.002812	0.2155	b1112	*bhsA*	-9.29	0.018499	166 -- 175	biofilm cell surface and signaling protein
45	0.002854	0.2155	b0956	*matP*	-8.02	0.050904	3 -- 11	Ter macrodomain organizer matS-binding protein
46	0.003025	0.2157	b0026	*ileS*	-9.01	0.023432	108 -- 119	isoleucyl-tRNA synthetase
47	0.00303	0.2157	b2039	*rfbA*	-9.52	0.015160	108 -- 122	glucose-1-phosphate thymidylyltransferase
48	0.003048	0.2157	b0214	*rnhA*	-7.81	0.059842	200 -- 209	ribonuclease HI degrades RNA of DNA-RNA hybrids
49	0.003204	0.2165	b4034	*malE*	-7.20	0.092020	283 -- 298	maltose transporter subunit
50	0.003206	0.2165	b1712	*ihfA*	-6.47	0.148705	264 -- 275	integration host factor (IHF) DNA-binding protein alpha subunit
51	0.003372	0.2165	b1110	*ycfJ*	-9.76	0.012287	278 -- 298	uncharacterized protein
52	0.003412	0.2165	b0131	*panD*	-10.80	0.004647	16 -- 27	aspartate 1-decarboxylase
53	0.003415	0.2165	b4013	*metA*	-8.37	0.039123	244 -- 270	homoserine O-transsuccinylase
54	0.003469	0.2165	b2526	*hscA*	-7.45	0.077331	59 -- 79	DnaK-like molecular chaperone specific for IscU
55	0.003528	0.2165	b1628	*rsxB*	-7.10	0.098388	196 -- 206	SoxR iron-sulfur cluster reduction factor component; putative iron-sulfur protein
56	0.003568	0.2165	b0142	*folK*	-6.30	0.165312	108 -- 122	2-amino-4-hydroxy-6-hydroxymethyldihyropteridine pyrophosphokinase
57	0.003754	0.217	b3895	*fdhD*	-8.30	0.041119	248 -- 288	formate dehydrogenase formation protein
58	0.003757	0.217	b3519	*treF*	-7.63	0.068007	201 -- 214	cytoplasmic trehalase
59	0.003768	0.217	b0073	*leuB*	-8.21	0.044110	199 -- 214	3-isopropylmalate dehydrogenase NAD(+)-dependent
60	0.003866	0.219	b2747	*ispD*	-12.38	0.000833	65 -- 92	4-diphosphocytidyl-2C-methyl-D-erythritol synthase
61	0.004509	0.2512	b3746	*ravA*	-5.75	0.228722	96 -- 105	hexameric AAA+ MoxR family ATPase putative molecular chaperone
62	0.004584	0.2512	b3826	*yigL*	-7.62	0.068460	268 -- 278	pyridoxal phosphate phosphatase
63	0.004783	0.258	b1768	*pncA*	-7.72	0.063842	174 -- 203	nicotinamidase/pyrazinamidase
64	0.004909	0.2607	b0760	*modF*	-8.94	0.024821	233 -- 245	molybdate ABC transporter ATPase
65	0.005058	0.2644	b3421	*rtcB*	-8.74	0.029270	182 -- 209	RNA-splicing ligase
66	0.005213	0.2684	b0801	*ybiC*	-8.56	0.033796	219 -- 231	putative dehydrogenase
67	0.005353	0.2703	b2802	*fucI*	-8.68	0.030486	274 -- 284	L-fucose isomerase

68	0.00541	0.2703	b3185	*rpmA*	-7.86	0.057650	24 -- 41	50S ribosomal subunit protein L27
69	0.005814	0.2841	b2563	*acpS*	-5.63	0.245401	176 -- 190	holo-[acyl-carrier-protein] synthase 1
70	0.005937	0.2841	b0576	*pheP*	-7.41	0.079578	132 -- 143	phenylalanine transporter
71	0.005937	0.2841	b0036	*caiD*	-8.15	0.046314	57 -- 67	carnitinyl-CoA dehydratase
72	0.006221	0.2897	b1274	*topA*	-9.92	0.010632	163 -- 200	DNA topoisomerase I omega subunit
73	0.006224	0.2897	b2302	*yfcG*	-6.93	0.110119	168 -- 187	GSH-dependent disulfide bond oxidoreductase
74	0.006553	0.2958	b0591	*entS*	-7.81	0.059591	176 -- 206	enterobactin exporter iron-regulated
75	0.006586	0.2958	b4191	*ulaR*	-8.03	0.050532	237 -- 261	transcriptional repressor for the L-ascorbate utilization divergent operon
76	0.006856	0.2958	b3444	*insA*	-5.08	0.329153	40 -- 51	IS1 repressor TnpA
77	0.00691	0.2958	b1097	*yceG*	-8.99	0.023748	142 -- 159	septation protein ampicillin sensitivity
78	0.007009	0.2958	b1000	*cbpA*	-9.30	0.018274	273 -- 291	DnaK co-chaperone; curved DNA-binding protein
79	0.007153	0.2958	b0649	*djlC*	-8.92	0.025259	178 -- 188	J domain-containing HscC co-chaperone; Hsc56
80	0.007163	0.2958	b2660	*lhgO*	-9.03	0.022933	182 -- 208	L-2-hydroxyglutarate oxidase
81	0.007189	0.2958	b4465	*yggP*	-8.04	0.050356	275 -- 288	putative Zn-binding dehydrogenase
82	0.007219	0.2958	b1780	*yeaD*	-9.23	0.019428	18 -- 40	D-hexose-6-phosphate epimerase-like protein
83	0.007225	0.2958	b2666	*yqaE*	-9.57	0.014595	270 -- 300	cyaR sRNA-regulated protein
84	0.007439	0.3009	b0963	*mgsA*	-6.42	0.153786	6 -- 18	methylglyoxal synthase
85	0.007653	0.3039	b2154	*yeiG*	-7.62	0.068531	283 -- 291	S-formylglutathione hydrolase
86	0.007691	0.3039	b2510	*yfgJ*	-8.32	0.040604	26 -- 35	DUF1407 family protein
87	0.008174	0.318	b2553	*glnB*	-7.31	0.085185	126 -- 158	regulatory protein P-II for glutamine synthetase
88	0.008749	0.318	b1586	*ynfD*	-7.82	0.059117	164 -- 185	DUF1161 family periplasmic protein
89	0.008881	0.318	b2925	*fbaA*	-6.66	0.131832	273 -- 291	fructose-bisphosphate aldolase class II
90	0.00897	0.318	b2704	*srlB*	-8.27	0.042126	230 -- 248	glucitol/sorbitol-specific enzyme IIA component of PTS
91	0.009114	0.318	b0456	*ybaA*	-7.35	0.083007	287 -- 298	DUF1428 family protein
92	0.009135	0.318	b3082	*higA*	-11.80	0.001614	141 -- 160	antitoxinof the HigB-HigA toxin-antitoxin system

93	0.009168	0.318	b2015	*yeeY*	-6.96	0.108187	256 -- 266	LysR family putative transcriptional regulator
94	0.00919	0.318	b2964	*nupG*	-8.56	0.033652	20 -- 30	nucleoside transporter
95	0.00921	0.318	b1660	*ydhC*	-9.51	0.015254	46 -- 75	putative arabinose efflux transporter
96	0.009221	0.318	b1854	*pykA*	-7.13	0.096406	17 -- 29	pyruvate kinase II
97	0.009452	0.318	b3554	*yiaF*	-9.91	0.010723	239 -- 251	barrier effect co-colonization resistance factor; DUF3053 family lipoprotein
98	0.009543	0.318	b0091	*murC*	-8.64	0.031617	181 -- 202	UDP-N-acetylmuramate:L-alanine ligase
99	0.009555	0.318	b0815	*opgE*	-8.99	0.023778	210 -- 222	OPG biosynthetic transmembrane phosphoethanolamine transferase
100	0.009581	0.318	b1627	*rsxA*	-8.30	0.041334	104 -- 122	SoxR iron-sulfur cluster reduction factor component; inner membrane protein of electron transport complex
101	0.009594	0.318	b3293	*yhdN*	-9.02	0.023108	62 -- 72	DUF1992 family protein
102	0.009684	0.318	b1085	*yceQ*	-7.84	0.058482	184 -- 192	uncharacterized protein
103	0.01001	0.318	b2479	*gcvR*	-5.72	0.232495	198 -- 205	transcriptional repressor regulatory protein accessory to GcvA
104	0.01007	0.318	b4196	*ulaD*	-7.72	0.063787	182 -- 200	3-keto-L-gulonate 6-phosphate decarboxylase
105	0.01035	0.318	b4242	*mgtA*	-7.66	0.066726	201 -- 212	magnesium transporter
106	0.01049	0.318	b3190	*ibaG*	-7.43	0.078141	23 -- 46	acid stress protein; putative BolA family transcriptional regulator
107	0.01054	0.318	b2763	*cysI*	-7.11	0.097725	37 -- 46	sulfite reductase beta subunit NAD(P)-binding heme-binding
108	0.01064	0.318	b0183	*rnhB*	-8.82	0.027417	8 -- 37	ribonuclease HII degrades RNA of DNA-RNA hybrids
109	0.01089	0.318	b0152	*fhuD*	-8.18	0.045208	80 -- 93	iron(3+)-hydroxamate import ABC transporter periplasmic binding protein
110	0.0109	0.318	b4398	*creB*	-6.73	0.125978	231 -- 244	response regulator in two-component regulatory system with CreC
111	0.01097	0.318	b3903	*rhaA*	-7.86	0.057548	21 -- 29	L-rhamnose isomerase
112	0.01097	0.318	b3329	*gspH*	-9.62	0.013917	168 -- 206	putative general secretory pathway component cryptic
113	0.01098	0.318	b3712	*yieE*	-5.10	0.325479	88 -- 121	phosphopantetheinyl transferase superfamily protein
114	0.01112	0.318	b2668	*ygaP*	-7.52	0.073644	227 -- 243	DUF2892 family inner membrane rhodanese
115	0.01117	0.318	b3946	*fsaB*	-6.95	0.108924	103 -- 116	fructose-6-phosphate aldolase 2

116	0.01119	0.318	b0171	*pyrH*	-8.10	0.047990	85 -- 99	uridylate kinase
117	0.01131	0.318	b4366	*bglJ*	-6.51	0.145361	286 -- 296	bgl operon transcriptional activator
118	0.01132	0.318	b0224	*yafK*	-7.19	0.092887	221 -- 233	L D-transpeptidase-related protein
119	0.01149	0.318	b4320	*fimH*	-6.63	0.134343	185 -- 205	minor component of type 1 fimbriae
120	0.01157	0.318	b4176	*yjeT*	-5.68	0.238877	220 -- 227	DUF2065 family protein
121	0.0116	0.318	b3935	*priA*	-4.28	0.476957	55 -- 63	Primosome factor n' (replication factor Y)
122	0.01163	0.318	b1765	*ydjA*	-7.45	0.077264	74 -- 101	putative oxidoreductase
123	0.01181	0.318	b0550	*rusA*	-7.07	0.100584	116 -- 126	DLP12 prophage; endonuclease RUS
124	0.01183	0.318	b3160	*yhbW*	-7.77	0.061294	6 -- 14	putative luciferase-like monooxygenase
125	0.01183	0.318	b3319	*rplD*	-5.66	0.241493	24 -- 33	50S ribosomal subunit protein L4
126	0.01194	0.318	b2040	*rfbD*	-7.60	0.069395	32 -- 46	dTDP-L-rhamnose synthase NAD(P)-dependent dTDP-4-dehydrorhamnose reductase subunit
127	0.01204	0.318	b3932	*hslV*	-7.73	0.063158	278 -- 286	peptidase component of the HslUV protease
128	0.01221	0.318	b0678	*nagB*	-10.01	0.009778	145 -- 155	glucosamine-6-phosphate deaminase
129	0.01226	0.318	b4115	*adiC*	-6.83	0.118212	215 -- 223	arginine:agmatine antiporter
130	0.01228	0.318	b1654	*grxD*	-7.81	0.059838	115 -- 123	glutaredoxin-4
131	0.01233	0.318	b2215	*ompC*	-6.72	0.126533	63 -- 95	outer membrane porin protein C
132	0.0125	0.318	b3209	*elbB*	-7.47	0.076113	43 -- 53	isoprenoid biosynthesis protein with amidotransferase-like domain
133	0.01282	0.318	b4255	*rraB*	-7.69	0.064856	244 -- 263	protein inhibitor of RNase E
134	0.01286	0.318	b0169	*rpsB*	-6.76	0.123314	225 -- 234	30S ribosomal subunit protein S2
135	0.01289	0.318	b1734	*chbF*	-4.50	0.433879	255 -- 294	phospho-chitobiase; general 6-phospho-beta-glucosidase activity
136	0.01294	0.318	b0726	*sucA*	-7.16	0.094376	19 -- 32	2-oxoglutarate decarboxylase thiamine triphosphate-binding
137	0.01298	0.318	b3429	*glgA*	-6.89	0.113112	218 -- 228	glycogen synthase
138	0.01299	0.318	b0531	*sfmC*	-7.82	0.059179	222 -- 231	putative periplasmic pilus chaperone
139	0.01304	0.318	b1263	*trpD*	-6.83	0.117705	281 -- 293	fused glutamine amidotransferase (component II) of anthranilate synthase/anthranilate phosphoribosyl transferase

140	0.01313	0.318	b0102	*zapD*	-6.30	0.165714	176 -- 203	FtsZ stabilizer
141	0.0132	0.318	b3963	*fabR*	-7.33	0.084073	177 -- 203	transcriptional repressor of fabA and fabB
142	0.01341	0.321	b0407	*yajC*	-8.50	0.035312	91 -- 129	SecYEG protein translocase auxillary subunit
143	0.01351	0.321	b3321	*rpsJ*	-7.25	0.088632	212 -- 224	30S ribosomal subunit protein S10
144	0.01377	0.3235	b2557	*purL*	-6.68	0.129857	181 -- 208	phosphoribosylformyl-glycineamide synthetase
145	0.01381	0.3235	b3904	*rhaB*	-7.93	0.054565	173 -- 182	rhamnulokinase
146	0.0145	0.3303	b0034	*caiF*	-8.04	0.050239	4 -- 27	cai operon transcriptional activator
147	0.01462	0.3303	b2398	*yfeC*	-7.30	0.085796	186 -- 196	DUF1323 family putative DNA-binding protein
148	0.01463	0.3303	b0123	*cueO*	-9.29	0.018474	47 -- 59	multicopper oxidase (laccase)
149	0.01464	0.3303	b1704	*aroH*	-6.93	0.110384	196 -- 204	3-deoxy-D-arabino-heptulosonate-7-phosphate synthase tryptophan repressible
150	0.01471	0.3303	b2504	*yfgG*	-11.85	0.001535	103 -- 140	uncharacterized protein
151	0.01484	0.3303	b0619	*citA*	-3.97	0.538995	165 -- 173	sensory histidine kinase in two-component regulatory system with CitB
152	0.015	0.3303	b3066	*dnaG*	-7.76	0.061867	171 -- 185	DNA primase
153	0.01502	0.3303	b0532	*sfmD*	-10.32	0.007363	145 -- 163	putative outer membrane export usher protein
154	0.01504	0.3303	b3816	*corA*	-7.25	0.088570	142 -- 162	magnesium/nickel/cobalt transporter
155	0.01521	0.3303	b0884	*infA*	-6.71	0.127877	200 -- 212	translation initiation factor IF-1
156	0.01524	0.3303	b0850	*ybjC*	-7.90	0.055925	52 -- 75	DUF1418 family protein
157	0.01534	0.3303	b3337	*bfd*	-6.48	0.148273	202 -- 223	bacterioferritin-associated ferredoxin
158	0.01536	0.3303	b1291	*sapD*	-6.68	0.129964	226 -- 255	antimicrobial peptide ABC transporter ATPase
159	0.0159	0.3383	b2527	*hscB*	-7.30	0.085804	100 -- 111	HscA co-chaperone J domain-containing protein Hsc56; IscU-specific chaperone HscAB
160	0.01613	0.3383	b3981	*secE*	-7.21	0.091158	219 -- 228	preprotein translocase membrane subunit
161	0.01615	0.3383	b4225	*chpB*	-9.50	0.015423	11 -- 20	toxin of the ChpB-ChpS toxin-antitoxin system
162	0.01621	0.3383	b1709	*btuD*	-8.64	0.031536	146 -- 173	vitamin B12 ABC transporter ATPase

163	0.01623	0.3383	b1136	*icd*	-9.68	0.013145	182 -- 193	isocitrate dehydrogenase; e14 prophage attachment site; tellurite reductase
164	0.01661	0.3399	b1314	*ycjR*	-7.87	0.057171	265 -- 282	putative TIM alpha/beta barrel enzyme
165	0.01667	0.3399	b2038	*rfbC*	-6.79	0.121514	241 -- 265	dTDP-4-deoxyrhamnose-3 5-epimerase
166	0.01672	0.3399	b1501	*ydeP*	-7.94	0.054327	25 -- 40	putative oxidoreductase
167	0.0168	0.3399	b1051	*msyB*	-8.56	0.033607	216 -- 226	multicopy suppressor of secY and secA
168	0.01682	0.3399	b0953	*rmf*	-10.85	0.004417	268 -- 281	ribosome modulation factor
169	0.0169	0.3399	b0431	*cyoB*	-8.53	0.034544	279 -- 294	cytochrome o ubiquinol oxidase subunit I
170	0.01721	0.344	b1107	*nagZ*	-6.17	0.179320	84 -- 95	beta N-acetyl-glucosaminidase
171	0.01745	0.344	b2669	*stpA*	-6.45	0.150793	285 -- 293	DNA binding protein nucleoid-associated
172	0.01746	0.344	b0754	*aroG*	-8.73	0.029455	224 -- 236	3-deoxy-D-arabino-heptulosonate-7-phosphate synthase phenylalanine repressible
173	0.01751	0.344	b1563	*relE*	-8.76	0.028633	141 -- 165	Qin prophage; toxin of the RelE-RelB toxin-antitoxin system
174	0.01793	0.3502	b3512	*gadE*	-7.73	0.063157	38 -- 47	gad regulon transcriptional activator
175	0.0184	0.3573	b3690	*cbrA*	-7.35	0.082726	178 -- 188	colicin M resistance protein; FAD-binding protein putative oxidoreductase
176	0.01928	0.3703	b1771	*ydjG*	-7.12	0.097256	169 -- 195	methylglyoxal reductase NADH-dependent
177	0.01929	0.3703	b1643	*ydhI*	-7.42	0.078769	203 -- 217	DUF1656 family putative inner membrane efflux pump associated protein
178	0.01952	0.3726	b1015	*putP*	-7.20	0.091701	276 -- 297	proline:sodium symporter
179	0.01993	0.3784	b2379	*alaC*	-3.26	0.686030	158 -- 167	glutamate-pyruvate aminotransferase; glutamic-pyruvic transaminase (GPT); alanine transaminase
180	0.02023	0.3784	b0807	*rlmF*	-6.92	0.111464	184 -- 193	23S rRNA m(6)A1618 methyltransferase SAM-dependent
181	0.02033	0.3784	b2166	*psuK*	-11.43	0.002434	203 -- 226	pseudouridine kinase
182	0.02034	0.3784	b0882	*clpA*	-7.33	0.083968	30 -- 39	ATPase and specificity subunit of ClpA-ClpP ATP-dependent serine protease chaperone activity
183	0.02042	0.3784	b2923	*argO*	-6.59	0.137877	253 -- 284	arginine transporter
184	0.02064	0.3784	b3154	*yhbP*	-8.29	0.041485	143 -- 182	UPF0306 family protein

185	0.02065	0.3784	b0926	*ycbK*	-7.85	0.057888	29 -- 53	M15A protease-related family periplasmic protein
186	0.02076	0.3784	b3788	*rffG*	-11.18	0.003147	186 -- 207	dTDP-glucose 4 6-dehydratase
187	0.02082	0.3784	b4481	*wecF*	-6.56	0.140567	155 -- 176	TDP-Fuc4NAc:lipidIIFuc4NAc transferase
188	0.02104	0.3802	b3632	*waaQ*	-9.83	0.011557	246 -- 262	lipopolysaccharide core biosynthesis protein
189	0.02134	0.3818	b3054	*ygiF*	-6.26	0.169237	194 -- 227	inorganic triphosphatase
190	0.02135	0.3818	b1427	*rimL*	-9.60	0.014150	22 -- 42	ribosomal-protein-L7/L12-serine acetyltransferase
191	0.02156	0.3835	b3665	*adeD*	-7.75	0.062402	140 -- 151	cryptic adenine deaminase
192	0.02239	0.3946	b3674	*yidF*	-7.28	0.087348	220 -- 232	putative Cys-type oxidative YidJ-maturating enzyme
193	0.02242	0.3946	b3161	*mtr*	-6.98	0.106839	2 -- 10	tryptophan transporter of high affinity
194	0.0226	0.3959	b2895	*fldB*	-4.63	0.408882	7 -- 13	flavodoxin 2
195	0.02277	0.3967	b1913	*uvrC*	-6.04	0.193571	211 -- 225	excinuclease UvrABC endonuclease subunit
196	0.0229	0.397	b2585	*pssA*	-6.72	0.127221	30 -- 38	phosphatidylserine synthase; CDP-diacylglycerol-serine O-phosphatidyltransferase
197	0.02345	0.4028	b3765	*yifB*	-3.03	0.731426	187 -- 194	magnesium chelatase family protein and putative transcriptional regulator
198	0.02347	0.4028	b1569	*dicC*	-8.77	0.028417	212 -- 219	Qin prophage; DNA-binding transcriptional regulator for DicB
199	0.02379	0.4058	b0734	*cydB*	-5.85	0.216995	99 -- 110	cytochrome d terminal oxidase subunit II
200	0.02397	0.4058	b4472	*yhdP*	-7.15	0.095031	111 -- 128	DUF3971-AsmA2 domains protein

CopraRNA summarized result table for rRF-33 mRNA targets

Table S5 – Full set of rRF-33 targets predicted by CopraRNA

Rank	CopraRNA p-value	CopraRNA fdr	Gene Name	Energy kcal/mol	IntaRNA p-value	Position mRNA	Annotation
1	1	0	*acpP*	-20.78	0.000004	199 -- 221	acyl carrier protein (ACP)
2	5.021e-05	0.07481	*oxyR*	-14.59	0.001175	106 -- 139	oxidative and nitrosative stress transcriptional regulator
3	7.777e-05	0.07725	*ycjQ*	-16.82	0.000191	109 -- 137	putative Zn-dependent NAD(P)-binding oxidoreductase
4	0.0001504	0.09155	*hemX*	-11.32	0.011620	43 -- 51	putative uroporphyrinogen III methyltransferase
5	0.0001806	0.09155	*gmk*	-11.38	0.011165	111 -- 119	guanylate kinase
6	0.0001843	0.09155	*yhjH*	-13.02	0.003705	281 -- 288	cyclic-di-GMP phosphodiesterase FlhDC-regulated
7	0.000232	0.09876	*holE*	-12.48	0.005378	273 -- 291	DNA polymerase III theta subunit
8	0.0003628	0.1205	*umuD*	-14.45	0.001308	68 -- 89	translesion error-prone DNA polymerase V subunit; RecA-activated auto-protease
9	0.000364	0.1205	*msrB*	-11.73	0.008898	170 -- 178	methionine sulfoxide reductase B
10	0.0005276	0.1452	*aceF*	-11.48	0.010476	43 -- 64	pyruvate dehydrogenase dihydrolipoyltransacetylase component E2
11	0.0005358	0.1452	*mntR*	-12.69	0.004637	148 -- 170	$Mn^{(2+)}$-responsive manganese regulon transcriptional regulator
12	0.0006649	0.1538	*grpE*	-9.35	0.038647	114 -- 125	heat shock protein
13	0.0006709	0.1538	*yfgG*	-13.57	0.002506	173 -- 205	uncharacterized protein
14	0.0009309	0.1857	*relA*	-11.11	0.013270	293 -- 300	(p)ppGpp synthetase I/GTP pyrophosphokinase
15	0.0009349	0.1857	*eutS*	-7.48	0.108034	206 -- 215	putative ethanol utilization carboxysome structural protein
16	0.001066	0.1985	*allD*	-11.03	0.013946	268 -- 275	ureidoglycolate dehydrogenase
17	0.001197	0.1988	*metH*	-11.45	0.010674	199 -- 207	homocysteine-N5-methyltetrahydrofolate transmethylase B12-dependent
18	0.001346	0.1988	*nrdD*	-11.60	0.009669	252 -- 268	anaerobic ribonucleoside-triphosphate reductase
19	0.001349	0.1988	*adhE*	-12.67	0.004704	81 -- 90	fused acetaldehyde-CoA dehydrogenase/iron-dependent alcohol dehydrogenase/pyruvate-formate lyase deactivase
20	0.001354	0.1988	*queE*	-11.76	0.008721	183 -- 200	7-carboxy-7-deazaguanine synthase; queosine biosynthesis

21	0.001401	0.1988	*yqcC*	-12.52	0.005225	10 -- 18	DUF446 family protein
22	0.001607	0.2016	*yfcF*	-6.78	0.154433	1 -- 11	glutathione S-transferase
23	0.001716	0.2016	*iscX*	-8.48	0.063397	179 -- 186	$Fe(^{2+})$ donor and activity modulator for cysteine desulfurase
24	0.001723	0.2016	*yceG*	-9.76	0.030410	181 -- 189	septation protein ampicillin sensitivity
25	0.001765	0.2016	*fucI*	-11.61	0.009613	268 -- 285	L-fucose isomerase
26	0.001784	0.2016	*aceA*	-9.97	0.026854	54 -- 67	isocitrate lyase
27	0.00183	0.2016	*acrA*	-12.00	0.007424	14 -- 46	multidrug efflux system
28	0.00203	0.2016	*tabA*	-10.46	0.019857	263 -- 283	biofilm modulator regulated by toxins; DUF386 family protein cupin superfamily protein
29	0.00215	0.2016	*mscL*	-11.69	0.009137	202 -- 224	mechanosensitive channel protein high conductance
30	0.002201	0.2016	*glgB*	-9.71	0.031344	24 -- 45	1 4-alpha-glucan branching enzyme
31	0.002251	0.2016	*yjiJ*	-11.31	0.011672	215 -- 234	DUF1228 family putative inner membrane MFS superfamily transporter
32	0.002277	0.2016	*ybgE*	-10.17	0.023712	89 -- 123	putative inner membrane protein in cydABX-ybgE operon
33	0.002375	0.2016	*yhbY*	-10.59	0.018426	55 -- 66	RNA binding protein associated with pre-50S ribosomal subunits
34	0.002383	0.2016	*rraB*	-9.26	0.040654	1 -- 21	protein inhibitor of RNase E
35	0.002436	0.2016	*yihV*	-7.70	0.096317	157 -- 193	6-deoxy-6-sulphofructose kinase
36	0.002541	0.2016	*cydX*	-11.55	0.009980	202 -- 236	cytochrome d (bd-I) ubiquinol oxidase subunit X
37	0.002542	0.2016	*glk*	-6.98	0.139396	158 -- 164	glucokinase
38	0.002571	0.2016	*valS*	-11.00	0.014219	246 -- 275	valyl-tRNA synthetase
39	0.002876	0.2197	*rhmA*	-12.28	0.006172	216 -- 238	2-keto-3-deoxy-L-rhamnonate aldolase
40	0.003103	0.2312	*nuoJ*	-10.27	0.022307	135 -- 151	NADH:ubiquinone oxidoreductase membrane subunit J
41	0.00319	0.2318	*ybhQ*	-7.83	0.089889	26 -- 35	inner membrane protein
42	0.003371	0.2325	*dppB*	-8.73	0.055018	144 -- 163	dipeptide/heme ABC transporter permease
43	0.003373	0.2325	*wzc*	-9.76	0.030310	165 -- 185	colanic acid production tyrosine-protein kinase; autokinase; Ugd phosphorylase
44	0.00348	0.2325	*ampE*	-8.14	0.076262	256 -- 264	ampicillin resistance inner membrane protein; putative signaling protein in beta-lactamase regulation

45	0.00351	0.2325	*tmcA*	-11.88	0.008014	97 -- 115	elongator methionine tRNA (ac4C34) acetyltransferase
46	0.003658	0.2369	*mltA*	-6.50	0.176759	232 -- 238	membrane-bound lytic murein transglycosylase A
47	0.003868	0.2453	*fdoI*	-9.47	0.036121	181 -- 198	formate dehydrogenase-O cytochrome b556 subunit
48	0.004048	0.2498	*hemN*	-8.66	0.057258	195 -- 202	coproporphyrinogen III oxidase SAM and NAD(P)H dependent oxygen-independent
49	0.004108	0.2498	*yciN*	-10.31	0.021835	265 -- 283	DUF2498 protein YciN
50	0.004356	0.2596	*clpP*	-9.34	0.039004	280 -- 288	proteolytic subunit of ClpA-ClpP and ClpX-ClpP ATP-dependent serine proteases
51	0.004539	0.2628	*ihfA*	-9.88	0.028378	149 -- 171	integration host factor (IHF) DNA-binding protein alpha subunit
52	0.004782	0.2628	*gss*	-10.95	0.014639	198 -- 215	glutathionylspermidine amidase and glutathionylspermidine synthetase
53	0.00481	0.2628	*kdgR*	-9.45	0.036411	268 -- 299	KDG regulon transcriptional repressor
54	0.004835	0.2628	*cspA*	-7.85	0.089149	120 -- 139	RNA chaperone and antiterminator cold-inducible
55	0.005014	0.2628	*kdpB*	-11.55	0.010003	188 -- 196	potassium translocating ATPase subunit B
56	0.005099	0.2628	*nadC*	-8.00	0.082443	235 -- 254	quinolinate phosphoribosyltransferase
57	0.005145	0.2628	*dsdX*	-10.07	0.025231	221 -- 229	D-serine transporter
58	0.005396	0.2628	*lpp*	-8.61	0.058958	75 -- 87	murein lipoprotein
59	0.005474	0.2628	*proV*	-5.96	0.226893	281 -- 297	glycine betaine/proline ABC transporter periplasmic binding protein
60	0.005529	0.2628	*ldtD*	-12.47	0.005407	3 -- 31	murein L D-transpeptidase
61	0.005534	0.2628	*nanT*	-7.18	0.126220	169 -- 178	sialic acid transporter
62	0.005606	0.2628	*dmsA*	-7.60	0.101490	177 -- 183	dimethyl sulfoxide reductase anaerobic subunit A
63	0.005606	0.2628	*dmsA*	-7.60	0.101490	177 -- 183	dimethyl sulfoxide reductase anaerobic subunit A
64	0.005645	0.2628	*uvrD*	-9.44	0.036762	199 -- 213	DNA-dependent ATPase I and helicase II
65	0.005811	0.2664	*hdfR*	-7.85	0.089126	194 -- 200	flhDC operon transcriptional repressor
66	0.006008	0.2713	*yiiD*	-9.72	0.031177	119 -- 141	GNAT family putative N-acetyltransferase
67	0.006118	0.2721	*tisB*	-10.15	0.024003	73 -- 82	toxic membrane persister formation peptide LexA-regulated
68	0.006477	0.2839	*yfdC*	-11.65	0.009370	248 -- 266	putative inner membrane protein

69	0.006719	0.2872	*yqiA*	-9.52	0.035034	193 -- 217	acyl CoA esterase
70	0.006746	0.2872	*ptsG*	-7.94	0.085165	187 -- 214	fused glucose-specific PTS enzymes: IIB component/IIC component
71	0.00693	0.2909	*yihL*	-7.77	0.093045	42 -- 49	putative DNA-binding transcriptional regulator
72	0.007325	0.3032	*rhsD*	-9.41	0.037395	190 -- 207	Rhs protein with DUF4329 family putative toxin domain; putative neighboring cell growth inhibitor
73	0.007565	0.3088	*kup*	-8.41	0.065715	149 -- 163	potassium transporter
74	0.008075	0.3203	*adiC*	-10.03	0.025915	36 -- 52	arginine:agmatine antiporter
75	0.008384	0.3203	*bglF*	-9.95	0.027175	44 -- 65	fused beta-glucoside-specific PTS enzymes: IIA component/IIB component/IIC component
76	0.00841	0.3203	*dnaQ*	-8.21	0.073296	202 -- 209	DNA polymerase III epsilon subunit
77	0.008502	0.3203	*ybdF*	-9.93	0.027412	82 -- 103	DUF419 family protein
78	0.008607	0.3203	*entD*	-8.46	0.064177	276 -- 283	phosphopantetheinyltransferase component of enterobactin synthase multienzyme complex
79	0.008678	0.3203	*ybjX*	-9.53	0.034822	43 -- 65	DUF535 family protein
80	0.00868	0.3203	*eutP*	-8.28	0.070841	99 -- 106	putative P-loop NTPase ethanolamine utilization protein
81	0.008723	0.3203	*nlpI*	-8.92	0.049479	13 -- 31	lipoprotein involved in osmotic sensitivity and filamentation
82	0.008814	0.3203	*argH*	-8.14	0.076150	203 -- 224	argininosuccinate lyase
83	0.008939	0.3209	*yhjE*	-9.05	0.045957	49 -- 67	putative MFS transporter; membrane protein
84	0.009585	0.3369	*clcA*	-10.53	0.019091	105 -- 135	H(+)/Cl(-) exchange transporter
85	0.009609	0.3369	*yaiY*	-9.80	0.029643	51 -- 78	DUF2755 family inner membrane protein
86	0.01032	0.3458	*puuA*	-6.85	0.149255	294 -- 300	glutamate--putrescine ligase
87	0.011	0.3458	*yncD*	-9.42	0.037133	71 -- 83	putative iron outer membrane transporter
88	0.01105	0.3458	*fliS*	-6.72	0.158696	72 -- 84	flagellar protein potentiates polymerization
89	0.01118	0.3458	*yjgA*	-4.04	0.490008	45 -- 56	ribosome-associated UPF0307 family protein
90	0.01119	0.3458	*sgbE*	-3.14	0.643955	129 -- 136	L-ribulose-5-phosphate 4-epimerase
91	0.01125	0.3458	*rnhB*	-8.97	0.048250	48 -- 55	ribonuclease HII degrades RNA of DNA-RNA hybrids
92	0.01149	0.3458	*yafK*	-7.38	0.113714	290 -- 297	L D-transpeptidase-related protein

93	0.01155	0.3458	*znuA*	-8.84	0.051669	211 -- 225	zinc ABC transporter periplasmic binding protein
94	0.01163	0.3458	*rlmA*	-10.75	0.016673	55 -- 78	23S rRNA m(1)G745 methyltransferase SAM-dependent
95	0.01164	0.3458	*rnk*	-6.64	0.165050	87 -- 93	regulator of nucleoside diphosphate kinase
96	0.01164	0.3458	*ycdY*	-10.07	0.025247	88 -- 108	redox enzyme maturation protein (REMP) chaperone for YcdX
97	0.01192	0.3458	*yhhY*	-8.42	0.065439	30 -- 37	aminoacyl nucleotide detoxifying acetyltransferase
98	0.01195	0.3458	*dnaN*	-9.59	0.033556	220 -- 228	DNA polymerase III beta subunit
99	0.01201	0.3458	*rluE*	-7.06	0.134432	43 -- 49	23S rRNA pseudouridine(2457) synthase
100	0.01245	0.3458	*yddE*	-4.50	0.416466	59 -- 65	PhzC-PhzF family protein
101	0.01257	0.3458	*gpp*	-6.17	0.206481	158 -- 166	guanosine pentaphosphatase/exopolyphosphatase
102	0.01258	0.3458	*dppF*	-7.75	0.094181	202 -- 208	dipeptide/heme ABC transporter ATPas
103	0.01271	0.3458	*ompR*	-6.68	0.162094	142 -- 190	response regulator in two-component regulatory system with EnvZ
104	0.01279	0.3458	*yjcE*	-9.79	0.029850	111 -- 119	putative cation/proton antiporter
105	0.01283	0.3458	*nagZ*	-7.81	0.090882	199 -- 206	beta N-acetyl-glucosaminidase
106	0.01291	0.3458	*gtrB*	-7.47	0.108570	171 -- 193	CPS-53 (KpLE1) prophage; bactoprenol glucosyl transferase
107	0.01293	0.3458	*iscR*	-9.04	0.046143	202 -- 225	isc operon transcriptional repressor; suf operon transcriptional activator; oxidative stress- and iron starvation-inducible; autorepressor
108	0.01308	0.3458	*yedW*	-7.90	0.086684	45 -- 63	response regulator family protein
109	0.01314	0.3458	*wrbA*	-9.39	0.037821	109 -- 136	NAD(P)H:quinone oxidoreductase
110	0.01331	0.3458	*carB*	-8.89	0.050382	269 -- 276	carbamoyl-phosphate synthase large subunit
111	0.01334	0.3458	*slyD*	-10.18	0.023551	134 -- 141	FKBP-type peptidyl prolyl cis-trans isomerase (rotamase)
112	0.01335	0.3458	*mutM*	-7.85	0.089295	1 -- 13	formamidopyrimidine/5-formyluracil/ 5-hydroxymethyluracil DNA glycosylase
113	0.01337	0.3458	*amiA*	-9.12	0.044169	199 -- 209	N-acetylmuramoyl-l-alanine amidase I
114	0.01354	0.3458	*ybiB*	-4.57	0.405869	174 -- 198	putative family 3 glycosyltransferase
115	0.01356	0.3458	*zraS*	-5.66	0.259317	179 -- 196	sensory histidine kinase in two-component regulatory system with ZraR
116	0.01366	0.3458	*ydcY*	-3.37	0.604367	248 -- 256	DUF2526 family protein

117	0.01368	0.3458	*yfiH*	-8.68	0.056741	154 -- 178	UPF0124 family protein
118	0.01379	0.3458	*tig*	-8.44	0.064913	173 -- 193	peptidyl-prolyl cis/trans isomerase (trigger factor)
119	0.01381	0.3458	*frvR*	-4.91	0.355643	10 -- 16	putative frv operon regulator; contains a PTS EIIA domain
120	0.01399	0.3473	*metJ*	-6.40	0.185585	239 -- 245	transcriptional repressor S-adenosylmethionine-binding
121	0.01431	0.3497	*helD*	-8.84	0.051847	109 -- 136	DNA helicase IV
122	0.0144	0.3497	*rpiA*	-6.57	0.170527	236 -- 248	ribose 5-phosphate isomerase constitutive
123	0.01473	0.3497	*yfgD*	-7.40	0.112682	149 -- 159	putative oxidoreductase
124	0.01485	0.3497	*gntU*	-8.70	0.056112	35 -- 58	gluconate transporter low affinity GNT 1 system
125	0.01493	0.3497	*fhuE*	-7.37	0.114715	191 -- 207	ferric-rhodotorulic acid outer membrane transporter
126	0.01497	0.3497	*rpsN*	-8.41	0.065814	44 -- 50	30S ribosomal subunit protein S14
127	0.01501	0.3497	*yciC*	-8.55	0.061107	91 -- 112	UPF0259 family inner membrane protein
128	0.01522	0.3497	*frr*	-7.23	0.122975	52 -- 58	ribosome recycling factor
129	0.01529	0.3497	*potH*	-6.76	0.155410	202 -- 208	putrescine ABC transporter permease
130	0.01534	0.3497	*uvrA*	-5.32	0.301250	8 -- 20	ATPase and DNA damage recognition protein of nucleotide excision repair excinuclease UvrABC
131	0.01565	0.3497	*malT*	-8.72	0.055473	256 -- 263	mal regulon transcriptional activator
132	0.01566	0.3497	*yiaJ*	-8.96	0.048404	213 -- 234	transcriptional repressor for the yiaKLMNO-lyxK-sgbHUE operon
133	0.01568	0.3497	*codA*	-7.09	0.132276	1 -- 8	cytosine/isoguanine deaminase
134	0.01615	0.3497	*srlQ*	-5.60	0.266954	178 -- 199	D-arabinose 5-phosphate isomerase
135	0.01629	0.3497	*yaiV*	-7.90	0.086934	187 -- 196	putative transcriptional regulator
136	0.01642	0.3497	*yicR*	-8.93	0.049309	176 -- 196	UPF0758 family protein
137	0.01648	0.3497	*rpoZ*	-2.87	0.689839	184 -- 199	RNA polymerase omega subunit
138	0.01661	0.3497	*nadD*	-8.99	0.047509	125 -- 133	nicotinic acid mononucleotide adenylyltransferase NAD(P)-dependent
139	0.01664	0.3497	*hcaB*	-8.39	0.066626	267 -- 290	2 3-dihydroxy-2 3-dihydrophenylpropionate dehydrogenase
140	0.01665	0.3497	*ulaG*	-9.96	0.027054	241 -- 248	L-ascorbate 6-phosphate lactonase
141	0.01668	0.3497	*fimC*	-7.17	0.127054	201 -- 207	periplasmic chaperone

142	0.01697	0.3497	*csdA*	-8.87	0.051010	194 -- 212	cysteine sulfinate desulfinase
143	0.01704	0.3497	*groL*	-8.33	0.068980	163 -- 190	Cpn60 chaperonin GroEL large subunit of GroESL
144	0.01707	0.3497	*pykA*	-6.14	0.209216	192 -- 216	pyruvate kinase II
145	0.01718	0.3497	*clpA*	-8.90	0.050036	279 -- 288	ATPase and specificity subunit of ClpA-ClpP ATP-dependent serine protease chaperone activity
146	0.01722	0.3497	*gmhB*	-7.42	0.111554	134 -- 140	D D-heptose 1 7-bisphosphate phosphatase
147	0.01725	0.3497	*ynfA*	-7.88	0.087741	238 -- 260	UPF0060 family inner membrane protein
148	0.01771	0.3503	*entF*	-8.38	0.067008	16 -- 28	enterobactin synthase multienzyme complex component ATP-dependent
149	0.0178	0.3503	*ispU*	-4.78	0.374870	165 -- 171	undecaprenyl pyrophosphate synthase
150	0.01783	0.3503	*xapR*	-10.29	0.022092	187 -- 209	transcriptional activator of xapAB
151	0.01784	0.3503	*chbG*	-8.65	0.057794	191 -- 211	chito-oligosaccharide deacetylase
152	0.01787	0.3503	*hofM*	-10.56	0.018669	267 -- 288	DNA catabolic putative pilus assembly protein
153	0.01828	0.3546	*amiD*	-13.27	0.003102	48 -- 70	1 6-anhydro-N-acetylmuramyl-L-alanine amidase Zn-dependent; OM lipoprotein
154	0.01832	0.3546	*ftnB*	-9.22	0.041807	199 -- 223	ferritin B putative ferrous iron reservoir
155	0.0185	0.3556	*ubiI*	-8.15	0.076083	177 -- 192	2-octaprenylphenol hydroxylase FAD-dependent
156	0.0187	0.3572	*pal*	-6.59	0.168839	94 -- 112	peptidoglycan-associated outer membrane lipoprotein
157	0.01892	0.3591	*mraZ*	-8.94	0.048989	236 -- 247	RsmH methytransferase inhibitor
158	0.01909	0.36	*rpmG*	-7.48	0.108349	266 -- 278	50S ribosomal subunit protein L33
159	0.01961	0.3614	*bolA*	-7.30	0.119010	211 -- 227	stationary-phase morphogene transcriptional repressor for mreB; also regulator for dacA dacC and ampC
160	0.01967	0.3614	*rpoA*	-8.78	0.053445	284 -- 299	RNA polymerase alpha subunit
161	0.01974	0.3614	*yhfS*	-6.58	0.170301	194 -- 213	FNR-regulated pyridoxal phosphate-dependent aminotransferase family protein
162	0.01977	0.3614	*aceE*	-7.40	0.112724	286 -- 296	pyruvate dehydrogenase decarboxylase component E1 thiamine triphosphate-binding
163	0.01978	0.3614	*ybeL*	-8.48	0.063328	143 -- 161	DUF1451 family protein
164	0.02008	0.3614	*cyoB*	-8.17	0.074878	115 -- 122	cytochrome o ubiquinol oxidase subunit I

165	0.02019	0.3614	*gtrA*	-6.83	0.150535	45 -- 51	CPS-53 (KpLE1) prophage; bactoprenol-linked glucose translocase/flippase
166	0.02022	0.3614	*yaiV*	-7.90	0.086934	187 -- 196	putative transcriptional regulator
167	0.0204	0.3614	*srmB*	-9.32	0.039270	4 -- 10	ATP-dependent RNA helicase
168	0.02049	0.3614	*mutS*	-5.22	0.313678	202 -- 208	methyl-directed mismatch repair protein
169	0.02049	0.3614	*fadE*	-8.60	0.059282	90 -- 96	acyl coenzyme A dehydrogenase
170	0.02096	0.3614	*menA*	-8.35	0.067959	74 -- 81	1 4-dihydroxy-2-naphthoate octaprenyltransferase
171	0.02097	0.3614	*speD*	-8.28	0.070821	282 -- 296	S-adenosylmethionine decarboxylase
172	0.0213	0.3614	*gntK*	-8.41	0.065897	202 -- 208	gluconate kinase 2
173	0.02134	0.3614	*ybdN*	-5.44	0.285374	246 -- 257	PAPS reductase-like domain protein
174	0.02134	0.3614	*ybdN*	-5.44	0.285374	246 -- 257	PAPS reductase-like domain protein
175	0.02135	0.3614	*truA*	-7.10	0.131562	248 -- 254	tRNA pseudouridine(38-40) synthase
176	0.02143	0.3614	*ybjM*	-7.31	0.118191	278 -- 300	inner membrane protein
177	0.02154	0.3614	*clpX*	-8.09	0.078326	164 -- 181	ATPase and specificity subunit of ClpX-ClpP ATP-dependent serine protease
178	0.02159	0.3614	*fumB*	-9.32	0.039296	65 -- 72	anaerobic class I fumarate hydratase (fumarase B)
179	0.02171	0.3614	*torR*	-9.24	0.041136	32 -- 38	response regulator in two-component regulatory system with TorS
180	0.02217	0.3666	*insA*	-5.71	0.254097	78 -- 87	IS1 repressor TnpA
181	0.02236	0.3666	*yigZ*	-8.08	0.078868	31 -- 47	UPF0029 family protein
182	0.02239	0.3666	*yidP*	-8.35	0.068077	84 -- 92	UTRA domain-containing GntR family putative transcriptional regulator
183	0.02281	0.3672	*atpF*	-7.74	0.094434	161 -- 178	F0 sector of membrane-bound ATP synthase subunit b
184	0.02287	0.3672	*acrD*	-5.21	0.315415	237 -- 244	aminoglycoside/multidrug efflux system
185	0.02299	0.3672	*ung*	-4.85	0.364490	132 -- 138	uracil-DNA-glycosylase
186	0.02302	0.3672	*birA*	-9.38	0.037998	178 -- 207	bifunctional biotin-[acetylCoA carboxylase] holoenzyme synthetase/ DNA-binding transcriptional repressor bio-5'-AMP-binding
187	0.02313	0.3672	*ligB*	-7.58	0.103006	251 -- 259	DNA ligase NAD(+)-dependent
188	0.02317	0.3672	*rspA*	-5.86	0.237265	103 -- 123	bifunctional D-altronate/D-mannonate dehydratase

189	0.02343	0.3672	*yceO*	-7.29	0.119547	68 -- 75	uncharacterized protein
190	0.0235	0.3672	*sapA*	-7.06	0.133888	242 -- 248	antimicrobial peptide transport ABC transporter periplasmic binding protein
191	0.02354	0.3672	*dcuA*	-8.99	0.047673	138 -- 145	C4-dicarboxylate antiporter
192	0.02385	0.3701	*fur*	-7.59	0.102480	238 -- 252	ferric iron uptake regulon transcriptional repressor; autorepressor
193	0.02427	0.3728	*yceK*	-8.09	0.078348	216 -- 223	outer membrane integrity lipoprotein
194	0.02431	0.3728	*alsR*	-11.22	0.012388	198 -- 231	d-allose-inducible als operon transcriptional repressor; autorepressor; repressor of rpiR
195	0.02453	0.3728	*ldtB*	-8.31	0.069603	115 -- 145	L D-transpeptidase linking Lpp to murein
196	0.02463	0.3728	*malP*	-8.65	0.057792	226 -- 245	maltodextrin phosphorylase
197	0.02465	0.3728	*ydfR*	-11.27	0.011979	281 -- 292	Qin prophage; uncharacterized protein
198	0.02539	0.3781	*cydD*	-9.01	0.046907	26 -- 49	glutathione/cysteine ABC transporter export permease/ATPase
199	0.02549	0.3781	*cmoB*	-7.72	0.095642	175 -- 198	tRNA (cmo5U34)-carboxymethyltransferase carboxy-SAM-dependent
200	0.02563	0.3781	*fau*	-8.71	0.055834	22 -- 29	5-formyltetrahydrofolate cyclo-ligase family protein

General discussion

5. Chapter 5: General discussion

5.1. General discussion

The work developed in this doctoral dissertation is primarily centered in the study of ribosomal RNA and ribosome biogenesis factors, highlighting its central role in translation regulation. We have uncovered new regulators involved in rRNA processing, folding and degradation pathways. For the first time, we demonstrate that the widely conserved RNA chaperone Hfq acts as a ribosomal assembly factor in bacteria, affecting not only rRNA processing but also ribosome levels. This function is suggested to be independent of its activity as sRNA-regulator. Furthermore, Hfq is found to interact with RNase R, a hydrolytic exoribonuclease. These two proteins cooperate not only in a novel RNA quality control pathway that eliminates superfluous rRNA fragments but also in rRNA maturation. Additionally, we provide evidences that rRNA may act as a reservoir of regulatory small non-coding RNAs. Overall, this work offers a new perspective on translation regulation studies, which are usually focused on messenger RNA.

5.2. Hfq affects ribosome biogenesis and translation fidelity

In Chapter 2 we describe a novel function for the RNA-binding protein Hfq as a ribosome biogenesis factor. Hfq depleted cells display phenotypes commonly found in mutants of ribosome biogenesis factors, such as misprocessing of rRNA and reduced levels of functional 70S ribosomes. We have also found that inactivation of Hfq results in a cold-sensitive phenotype, typically associated with ribosome biogenesis mutants.

Hfq is widely studied for its role in post-transcriptional gene regulation by facilitating basepairing between small non-coding RNAs and messenger RNA. We now expand the known list of Hfq substrates to include rRNA, the most abundant RNA molecule within the cell. This was essential to demonstrate a novel function for Hfq as a ribosome biogenesis factor in bacteria. Inactivation of Hfq results in defects in rRNA processing and ribosome assembly leading to a reduced pool of functional 70S ribosomes. Moreover, our data suggests that the activity of Hfq as a ribosome biogenesis factor is independent of its known function in promoting sRNA-mRNA interactions. Mutations in

the surfaces that Hfq uses to bind sRNAs (proximal face and rim) did not affect ribosome levels, which argues that this new function is independent of its activity as a sRNA mediator.

Despite its importance, nowadays it became debatable if Hfq is truly as important for sRNA biology as initially thought. In *E. coli* the majority of small non-coding RNAs is deemed to bind to Hfq, however, not all sRNAs depend on Hfq for stability and mode of action. In other bacteria, like *Listeria monocytogenes*, Hfq is not even required for the sRNA dependent pathways (Christiansen *et al*, 2006; Rochat *et al*, 2015), suggesting other yet undefined function(s) of Hfq beyond regulation of sRNA activity. Recently new sRNA regulators have been described, such as ProQ and CsrA. Although some small non-coding RNAs can bind both Hfq and at least one of these regulators, other sRNAs bind specifically to ProQ or CsrA and not to Hfq (Smirnov *et al*, 2017; Holmqvist & Vogel, 2018). There is a bulk of information that suggests that Hfq may play other roles beyond the accepted role in promoting the basepairing between sRNA-mRNA.

The 16S rRNA is a large and highly structured molecule that may exhibit alternative non-functional conformations. Ribosome biogenesis factors assist the correct folding of rRNA. RNA chaperones are key players in this process as they can remodel RNA secondary structures. Remarkably, Hfq was initially identified as a host factor required to melt part of the invading RNA from the bacteriophage Qβ, allowing its efficient replication (Franze de Fernandez *et al*, 1968). Our results have now shown that Hfq is required for the correct folding of the 16S rRNA (Andrade *et al*, 2018), affecting the formation of helix 1 and the central pseudoknot, physical structures critical for the overall folding and function of this large RNA molecule (Brink *et al*, 1993). The exacerbated sensitivity of the Δ*hfq* strain to aminoglycosides and the cumulative translation errors induced by *hfq* deletion could be explained by the formation of abnormal 30S subunits bearing misfolded pseudoknot structures (Lodmell & Dahlberg, 1997). We also show that Hfq is required for the maturation of the 16S rRNA. In both scenarios, Hfq may play a role in melting RNA secondary structures assisting not only in the correct 3D structure but also its processing.

Several auxiliary factors associate with the ribosome at different maturation stages during the intricate process of ribosome assembly, assisting in r-protein binding

and rRNA folding steps (Kaczanowska & Rydén-Aulin, 2007). The way that some of these assembly factors act is well known and is possible to establish a hierarchy of events. It has been shown that overexpression of some factors may compensate for the lack of others, which indicates possible overlapping functions (Shajani *et al*, 2011). For example, overexpression of the essential Era GTPase was found to rescue the defects from RbfA deletion, meaning that RbfA probably acts later than Era during ribosome assembly. Moreover, RbfA overexpression partially helped mitigating the effects of RimM inactivation. RbfA and RimM are thus thought to act on a later stage of assembly (Shajani *et al*, 2011). RbfA is one of the best characterized ribosome assembly factors, affecting the maturation of 30S subunits. This protein binds the 5' end of the 16S rRNA and is critical for the formation of the correct folding of helix 1. RbfA has been found to associate with free 30S subunits but not with 70S or polysomes.

There is mounting evidence that RbfA and Hfq act similarly. RbfA and Hfq were shown to bind to the 5' end of the 16S rRNA affecting helix 1 folding (Dammel & Noller, 1995; Bylund *et al*, 1998). Both Δ*hfq* and Δ*rbfA* mutants exhibit an accumulation of unprocessed 17S rRNA, increased levels of immature 30S subunits and a reduction in the pool of 70S ribosomes. As previously described for RbfA (Dammel & Noller, 1995), we now show that Hfq also specifically binds to immature 30S subunits and do not associate with mature 30S. Therefore, it seems likely that Hfq act as a late stage 30S subunit assembly factor in close resemblance to RbfA. This fact raises the question whether Hfq and RbfA may compensate for the absence of each other. Additional questions that raises from this work concerns the timing of Hfq binding to pre-30S complexes during 30S biogenesis and how Hfq interacts with the RNA and protein components of the ribosome.

We show that the absence of Hfq has a global impact in translation. This novel function of Hfq provides an additional explanation for the pleiotropic effects of Hfq deletion on bacterial growth and stress response beyond the widely characterized role as sRNA regulator. As expected for an important ribosome biogenesis factor, we found that Hfq exhibits features of aberrantly assembled ribosomes such as compromised translation efficiency and fidelity. Interestingly, ribosome profiling analysis revealed that inactivation of Hfq particularly affected the protein synthesis of many translation-related gene classes.

5.3. Hfq and RNase R cooperate in a novel rRNA quality control pathway

Further work demonstrated that Hfq interacts with RNase R, a nuclease that also has rRNA as a substrate. RNase R is very effective in the degradation of highly structure RNA molecules. This is a feature that distinguishes RNase R from the remaining exoribonucleases, like RNase II and PNPase, as these get stalled in the presence of stable stem-loops. Therefore, it is not surprising to find that amongst 3'-5' exoribonucleases, RNase R exhibits the highest affinity towards rRNA (Cheng & Deutscher, 2002).

Interaction between Hfq and RNase R was found to be important for the hitherto unrecognized RNA quality control pathway that is key for the elimination of aberrant rRNA fragments. The simultaneous inactivation of Hfq and RNase R results in the accumulation of rRNA fragments derived from 16S and 23S rRNAs when compared to the wild-type. The build up of these fragments was shown to be deleterious for the cell (Cheng & Deutscher, 2003). In fact, the double mutant Δ*hfq* Δ*rnr* exhibits a strong growth impairment worse than any of the single mutants. How to explain that RNase R may required association with Hfq for the degradation of these fragments? Although RNase R is unique in the degradation of structured molecules, it exhibits a requirement for 3' RNA overhangs to initiate unwinding and processive degradation (Matos *et al*, 2009; Vincent & Deutscher, 2009). It seems plausible that Hfq, as an RNA chaperone, may bind to the structured rRNA fragments and remodels the 3' end producing a linear stretch to which RNase R is then able to bind and initiate degradation. Although Hfq and RNase R are required for the elimination of these rRNA fragments, these two enzymes are probably not involved in their production. Endoribonucleases, like RNase E and RNase III, are thought to be the main responsible for this (Zundel *et al*, 2009; Basturea *et al*, 2011; Sulthana *et al*, 2016).

Furthermore, our results are in line with the report that Hfq and RNase R can bind to the r-protein S12 in the small subunit in stationary phase *E. coli* cells (Strader *et al*, 2013). S12 is a key mediator of fidelity of translation in both prokaryotes and eukaryotes and is positioned in helix 44 of the 16S rRNA that is known to form extensive contacts with the large subunit (Yusupov et al., 2001; Cukras et al., 2003). Moreover, helix 44 houses the anti-SD sequence vital for correct codon/anticodon interactions. Consequently, association of Hfq and RNase R with S12 suggests an important role for maintaining the

correct fold of helix 44 in the mature 16S rRNA. The association of Hfq and RNase R with S12 may also provide insights to the observed increased sensitivity of the Δ*hfq* strain to sub-lethal dosages of aminoglycosides. Aminoglycosides specifically target helix 44 at the decoding center, altering the translation proof-reading leading to decreased translation fidelity.

Hence, interaction of Hfq with an RNase may constitute an advantage for modulation of RNA stability, namely of structured RNAs. This seems to be more frequent than what was initially envisaged. Indeed, Hfq is found to associate with other enzymes of the RNA degradation machinery besides RNase R. One of these enzymes is the endonuclease RNase E. (Morita *et al*, 2005; Ikeda *et al*, 2011). Hfq was found to copurify with RNase E and this complex was proposed to be important to guide RNase E to the target mRNA by specific regulatory sRNAs (Morita & Aiba, 2011). RNase E provides a scaffold for a large multiprotein complex involved in RNA degradation that is known as the degradosome. Hfq was found to copurify specifically with RNase E but not with other components of the degradosome. However, one other report shows that Hfq copurifies with the phosphorolytic 3'-5' exoribonuclease PNPase, one of the members of the degradosome (Mohanty *et al*, 2004). This suggests that Hfq may form complexes with PNPase in a degradosome-independent manner. The biological meaning of Hfq/PNPase interaction is still not fully understood, but it may confer an advantage in the degradation of small non-coding RNAs, one other class of usually highly structured RNAs (Andrade & Arraiano, 2008; Andrade *et al*, 2012, 2013; De Lay & Gottesman, 2011; Bandyra *et al*, 2016). In addition, Hfq can also form a complex with poly(A) polymerase I (PAP I) (Mohanty *et al*, 2004). This association may explain why Hfq was shown to stimulate the biosynthetic activity of PAP I on mRNAs (Hajnsdorf & Régnier, 2000; Le Derout *et al*, 2003). Furthermore, Hfq was shown to bind to elongated poly(A) tails protecting the mRNA from exonucleolytic degradation (Folichon *et al*, 2005).

Following the work developed in Chapter 2 we hypothesized that cooperation between Hfq and RNase R could also affect rRNA processing. Remarkably, we found that Hfq and RNase R are required for the correct maturation of the 16S and 23S rRNAs. The precursors of 16S and 23S rRNA were found to strongly accumulate in the Δ*hfq* Δ*rnr*

double mutant when compare either to wild-type or any of the single mutants. These defects in rRNA processing were correlated with an altered ribosomal profile upon inactivation of Hfq and RNase R. Accumulation of 17S rRNA is consistent with increasing amounts of immature 30S small subunits whereas higher levels of pre-23S rRNA correlates with the increase in immature 50S subunits in the double mutant. Consequently, the Δ*hfq* Δ*rnr* strain exhibits reduced levels of 70S ribosomes. Therefore, defects in ribosome synthesis leads to a decrease in functional ribosomes that impairs translation and negatively impacts cellular fitness cascades (Connolly et al., 2008; Sharpe Elles et al., 2009; Leong et al., 2013). The decrease in the pool of 70S particles may be a result of assembly defects from the damaged subunits and/or increased degradation of the faulty ribosomes. A quality control system in which the endonuclease YbeY targets defective ribosomes with immature 30S subunits was recently described (Jacob *et al*, 2013). To the best of our knowledge no such system has been described to target ribosomes with immature 50S subunits.

5.4. Precursor rRNA-derived fragments arise from 16S rRNA processing

The work described in this dissertation highlights the important role of rRNA processing. rRNA misprocessing events result in the accumulation of precursors that lead to the synthesis of immature subunits and defective ribosomes.

rRNA processing is commonly seen only as a mean to produce the mature and functional forms of rRNAs. The excised rRNA precursor fragments are commonly regarded as mere dispensable byproducts whose fate is its rapid elimination. In a word, these sequences were simply viewed as “junk” RNA. However, this work shows that fragments originated from the precursor rRNA accumulate to high levels in fast growing cells. Furthermore, we suggest that these precursor rRNA fragments have the potential to act as regulatory RNAs. Strikingly, rRNAs may act as reservoirs of small RNAs.

Small regulatory RNAs are firmly established as important post-transcriptional regulators that govern cell physiology. In bacteria, sRNA-mediated regulation enables the cell to quickly change its transcriptome in order to maintain homeostasis. The first small RNAs to be identified were originated from intergenic sequences. Later on, alternative

sRNA biogenesis pathways have been described that greatly extended the list of known sources of sRNAs. These regulatory RNAs may be expressed from independent transcriptional units or result from processing of other transcripts. In addition to the intergenic regions sRNAs were found to derive not only from 5' UTRs and 3' UTRs (Kawano *et al*, 2005; Chao & Vogel, 2016; Miyakoshi *et al*, 2015) but also from the coding sequence of some mRNAs (Dar & Sorek, 2018). Nevertheless, the origin of sRNAs is not confined to mRNA and other RNA molecules may act as non-canonical sources of sRNAs, this is the case of transfer RNAs. tRNAs can be transcribed as polycistronic RNAs that undergo processing steps to yield their mature functional forms. The 3' external transcribed sequence of the $tRNA^{LeuZ}$ was shown to harbor a functional sRNA molecule capable of modulation the action of other sRNAs, that is generated by normal processing of the tRNA (Lalaouna *et al*, 2015). The functional RNAs that originate from the tRNA precursor sequences were termed tRNA-derived fragments (tRFs) (Kim *et al*, 2017).

We propose that a similar mechanism is occurring with the processing of rRNA precursors. With this work we demonstrated that the precursor flanking sequences excised from the 17S rRNA precursor are stable and are predicted to possess regulatory potential. Accordingly, these are precursor rRNA-derived fragments as rRFs. After processing the 17S rRNA precursor originates two rRFs: the rRF-115 corresponds 5' end flanking sequence of the 17S rRNA precursor, while the rRF-33 corresponds to the 3' end flanking sequence.

A set of different software prediction tools identified a list of target candidates that exhibit strong hybridization energies with these rRFs. The top candidates included not only mRNAs but as well as other sRNAs. We propose two mechanisms by which rRF-115 and rRF-33 could carry out their regulatory roles depending on the target. rRFs could interact with mRNAs to modulate their translation and stability, as commonly observed for other sRNAs. However, if the target is a small non-coding RNA it is possible that rRFs may behave as molecular sponges.

The levels of rRFs are dependent on the levels of rRNA that depend on the metabolic state of the cell. In fast growing cells rRNA is highly synthesized and consequently processed, giving raise to high levels of rRFs. In contrast, in non-dividing

cells rRNA synthesis is low and so is rRF levels. This suggests that precursor rRNA-derived fragments may act as regulatory molecules that modulate the metabolic state of the cell. We propose that modulation of the Csr system by rRFs might be part of a genetic network controlling cellular growth. A computational study revealed that rRF-115 and rRF-33 have multiple predicted binding sites in the sRNA CsrB, a modulator of the carbon storage regulator A (CsrA). CsrA is a transcriptional repressor that controls the metabolic state of the cell. Amongst its targets, it negatively regulates the expression of the glycogen biosynthetic gene *glgC* and a carbon starvation-induced gene *cstA*. The small non-coding regulatory RNA CsrB contains multiple CsrA binding sites and functions as CsrA antagonists by sequestering this protein. In fast growing cells rRF-115 and rRF-33 can potentially bind and sequester CsrB which results in the release of CsrA. When conditions change and the cell stops growing, the levels of rRFs decrease and consequently CsrB is free to sequester CsrA, that no longer can act as a repressor of carbon storage genes. Future work has to be performed in order to validate this model.

Hfq is able to bind to these sequences in vitro, as demonstrated by the EMSA experiments in Chapter 2. Accordingly, rRF binding to RNA chaperones like Hfq could pose as a discriminatory decision of whether rRFs would target mRNAs or other regulatory RNA molecules. Hfq could catalyze the imperfect basepairing between rRFs and their mRNA targets, which could lead to positive or negative regulation of the transcript through a variety of pathways, as described for other sRNAs (Vogel & Luisi, 2011; Holmqvist & Vogel, 2018). On the other hand, rRFs that are not bound to Hfq could behave as sRNA sponges, as we proposed for the Csr system.

Due to their vital biological functions, rRNA genes are maintained in multiple copy numbers in all organisms (Kaczanowska & Rydén-Aulin, 2007). The *E. coli* genome bears 7 rDNA operons which contribute to the extremely high levels of rRNAs within the cell. However, not all rDNA operons are equally transcribed, and their transcriptional regulation is tightly coupled with cell physiology, leading to heterogeneity in the pool of 70S ribosomes (Byrgazov *et al*, 2013). This further corroborates the high regulatory potential of rRNA-derived sequences. Moreover, a recent report shows that endogenous rRNA sequence variation alone is able to modulate ribosomes and gene expression in *E.*

coli to respond to specific stresses (Kurylo *et al*, 2018). It is therefore conceivable that specialized functional rRNA-derived fragments could arise from the processing of the long transcript of this rRNA operon.

We believe that identification of rRNA precursors as reservoirs of regulatory sRNAs offers a promising strategy to control gene expression. This seems to be a conserved regulatory mechanism since rRFs were detected using high-throughput approaches on eukaryotic cells; however, no functions have been characterized up to the moment (Asha & Soniya, 2017).

Ribosomal RNA constitutes more than 80% of total RNA within the cell. Despite this large abundance, most of the high throughput current studies neglected rRNA. One of the first steps in RNA sequencing is removal of "contaminating" rRNA. Consequently, information regarding rRNA biology is scarce. The work described in this dissertation expanded our knowledge on the regulatory features of rRNA. We expanded the list of natural substrates of the widely conserved Hfq RNA chaperone, that now includes ribosomal RNA. We show that Hfq is a central regulator affecting different levels of ribosome biogenesis, that include not only the processing but also ribosome assembly. In addition, we showed that Hfq is able to associate with another rRNA binding enzyme, the exoribonuclease RNase R. This association may provide a synergistic effect not only in the processing of rRNA but also in the degradation of rRNA fragments. Future work will certainly elucidate the potential regulatory role of the predicted precursor rRNA-derived fragments. The first step to this is validation of the predicted targets and elucidation of regulatory networks.

5.5. References

Andrade JM, Pobre V & Arraiano CM (2013) Small RNA modules confer different stabilities and interact differently with multiple targets. *PLoS One* **8:** e52866

Andrade JM, Pobre V, Matos AM & Arraiano CM (2012) The crucial role of PNPase in the degradation of small RNAs that are not associated with Hfq. *RNA* **18:** 844–55

Andrade JM, Dos Santos RF, Chelysheva I, Ignatova Z & Arraiano CM (2018) The RNA-binding protein Hfq is important for ribosome biogenesis and affects translation fidelity. *EMBO J.* **37:** e97631

Andrade JM & Arraiano CM (2008) PNPase is a key player in the regulation of small RNAs that control the expression of outer membrane proteins. *RNA* **14:** 543–51

Asha S & Soniya EV (2017) The sRNAome mining revealed existence of unique signature small RNAs derived from 5.8SrRNA from Piper nigrum and other plant lineages. *Sci. Rep.* **7:** 41052

Bandyra KJ, Sinha D, Syrjanen J, Luisi BF & De Lay NR (2016) The ribonuclease polynucleotide phosphorylase can interact with small regulatory RNAs in both protective and degradative modes. *RNA* **22:** 360–372

Basturea GN, Zundel MA & Deutscher MP (2011) Degradation of ribosomal RNA during starvation: comparison to quality control during steady-state growth and a role for RNase PH. *RNA* **17:** 338–45

Brink MF, Verbeet MP & de Boer HA (1993) Formation of the central pseudoknot in 16S rRNA is essential for initiation of translation. *EMBO J.* **12:** 3987–96

Bylund GO, Wipemo LC, Lundberg LA & Wikström PM (1998) RimM and RbfA are essential for efficient processing of 16S rRNA in *Escherichia coli*. *J. Bacteriol.* **180:** 73–82

Byrgazov K, Vesper O & Moll I (2013) Ribosome heterogeneity: another level of complexity in bacterial translation regulation. *Curr. Opin. Microbiol.* **16:** 133–9

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

Cheng Z-F & Deutscher MP (2003) Quality control of ribosomal RNA mediated by polynucleotide phosphorylase and RNase R. *Proc. Natl. Acad. Sci. U. S. A.* **100:** 6388–93

Cheng Z-FF & Deutscher MP (2002) Purification and characterization of the *Escherichia coli* exoribonuclease RNase R. Comparison with RNase II. *J. Biol. Chem.* **277:** 21624–9

Christiansen JK, Nielsen JS, Ebersbach T, Valentin-Hansen P, Søgaard-Andersen L & Kallipolitis BH (2006) Identification of small Hfq-binding RNAs in *Listeria monocytogenes*. *RNA* **12:** 1383–96

Connolly K, Rife JP & Culver G (2008) Mechanistic insight into the ribosome biogenesis functions of the ancient protein KsgA. *Mol. Microbiol.* **70:** 1062–75

Cukras AR, Southworth DR, Brunelle JL, Culver GM & Green R (2003) Ribosomal proteins S12 and S13 function as control elements for translocation of the mRNA:tRNA complex. *Mol. Cell* **12:** 321–8

Dammel CS & Noller HF (1995) Suppression of a cold-sensitive mutation in 16S rRNA by overexpression of a novel ribosome-binding factor, RbfA. *Genes Dev.* **9:** 626–37

Dar D & Sorek R (2018) Bacterial noncoding RNAs excised from within protein-coding transcripts. *MBio* **9:** e01730-18

Le Derout J, Folichon M, Briani F, Dehò G, Régnier P & Hajnsdorf E (2003) Hfq affects the length and the frequency of short oligo(A) tails at the 3' end of *Escherichia coli rpsO* mRNAs. *Nucleic Acids Res.* **31:** 4017–23

Folichon M, Allemand F, Régnier P & 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:** 454–63

Franze de Fernandez MT, Eoyang L & August JT (1968) Factor fraction required for the synthesis of bacteriophage Qbeta-RNA. *Nature* **219:** 588–90

Hajnsdorf E & 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:** 1501–5

Holmqvist E & Vogel J (2018) RNA-binding proteins in bacteria. *Nat. Rev. Microbiol.* **16:** 601–615

Ikeda Y, Yagi M, Morita T & Aiba H (2011) Hfq binding at RhlB-recognition region of RNase E is crucial for the rapid degradation of target mRNAs mediated by sRNAs in *Escherichia coli*. *Mol. Microbiol.* **79:** 419–432

Jacob AI, Köhrer C, Davies BW, RajBhandary UL & Walker GC (2013) Conserved bacterial RNase YbeY plays key roles in 70S ribosome quality control and 16S rRNA maturation. *Mol. Cell* **49:** 427–38

Kaczanowska M & Rydén-Aulin M (2007) Ribosome biogenesis and the translation process in *Escherichia coli*. *Microbiol. Mol. Biol. Rev.* **71:** 477–94

Kawano M, Reynolds AA, Miranda-Rios J & Storz G (2005) Detection of 5'- and 3'-UTR-derived small RNAs and cis-encoded antisense RNAs in *Escherichia coli*. *Nucleic Acids Res.* **33:** 1040–1050

Kim HK, Fuchs G, Wang S, Wei W, Zhang Y, Park H, Roy-Chaudhuri B, Li P, Xu J, Chu K, Zhang F, Chua M-S, So S, Zhang QC, Sarnow P & Kay MA (2017) A transfer-RNA-derived small RNA regulates ribosome biogenesis. *Nature* **552:** 57–62

Kurylo CM, Parks MM, Juette MF, Zinshteyn B, Altman RB, Thibado JK, Vincent CT & Blanchard SC (2018) Endogenous rRNA Sequence Variation Can Regulate Stress Response Gene Expression and Phenotype. *Cell Rep.* **25:** 236–248.e6

Lalaouna D, Carrier M-C, Semsey S, Brouard J-S, Wang J, Wade JTT & Massé E (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

De Lay N & Gottesman S (2011) Role of polynucleotide phosphorylase in sRNA function in *Escherichia coli*. *Rna* **17:** 1172–1189

Leong V, Kent M, Jomaa A & Ortega J (2013) *Escherichia coli rimM* and *yjeQ* null strains accumulate immature 30S subunits of similar structure and protein complement. *RNA* **19:** 789–802

Lodmell JS & Dahlberg AE (1997) A conformational switch in *Escherichia coli* 16S ribosomal RNA during decoding of messenger RNA. *Science* **277:** 1262–7

Matos RG, Barbas A & Arraiano CM (2009) RNase R mutants elucidate the catalysis of structured RNA: RNA-binding domains select the RNAs targeted for degradation. *Biochem. J.* **423:** 291–301

Miyakoshi M, Chao Y & Vogel J (2015) Cross talk between ABC transporter mRNAs via a target mRNA-derived sponge of the GcvB small RNA. *EMBO J.* **34:** 1478–1492

Mohanty BK, Maples VF & Kushner SR (2004) The Sm-like protein Hfq regulates

polyadenylation dependent mRNA decay in *Escherichia coli*. *Mol. Microbiol.* **54:** 905–20

Morita T & 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:** 294–8

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

Rochat T, Delumeau O, Figueroa-Bossi N, Noirot P, Bossi L, Dervyn E & Bouloc P (2015) Tracking the Elusive Function of *Bacillus subtilis* Hfq. *PLoS One* **10:** e0124977

Shajani Z, Sykes MT & Williamson JR (2011) Assembly of bacterial ribosomes. *Annu. Rev. Biochem.* **80:** 501–26

Sharpe Elles LM, Sykes MT, Williamson JR & Uhlenbeck OC (2009) A dominant negative mutant of the *E. coli* RNA helicase DbpA blocks assembly of the 50S ribosomal subunit. *Nucleic Acids Res.* **37:** 6503–14

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

Strader MB, Hervey WJ, Costantino N, Fujigaki S, Chen CY, Akal-Strader A, Ihunnah CA, Makusky AJ, Court DL, Markey SP & Kowalak JA (2013) A coordinated proteomic approach for identifying proteins that interact with the *E. coli* ribosomal protein S12. *J. Proteome Res.* **12:** 1289–99

Sulthana S, Basturea GN & Deutscher MP (2016) Elucidation of pathways of ribosomal RNA degradation: an essential role for RNase E. *RNA* **22:** 1163–71

Vincent HA & Deutscher MP (2009) Insights into how RNase R degrades structured RNA: analysis of the nuclease domain. *J. Mol. Biol.* **387:** 570–83

Vogel J & Luisi BF (2011) Hfq and its constellation of RNA. *Nat. Rev. Microbiol.* **9:** 578–89

Yusupov MM, Yusupova GZ, Baucom A, Lieberman K, Earnest TN, Cate JH & Noller HF (2001) Crystal structure of the ribosome at 5.5 A resolution. *Science* **292:** 883–96

Zundel MA, Basturea GN & Deutscher MP (2009) Initiation of ribosome degradation during starvation in *Escherichia coli*. *RNA* **15:** 977–83