PhD thesis
Mette Kongstad

Bacterial growth physiology
With focus on the tRNA-linked-repeats and tRNA regulation during starvation

Submitted: 30/04/2016

Supervisor: Michael Askvad Sørensen, Co-supervisor: Sine Lo Svenningsen
Institute: Institute of Biology

Name of department: Department Biomolecular Science

Author: Mette Kongstad, M.Sc. Nano Science

Title: Bacterial growth Physiology

Subtitle: With focus on the tRNA-linked-repeats and tRNA regulation during starvation

Supervisor: Michael Askvad Sørensen, Associated Professor Department of Biology, University of Copenhagen, Denmark

Co supervisor: Sine Lo Svenningsen, Associated Professor Department of Biology, University of Copenhagen, Denmark

Assessment committee: Anders Løbner-Olsen, Professor Department of Biology, University of Copenhagen, Denmark

Katherine Long, Senior Scientist DTU Biosustain, Danmarks Tekniske Universitet, Denmark

Kai Papenfort, Professor Department of Biology I, Microbiology Ludwig-Maximilians Universität München, Germany

Submitted: 30. April 2016

Cover: Lord Louí
Acknowledgements

First, I would gratefully thank my supervisor Michael Askvad Sørensen for always taking the time to discuss the project and for excellent guidance along the way, together with my co-supervisor Sine Lo Svenningsen who has always been supportive, especially during the period of writing. Secondly, I would like to thank Brigitte H. Kallipolitis for her help with interpretation of the in vitro work and Godefroid Charbon for his help with the phenotypical characterization, Pilar Menéndez Gil for her guidance in the vitro work, and Anders Boysen for supply of Hfq. Furthermore, I would like to thank the members of our group; especially Thomas S. Stenum, Ana Josefa Muñoz Gomez and Marit Warren for valuable discussions about experimental set up and for making the laboratory work joyful, but also, Anna-Leigh Fehler, Sarah Dupont Søgaard, Iman Mahmoud Kurdi, Michael Sass Eskebjerg, Ida Sjøgaard, Nina Molin Høyland-Kroghsbo, Maria Schaub, Bertil Gummesson, Mathias Isager Fessler, Esben S Juul Løvenhardt Sørensen, Katrine Faxøe Nannestad, for being wonderful group mates.

I would also like to thank my current and former office and laboratory friends and colleagues at the University of Copenhagen, especially Patricia Dominguez Cuevas, Paul Santner, Jakob Frimodt-Møller, Louise Bjørn, Pia Lykke Hovgård Christensen, Christopher Campion, Elsa Monique Albine Maisonneuve, Georgia Hann, Kathryn Turnbull, Mohammad Roghanian, Sidsel Henriksen, Stanley Brown, Steen Pedersen, Thomas Thyge Thomsen and Søren Lindemose. But also colleagues from the Southern Danish University: Jakob Møller-Jensen, Maria Storm Mollerup, Joe Ross, Stine Lambert Nielsen, Eva Maria Sternkopf Lillebæk for being welcoming and helpful.

Finally, I would like to thank my housebound Christian Krarup, our children Sally Marie Kongstad-Krarup, Loui Kongstad-Krarup, Thomas Krarup, Felix Krarup, my mother Elisabet Kongstad, her housebound Uwe Koch, my mother in law Lone Krarup, my friends Ditte Anine Bertram, Charlotte Westi, for love, support, help throughout the good and bad times and not the least their patience with me.

Mette Kongstad, April 2016
Abstract

Transfer RNA (tRNA) mediates the translation of the code, transiently stored in the messenger RNA (mRNA), to the final protein. The entity of tRNA has for decades been assumed to be stable for hours in any circumstance, but my supervisor Michael A. Sørensen noticed during his work with charging level of tRNA that the tRNA level was seemingly reduced much faster than observed before, after the induction of amino acid starvation. Therefore, we developed a method for testing this observation. It turned out that tRNA degrades to ~75% of the initial level within 40 minutes after induction of amino acid starvation. Our next task was to explore the mechanism behind the starvation-induced downregulation of tRNA, and the obvious candidates as RelA, SpoT, and the E.coli toxin-anti toxin system (TA) were tested with no effect. Interestingly, the lifetime of tRNA was prolonged in an isogenic Δhfq after the induction of amino acid starvation. Indeed, the RNA chaperone Hfq is known to facilitate a pairing of small RNA (sRNA) with their target mRNA, affecting gene expression. Thus, our attention was drawn towards sRNA and other regulatory RNAs. The tRNA-linked-repeats (TLR) acting as regulatory RNA (regRNA) in collaboration with Hfq seemed to be good candidates for starvation-induced down regulation of tRNA. This is because TLRs are small intergenic repeats located in operons of tRNA and rRNA that share 18-19 nt of homology with the RNA they are located next to with yet an unknown function. It turned out that the TLRs were not involved in starvation-induced down regulation of tRNA. As such, my project was subdivided into the following segments:

1 Part: starvation-induced regulation of tRNA

2 Part: TLRs.

The TLRs were not involved in tRNA regulation. Instead, three TLRs showed themselves to be enriched at Hfq, and these TLRs displayed increased stability upon Hfq interaction. From this and other experiments, we propose that the three TLRs transcribed from the rrnA, rrnB, and rrnD operons act as regRNAs, assisted by Hfq.
## TABLE OF CONTENT

ACKNOWLEDGEMENTS 3

ABSTRACT 4

MOTIVATION 7

OVERVIEW OF THESIS 7

1 PART 9

1.1 INTRODUCTION: tRNA REGULATION 9

1.1.1 Cell physiology and steady-state growth 9

1.1.1.1 Defining steady state 9

1.1.1.2 The choice of growth medium and reproducibility 9

1.1.1.3 Macromolecular composition during steady state 10

1.1.2 Cell physiology and starvation 12

1.1.2.1 Methods of inducing amino acid starvation 12

1.1.2.2 Triggers of (p)ppGpp induction 14

1.1.2.3 RelA and SpoT catalyzes the synthesis of (p)ppGpp 15

1.1.2.4 How SpoT and RelA is activated 15

1.1.2.5 The effect of (p)ppGpp 17

1.1.3 tRNA 18

1.1.3.1 Canonical function of tRNA 18

1.1.3.2 Error of translation 19

1.1.3.3 Error frequency and level of aminoacyl tRNA 21

1.1.4 RNases 22

PAPER 1 25

1.2 DISCUSSION 54

1.2.1 Normalization 54
1.2.2 **RNases involved** 55

1.2.3 **Our hypothesis in context** 56

1.3 **CONCLUSION** 58

1.4 **FUTURE PERSPECTIVES** 59

## 2 PART 61

### 2.1 INTRODUCTION: TLR AND Hfq 61

#### 2.1.1 The family of tRNA-linked repeats 61

#### 2.1.2 Regulatory RNA in general 64

#### 2.1.3 The chaperone Hfq 65

- 2.1.3.1 Conservation of Hfq 65
- 2.1.3.2 Structure of Hfq 65
- 2.1.3.3 Properties of Hfq 67
- 2.1.3.4 Cellular concentration and location of Hfq 69
- 2.1.3.5 RNA cycling at Hfq 71

### PAPER 2 73

#### 2.2 DISCUSSION 115

- 2.2.1 The function of TLR 115
- 2.2.2 Phenotypic characterization 115
- 2.2.3 Homology 116
- 2.2.4 Hfq 116
- 2.2.5 Possible errors with the experiments 117
  - 2.2.5.1 C-terminal Epitope-tagged Hfq 117
  - 2.2.5.2 Co-precipitation and re-association of lysis components 117

#### 2.3 CONCLUSION 118

#### 2.4 FUTURE PERSPECTIVES 118
Motivation

The primary habitat of *Escherichia coli* (*E.coli*) is the gut of mammals. As such, *E.coli* mainly live as an aerobic organism in symbiosis with its host. When *E.coli* is challenged with a sudden change, as when it is transferred from the gut to open environments, the availability of nutrients and energy sources are often limited or changed, and the cell needs to adapt quickly in order to survive (van Elsas et al., 2011). Exploring bacteria adaptation therefore increases the very basic and important knowledge of bacteria life and survival strategies. Gaining insight to this field has indeed been very fascinating. Hopefully, the mechanism that we have explored and the little progress we have made will be a useful piece in the puzzle of understanding the amazing life of the bacteria.

OVERVIEW OF thesis

1. Part

The stringent response is crucial for understanding the global regulatory processes of bacteria (Durfee et al., 2008). We propose that the ability to regulate the level of tRNA is essential in the first minutes of the stringent response, and the stringent response can be triggered by amino acid starvation. Consequently, the main focus in the introduction of the first part is as follows: methods for inducing the stringent response by amino acid starvation, how the bacteria respond to the starvation, tRNA and error in translation, and how the level between the different tRNA affects the error frequency. At last, a description of the RNases mainly responsible for RNA turnover is included. After Paper 1, ‘Transfer RNA is rapidly degraded’, I have tried to bring our results into the context of what is former known about the stringent response, then I conclude and the first part is closed by future perspective.
In order to perform reproducible starvation experiments, the culture of growing bacteria must be in steady-state growth. Therefore, the first part of this study is started by briefly describing fundamentals about bacteria growth physiology.

2. Part

The subject of the second part is focused on the transcripts belonging to a family of the TLR and at the RNA chaperone Hfq. The second part of the report starts with a summary of previous literature about the family of TLR. This section is then followed by a brief description of regRNA in general. Since we found three of the TLRs to be enriched at Hfq, Hfq and RNA interaction are treated. This includes the structural basis for the RNA-binding properties of Hfq, the cellular concentration, and the location of Hfq. After Paper 2, ‘Transcripts from the family of tRNA-linked repeats interact with Hfq’, I discuss our attempts to characterize the phenotype of a mutant lacking the TLRs. I do this in order to gain insight into their function. In addition, a lack of basic knowledge of Hfq is discussed, and at last, I conclude and suggest few experiments that perhaps could increase our basic knowledge of the TLRs and Hfq.
1 Part

1.1 Introduction: tRNA regulation

1.1.1 Cell physiology and steady-state growth

One important way of measuring cell fitness is to record unicellular growth development over time at a certain set of conditions (Monod, 1949). According to Monod, the grow phases of the cultured bacteria can be subdivided into six phases: 1) lag phase (no growth), 2) acceleration phase (growth rate increases), 3) exponential phase (growth rate constant), 4) retardation phase (growth rate decrease), 5) stationary phase (no growth), and 6) decline (negative growth) (Monod, 1949). During the exponential phase, the biomass \( x \) of the culture increases, and the exponential rate constant \( (\mu) \) – also named the specific growth constant – can be described by the equation: 

\[
\mu = \frac{\ln(x_2) - \ln(x_1)}{t_2 - t_1},
\]

where \( t_2 - t_1 \) is any time interval during exponential phase, and \( x_2 \) and \( x_1 \) are biomass at time \( t_2 \) and \( t_1 \) respectively. The time passing before the biomass is doubled \( (T_2) \) is therefore given by \( T_2 = \ln(2)/\mu \).

1.1.1.1 Defining steady state

Since reproducibility is crucial, physiological experiments with bacteria are mostly (or at least they are claimed to be) done at a steady state of growth. A rigorous definition of steady-state growth was made by Fishov et al.: Steady-state growth is a time invariant state of a system that is open to its environment; it implies that all various components that make up the cells are increasing by the same rate, including cell number and mass. (Fishov, Zaritsky, & Grover, 1995). Steady-state growth includes balanced growth, but the opposite is not true (Hadas, Einav, Fishov, & Zaritsky, 1995). Also, exponential growth does not imply steady-state growth or balanced growth (Pritchard & Zaritsky, 1970).

1.1.1.2 The choice of growth medium and reproducibility

Since we use the composition of the media to perturb the state of the cell, a brief introduction to the medium and the composition is given here. *E.coli* belongs to a group of chemoorganoheterotrophs (Madigan, 2012). This is a group of organisms that uses organic compounds as carbon/energy sources and also needs other nutrients such as nitrogen, phosphorus, sulfur, potassium, and magnesium in building up the biomass. Obviously, the
bacteria growth medium should contain these components, and a defined medium typically selects one source (often glucose) as a limiting factor present at such a level that change in pH and the accumulation of metabolite are largely eliminated during the period of steady-state growth. The rest of the components must be in excess in order to ensure reproducible growth (Egli, 2015; Monod, 1949; Shehata and Marr, 1971). An example of a defined medium for enterobacteria is the MOPS medium, developed by Neidhardt (Neidhardt et al., 1974b).

Experiments could also be performed in a complex medium such as the Luria-Bertani broth (LB), but such a medium is difficult to use if the intention is to get the growing cells into a reproducible steady state. There are several reasons for this: 1) In LB, the growth rate is often so fast that the time for conducting experiments in steady state is too short; 2) There can be variations from batch to batch, and the composition of the medium can change over time when stored; besides this, the composition of the medium can change during autoclaving; 3) Most importantly, the cells are growing by using amino acids as carbon and energy sources. Since the specific stoichiometry of amino acids in the medium is unknown, it is hard to know which amino acid the cell has chosen as its carbon/energy source. Also, when limitations for that particular amino acid occur, the metabolic state of the cell is hard to predict (Egli, 2015; Sezonov et al., 2007).

1.1.1.3 Macromolecular composition during steady state

The growth rate, defined as the number of doublings per hour, can be calculated by $60/T_2$, if $T_2$ is given in minutes. The higher the growth rate, the higher the average cell size for steady-state growing *Salmonella typhimurium* (Schaechter et al., 1958). The mean cell volume of the bacteria is not only determined by the growth rate but also the specific macromolecular composition and size of the bacteria, dependents on the growth condition, and nature of the nutrient that limits the growth (Egli, 2015; Vadia and Levin, 2015). The most abundant macromolecules in *Salmonella typhimurium* are proteins, making up more than half of the cellular biomass, and the amount of protein per dry cell mass increases at higher growth rates (Maaløe and Kjeldgaard, 1966b). Dennis et al. measured the *E.coli B/r* steady-state composition of macromolecules in minimal media supplemented with glucose/amino acids, glucose, or succinate. For the three supplements, they recorded the bacteria growth rate and related this growth rate to the cellular composition of macromolecules. (See Fig. 1).
Figure 1. Cellular composition of macromolecules: The cellular amount of DNA, RNA, and proteins at different growth rates; for more information, see text. The figure is adapted from (Dennis and Bremer, 1974).

The absorption at 460 nm (A_{460}) per cell is assumed to be proportional to bacterial biomass per cell and is increasing non-exponentially as a function of the growth rate. Interestingly, proteins (calculated as numbers of amino acids per cell) and DNA (calculated as genomes per cell) follow the same slope when the growth rate is larger than 1.2. This suggests that DNA replication is coupled with the amount of protein (or mass). Moreover, the amount of total RNA (calculated as numbers of ribo-nucleotides per cell) increases faster than the growth rate. This suggests that one of the components filling up the larger cells at the higher growth rates is RNA. From these data, the researchers also deduced that the level of mRNA decreases as the growth rates increase, showing that at the highest growth rates, the cell focuses on producing rRNA (ribosomes) and tRNA in order to divide at such a high rate. Importantly, the internal ratio of stable RNA: rRNA / (rRNA + tRNA) is constant at both high and low growth rates. This indicates that the ratio between tRNA and rRNA is somehow important for the cell (Dennis, 1972; Dennis and Bremer, 1974).
1.1.2 Cell physiology and starvation

So far, only the growth of cells in steady state has been discussed. Now we turn the focus towards adaptation, starting with how to induce amino acid starvation and following with its effect.

1.1.2.1 Methods of inducing amino acid starvation

*E. coli* growing in their natural environment will possibly never experience the joy of steady-state growth but can undergo short bursts of fast growth followed by periods of starvation (Kolter et al., 1993). Inducing amino starvation is used as a controlled method to investigate cellular adaptation towards starvation in general.

Amino acid starvation might not only be a laboratory artifact since recently, Germain et al. discovered that HipA, a protein involved in persistence, modifies glutamyl-tRNA synthetase and thereby inhibits the catalytic formation cognate aminoacyl tRNA (charged tRNA). This process leads to amino acid starvation (Germain et al., 2013; Maisonneuve and Gerdes, 2014). The exogenous induction of amino acid starvation can be induced in at least three ways:

1. *Excess valine*

In *E. coli* K-12, the biosynthesis pathway of valine and isoleucine share some of the same enzymes. (See Fig. 2A.) Three of the enzymes are the heterodimeric acetohydroxy acid synthases: AHAS I, AHAS II, and AHAS III, which are encoded by the genes *ilvBN*, *ilvGM*, and *ilvIH* respectively. AHAS I and AHAS III are feedback inhibited by valine but not AHAS II. In *E. coli*-K12, AHASII contains a frameshift mutation in the gene coding *ilvG*, corrupting the function of AHASII. (See Fig.2B.) Therefore, it is possible to inhibit the formation of isoleucine by adding an excess of valine (Leavitt and Umbarger, 1962; Valle et al., 2008).
Figure 2. A: Biosynthesis pathway of valine and isoleucine. The heterodimeric acetohydroxy acid synthases: AHAS I, AHAS II, and AHAS III, are encoded by the genes ilvBN, ilvGM, and ilvIH respectively. At high levels of valine, AHAS I and AHAS III are inhibited but not AHAS II. (See dashed lines.) The activity of the enzyme product of ilvA is allosterically upregulated in the presence of valine and downregulated in the presence of isoleucine. B: E. coli K-12 the frameshift mutation in ilvG: The gene product of ilvG is inactivated by a frameshift mutation in MG1655. The above figure displays the MG1655 ilvG+ mutant aligned with MG1655. The figure is adapted from (Valle et al., 2008).

It has been reported that valine-induced isoleucine starvation possibly accumulates α-ketobutyrate (Herring et al., 1995). This leads to the inhibition of the glucose carbohydrate phosphotransferase system (Deutscher et al., 2006). Moreover, valine allosterically activates the enzyme product of IlvA (Chen et al., 2013). (See dashed lines, Fig. 2A). Therefore, this method might have unexpected side effects.

The amino acid starvation induced by this method is ‘soft’ since the feedback mechanism is a little leaky (Herring et al., 1995; Leavitt and Umbarger, 1962). In addition, the ‘old’ isoleucine that exists from before the excess valine was added is still available when the ‘old’ proteins are degraded.

2. Serine starvation

Serine hydroxamate disables the transfer of serine to its conjugated tRNA as performed by the seryl-tRNA synthetase. As such, the seryl-tRNA synthetase has a higher affinity towards serine
hydroxamate than serine, so serine hydroxamate probably inhibits acetylation of serine tRNA by excluding the amino acid substrate from the enzyme. Consequently, addition of serine hydroxamate reduces seryl-tRNA synthetase activity and inhibits protein synthesis and growth (Tosa and Pizer, 1971). Inducing serine starvation by adding Serine hydroxamate is causing ‘severe’ or complete starvation since serine tRNA remains deacylated (uncharged).

3. Medium change

If the cells are auxotroph for a specific amino acid, they will only enter steady-state growth if the medium is supplemented with this particular amino acid. By first including the amino acid into the medium during steady-state growth and then changing the medium to one lacking the amino acid, starvation is triggered (Kolter et al., 1993). This method can also have side effects, if the strain is auxotroph (e. g., threonine), removal of exogenous threonine will mostly affect the level of isoleucine. This is because threonine is converted to isoleucine by the enzyme product of ilvA. (See dashed lines, Fig. 2A.) Therefore, lowering the concentration of threonine n also reduces the concentration of isoleucine (Yegian and Stent, 1969). This process illustrates how the biosynthesis pathway for amino acids is complex (Umbarger, 1969; Vitreschak et al., 2006). Therefore, disturbing the synthesis of one amino acid can affect others. This method induces ‘medium’ amino acid starvation because ‘old’ amino acids from degraded proteins can be recycled.

1.1.2.2 Triggers of (p)ppGpp induction

During protein synthesis, the code transiently stored in the mRNA is translated by matching the codon of the mRNA and anti-codon of the charged tRNA subsequently leading to elongation of the growing peptidyl chain. This process is catalyzed by the huge ribo-enzyme the ribosome (Blanchard et al., 2004; Moazed and Noller, 1989). The charged tRNA entering the acceptor site (A site) of the ribosome is associated with elongation factor Tu (EF-Tu) and guanosine-5'-triphosphate (GTP), the so called tertiary complex (Rodenina et al., 2005a). During amino acid starvation, the chance for a near-cognate or an uncharged tRNA entering the A site is much higher because of the sudden lowering of cognate-charged tRNA. The entering of an uncharged tRNA into the A site of the ribosome stalls the ribosome and leads to an increase in the cellular concentration of guanosine pentaphosphate and guanosine tetraphosphate, together abbreviated as (p)ppGpp. This increase in (p)ppGpp starts the cellular adaptation ‘program’ called the stringent response (Goldman and Jakubowski, 1990; Haseltine and Block,
Neidhardt, 1966). Not only amino acid starvation but also fatty acid inhibition, oxidative stress, and carbon; nitrogen; and phosphorus starvation are conditions where (p)ppGpp accumulation is triggered (Battesti and Bouveret, 2006; Bougdour and Gottesman, 2007; Brown et al., 2014; Gong et al., 2002; Seyfzadeh et al., 1993; Xiao et al., 1991). Therefore the level of (p)ppGpp is linked to the availability of nutrients, growth (Potrykus et al., 2011) and stress response (Potrykus and Cashel, 2008).

### 1.1.2.3 RelA and SpoT catalyzes the synthesis of (p)ppGpp

The enzyme catalyzing (p)ppGpp accumulation during amino acid starvation is RelA. RelA is also responsible for the accumulation of (p)ppGpp during Nitrogen starvation (Brown et al., 2014). Fatty acid inhibition, carbon starvation, and oxidative stress also lead to the accumulation of (p)ppGpp, but the enzyme catalyzing the (p)ppGpp accumulation is SpoT (Battesti and Bouveret, 2006; Gong et al., 2002; Seyfzadeh et al., 1993). RelA and SpoT catalyze the synthesis of (p)ppGpp by condensation according to the equation below (Starosta et al., 2014):

\[
\text{GTP/GDP} + \text{ATP} \rightarrow (\text{p})\text{ppGpp} + \text{AMP} + \text{H}_2\text{O}
\]

But only SpoT are able to catalyze the hydrolysis of (p)ppGpp, leading to its degradation. Thus, RelA and SpoT in *E.coli* have homologous domains containing the synthase function, but only SpoT contains a functional hydrolase domain. The synthase domains of RelA/SpoT and the hydrolase domain of SpoT belong to the family of RelA-SpoT-Homologs (RSH). In many gamma and beta proteobacteria, the function of RSH is maintained by two enzymes whereas there is only one group of RSH in most other eubacteria (Atkinson et al., 2011; Mittenhuber, 2001a; Potrykus and Cashel, 2008).

### 1.1.2.4 How SpoT and RelA is activated

How SpoT and RelA are activated is still debated and not clear. The dual function of SpoT has not been well explored, but a minor domain upstream the conserved C-terminal is involved in switching between the synthetase and the hydrolase function. During fatty acid starvation, the acyl carrier protein, an important component in the synthesis of fatty acid, binds the minor domain. This action probably leads to the activation of synthetase function of SpoT (Battesti and Bouveret, 2006).

The activation of RelA during amino acid starvation has been better studied, and several models exist (English et al., 2011; Richter, 1976; Wendrich et al., 2002). One model ‘the
extended hopping model’, (English et al., 2011) tries to connect all of the facts found. These facts are listed below:

- Point mutations in the conserved C-terminal domain of RelA, abolish activation under Serine hydroxamate starvation (Gropp et al., 2001); this is also the case if L11, a ribosomal protein located at the 50S subunit, is mutated (Yang and Ishiguro, 2001).
- An electron microscope image of the complex, ribosome, uncharged tRNA, and RelA, shows direct interaction between RelA, L11, and the uncharged tRNA (Agirrezabala et al., 2013b).
- In vitro experiments substantiated that RelA binding to the ribosome primarily depends on mRNA presence at the ribosome; RelA synthesis of (p)ppGpp depends on L11. During the synthesis of (p)ppGpp, RelA lowers the affinity towards the ribosomes; the affinity between ribosomes and the uncharged tRNA is not changed when (p)ppGpp is expressed (Wendrich et al., 2002)
- The presence of a 3’OH group at the terminal adenosine at tRNA from yeast showed itself to be important for the induction of (p)ppGpp synthesis in an ex-vivo ribosome-stringent-factors-system purified from E.coli (Sprinzl and Richter, 1976)
- RelA synthesis of (p)ppGpp is enhanced by a positive allosteric feedback mechanism, but high levels of (p)ppGpp are not sufficient for the full activation of RelA (Shyp et al., 2012).
- During starvation, RelA spends most of the time dissociated from the ribosome, and before starvation, RelA is bound to the ribosome (English et al., 2011)
- The numbers of RelA per Ribosomes is 1: 200 (Pedersen and Kjeldgaard, 1977)
- The maximum synthesis rate of (p)ppGpp is measured to ~ 50,000 (p)ppGpp/cell/second (Lund and Kjeldgaard, 1972)

The ‘Hopping-model’ combines the results and explains the activation of RelA as such: RelA is bound to the ribosome during translation; when the ribosomal A-site is occupied by an uncharged tRNA, the ribosome/tRNA complex is subject to structural changes, activating RelA by direct interaction between the ribosome/tRNA and RelA. RelA starts catalyzing the synthesis of (p)ppGpp, leading to a decrease in the affinity between the RelA and the ribosome. This process leads to the dissociation of RelA, but the synthesizing of (p)ppGpp continues for several rounds. The higher concentration of (p)ppGpp upregulates the
expression of genes responsible for the biosynthesis of the starved amino acid, restoring the level of the starved amino acid (English et al., 2011).

Because this model proposes the continuous synthesis of (p)ppGpp after RelA has dissociated from the ribosome, it explains how relatively few RelA enzymes per cell (in hundreds) are able to produce (p)ppGpp at a synthesis rate of ~ 50,000 (p)ppGpp per cell per second (Lund and Kjeldgaard, 1972). We will note that care should be taken not to over interpret the exact rate of (p)ppGpp per cell per second since an estimation of the cellular (p)ppGpp level depends on these factors: the severity and type of starvation, the genotype of the cell, the specific amino acid starved for (Payoe and Fahlman, 2011), and the method used when extracting the (p)ppGpp ((Buckstein et al., 2008; Lagosky and Chang, 1978; Lagosky and Chang, 1980).

1.1.2.5 The effect of (p)ppGpp

Globally, the effect of amino acid starvation is the reduction in protein and RNA synthesis, thus reducing growth (O’Farrell, 1978; Parker and Friesen, 1980). The severity of the reduction depends on which amino acid starved for (Parker and Friesen, 1980) and how the amino acid starvation is induced. (See section 1.1.2.1.)

As discussed in section 1.1.1.3, there is a connection between the growth rate during steady state and the macromolecular composition of DNA, RNA, and proteins. In cells unable to produce (p)ppGpp, there is an elevated level of RNA in balanced, slowly growing cells. This indicates that cells incapable of producing (p)ppGpp are not able to adjust the macromolecular level to the nutrient condition of their surrounding (Potrykus et al., 2011). Thus, not surprisingly, the induction of (p)ppGpp directly affects transcription, and for some operons, the effect is only achieved when (p)ppGpp act in concert with the protein DksA (Gopalkrishnan et al., 2014; Gummesson et al., 2013; Paul et al., 2005).

Exactly how (p)ppGpp (and DksA) regulate transcription is still debated, and one model, the ‘saturating model’, predicts that an altered level of free RNA polymerase (RNAP) affects each promoter differently based on the saturation properties of the specific promotor (Jensen and Pedersen, 1990). Sørensen et al. measured the RNA chain growth rate and found the RNA chain growth rate to be inversely correlated with the concentrations of (p)ppGpp. This inverse relation was seen during amino acid starvation and under non-starving conditions (Sørensen et al., 1994; Vogel et al., 1992). The ‘saturating model’ predicts that an increase in (p)ppGpp
decreases the chain growth rate and occupies a larger fraction of RNAP in the transcription phase, thus lowering the amount of free RNAP. The model subdivides transcription initiation into two steps: formation of the open complex (RNAP/opened-DNA) and the rate of promoter clearance. Stringent controlled promoters are promoters considered difficult to saturate since both the rate of the open complex formation and the rate of promoter clearance are very high. Meanwhile, non stringent promoters are easy to saturate since formation of the open complex and/or the rate of promoter clearance are low, making the promoter ineffective. The reduction of free RNAP affects promoter activity of the stringently controlled promoters the most, redirecting transcription from genes encoding rRNA, tRNA, and ribosomal proteins (Murray et al., 2003) to genes encoding proteins used when maintaining the cell in a state of stress.

Interestingly, expression of the ribosomal modulation factor (RMF) requires the presence of (p)ppGpp, and an increase in RMF leads to the accumulation of the 100S particles, a complex consisting of two RMF bound to two ribosomes (Izutsu et al., 2001). Ribosomes bound in the 100S particles are transcriptionally inactive but are still ready for transcription within minutes if conditions change to being favorable for growth (Wada et al., 1995). However, if the gene coding RMF are disrupted, cell viability in the stationary phase is shortened (Yamagishi et al., 1993).

1.1.3 tRNA

1.1.3.1 Canonical function of tRNA

The molecules of tRNA are functionally highly conserved in all domains of life and are fundamental components of the translational machinery (Giegé R, 2013; Quax et al., 2015). Transfer RNA mediates protein synthesis by translating the code transiently stored in the mRNA into the growing polypeptidyl chain. The decoding is accomplished by matching the codon of the mRNA to the anti-codon of the tRNA; thus, the sequence in the decoded mRNA is reflected in the final protein (Ling et al., 2009). This important and highly regulated process takes place in the ribosome, keeping the error rate as low as possible since errors can lead to non-functional and incorrectly folded proteins with detrimental physiological consequences (Zaher and Green, 2009). Even though reports have stated that errors of translation can be beneficial to cells, under sudden external pressures (Reynolds et al., 2010). Therefore, a balance between the need for accuracy and the benefits of adaptation must have evolved (Zaher and Green, 2009).
During elongation of the polypeptide chain, the average error of missense translation is $10^{-4}$-$10^{-3}$ per codon. As such, missense error is understood as incorporation of an amino acid different than as dictated by the codon of the mRNA (neglecting error in replication and transcription) (Bouadloun et al., 1983; Edelmann and Gallant, 1977; Parker, 1989). The processes having impact on the missense error can be subdivided into two overall steps: fidelity in the process of decoding the mRNA, and fidelity in the process of charging the correct amino acid to the cognate tRNA.

The process of decoding has an error frequency of $10^{-3}$-$10^{-4}$ per translated codon (Bouadloun et al., 1983; Edelmann and Gallant, 1977; Parker, 1989), and the process of charging has an error frequency of $10^{-1}$-$10^{-5}$ per charging event (Edelmann and Gallant, 1977). Thus, decoding is the process dominating the average error frequency.

Models are developed in trying to capture the error-dominating processes in the decoding of mRNA. The ‘kinetic model’ describes the selection between cognate or near cognate tRNA in terms of second- and first-order reaction constants. (See Fig. 3 and Fig. 3 legend.) (Gromadski et al., 2006; Gromadski and Rodnina, 2004; Rodnina et al., 2005b). According to the ‘kinetic model’, the fidelity in the decoding process can be subdivided into two overall steps: Initial selection and proofreading. The two steps are separated by the non-reversible hydrolysis of GTP as determined by the rate constant $k_{GTP}$. The initial selection includes initial binding ($k_1, k_{-1}$) codon recognition ($k_2, k_{-2}$), and GTPase activation ($k_3$). Meanwhile, the proofreading includes the dissociation of EF-Tu ($k_6$) simultaneously taking place with tRNA accommodation ($k_5$) and peptide bond formation ($k_{pep}$) or rejection of tRNA ($k_7$).
**Figure 3. An overview of the kinetic model.** In the first step, the ternary complex transiently binds the ribosome ($k_1$ for forward rate constant, $k_{-1}$ for backward), followed by codon recognition ($k_2$ for forward rate constant, $k_{-2}$ for backward). Codon recognition activates the GTPase at EF-Tu ($k_3$). The rate of the activation limits the downstream processes; therefore, this step is rate-limiting for the hydrolysis of GTP ($k_{GTP}$). The release of inorganic phosphate induces a conformational change in EF-Tu ($k_4$), reducing affinity towards the charged tRNA and EF-Tu dissociates, (this step is not shown for sake of simplicity ($k_d$)). The charged tRNA accommodates into the A site ($k_5$), and either rapid peptide bonds are formed ($k_{pep}$), or the charged tRNA is rejected from the ribosome ($k_7$). All rates constant are first order except for $k_d$, which is second order and dependent on the concentration of the ribosome. See lower area below the figure for the meaning of symbols. The figure is adapted from (Rodnina et al., 2005b).

On the basis of a large body of experiments Gromadski and Rodnina have concluded the following: The initial selection for cognate or near cognate tRNA cannot be explained by selectively stabilization of cognate but not near-cognate, codon-anticodon complexes mediated by the ribosome. Rather, the selectivity in the initial selection is a non-equilibrium process driven by the rapid irreversible forward reactions of GTP hydrolysis. This leads to a selectivity of 1 wrongly incorporated amino acid for every 30 correctly incorporated amino acids in the final protein. During the process of proofreading, the selection is obtained by selective stabilization and a much higher accommodation rate of cognate tRNA than of near cognate tRNA. The selectivity during proofreading is 1 wrongly incorporated amino acid for every 15 correctly incorporated amino acids in the final protein.

Since the overall fidelity can be calculated as the product of the two steps, the setup of Gromadski and Rodnina could account for an error frequency of $2.2 \cdot 10^{-3}$ per translated codon (Gromadski et al., 2006; Gromadski and Rodnina, 2004; Rodnina et al., 2005b). The ‘kinetic model’ could therefore almost fully account for the error frequency of $10^{-7}$-$10^{-6}$ per translated codon. But the kinetic measurements were done in vitro where the concentration of the tested cognate or near cognate tRNA was mixed 1:1, or the kinetic constants were determined by assays using only one type of tRNA at a time. Besides, the concentration of ribosome was varied while the amount of tRNA was held constant thus far from the in vivo situation in *E.coli* (Jakubowski and Goldman, 1984).

Consequently, varying the level between the cognate and near-cognate tRNAs could have an impact on the fidelity in vivo. This was confirmed by the experiments of Kramer et al. when they measured the translational fidelity in vivo. They found the fidelity to be highly dependent on the competition between the cognate and the near cognate tRNA (Kramer and Farabaugh, 2007).
The average charging level of tRNA in exponential growing cells is ~80% (Lewis and Ames, 1972) reflecting a balance between the biosynthesis of amino acids and tRNA, the rate for charging the tRNA, and the successive addition of polypeptides. If competition between charged cognate and near cognate tRNA is having an effect on translation fidelity, as shown by Kramer et. al., the error rate should increase if one amino acid suddenly goes missing. This was confirmed by the experiments of Parker et al. and Sørensen et al. when they found an increase in error frequency after the induction of amino acid starvation. The error frequency was largest shortly after induction of starvation than later on (Parker and Friesen, 1980; Sørensen et al., 1994). If the cells are unable to express (p)ppGpp, a further increase in the error frequency will be seen (O’Farrell, 1978; Parker and Friesen, 1980). This process links the error frequency after amino acid starvation to the ability of the cell to induce (p)ppGpp. As described in section 1.1.2.5, cells unable to induce (p)ppGpp lack control of the RNA synthesis. But how does this affect the error frequency during translation?

M. A. Sørensen measured the tRNA charging level in *E. coli* and in an isogenic *E. coli* unable to induce (p)ppGpp during amino acid starvation. The level of charged tRNA was recorded after the induction of arginine starvation, and the experiment was repeated for 3 other amino acids with other identities. The tRNA from the *E. coli* unable to induce (p)ppGpp had 5-10 times lower the charging level than the cells capable of inducing (p)ppGpp.

M. A. Sørensen developed a model explaining how the charging level of tRNA affects the error of translation: Induction of (p)ppGpp reduces the rate of RNA synthesis, and this leads to a reduction in the mRNA pool (Section 1.1.2.5). Lowering in the mRNA pool reduces the demand for charged tRNA, making the pool of mRNA the limiting factor. In this way, the competition between cognate and near cognate tRNA is reduced and the error frequency in the growing polypeptide chain lowered.

For the *E. coli* incapable of inducing (p)ppGpp, the mRNA pool is still high after the induction of amino acid starvation, increasing the demand for charged tRNA. In this situation, the limiting factor is the low amount of charged cognate tRNA, which increases the competition from near cognate tRNAs. Hence, the stronger competition between near and cognate tRNA increases the change for incorporation of a near cognate tRNA, instead of cognate tRNA, and thereby, the error frequency in the growing polypeptide chain is enlarged (Sorensen, 2001).
1.1.4 RNases

The components rRNA and tRNA are considered stable housekeeping RNA, but treating the cell with any components that are perturbing membrane stability such as streptomycin (Dubin, 1964), mitomycin C (Suzuki and Kilgore, 1967), or heat induces rRNA degradation (Deutscher, 2009). The agent responsible is probably the RNase I since destabilization of the inner membrane allows for the entry of RNase I from the periplasmic space to the cytoplasm (Deutscher, 2003; Maiväli et al., 2013). Likewise does the stability of rRNA decrease during conditions of slow growth, (See section 1.1.2.5.), thus limitation of nitrogen (Ben-Hamida and Schlessinger, 1966), phosphorus (Maruyama and Mizuno, 1970), carbon (Jacobson and Gillespie, 1968b), and even mg$^{2+}$ trigger rRNA degradation (McCarthy, 1962). Degradation of the rRNA enables the cell to reuse the ribosomal components. Thus the ribosomes can be seen as a storehouse for nutrients, but degradation of rRNA also happens as a result of quality control. The degradation of rRNA due to quality control or starvation share common characteristics: the first cleavage is induced by endonucleolytic cleavages, and this cleavage is followed by exoribonucleolytic degradation.
Table 1 Overview of selected RNases from *E.coli*, listing their action, location, and function. The table is based on information from (Deutscher et al., 2006) except where marked with a number in parenthesis: 1 (Basturea et al., 2011), 2 (Beppu and Arima, 1969; Maiväli et al., 2013), 3 (Gopalan et al., 2002), 4 (Xiao et al., 2009), 5 (Spickler and Mackie, 2000), 6 (Li and Deutscher, 2002), 7 (Dutta et al., 2012), 8 (Li and Deutscher, 1995), 9 (Awano et al., 2010), 10 (Deana and Belasco, 2004), and 11 (Deutscher, 2009)

<table>
<thead>
<tr>
<th>Action</th>
<th>Location</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNase I</td>
<td>Nonspecific endoribonuclease digest mainly single-stranded RNA.</td>
<td>Mainly periplasmic space (11)</td>
</tr>
<tr>
<td>RNase BN</td>
<td>Endo and exoribonuclease (7)</td>
<td>Cytosol</td>
</tr>
<tr>
<td>RNase P</td>
<td>Initial cleavages in degradation and maturation. Initial cleavage is endonucleolytic cleavage, dependent on divalent metal ions (3)</td>
<td>Cytosol</td>
</tr>
<tr>
<td>RNase III</td>
<td>Initial cleavages in degradation and maturation. Double strand specific endonuclease (4)</td>
<td>Cytosol</td>
</tr>
<tr>
<td>RNase II</td>
<td>Terminal degradation. Exoribonuclease (5)</td>
<td>Cytosol</td>
</tr>
<tr>
<td>RNase R</td>
<td>Terminal degradation. Exoribonuclease with helicase activity (9)</td>
<td>Cytosol</td>
</tr>
<tr>
<td>RNase PH</td>
<td>3’end maturation. Endoribonuclease</td>
<td>Cytosol</td>
</tr>
<tr>
<td>RNase E</td>
<td>Initial cleavages in degradation and maturation. Endonuclease (6)</td>
<td>Suggested to be associated in the degradosome</td>
</tr>
<tr>
<td>PNPase</td>
<td>Terminal degradation. Exoribonuclease (5).</td>
<td>Cytosol</td>
</tr>
<tr>
<td>RNase G</td>
<td>Initial cleavages in degradation and maturation. Endonuclease (10)</td>
<td>Cytosol</td>
</tr>
<tr>
<td>RNase T</td>
<td>3’end maturation Exoribonuclease (8)</td>
<td>Cytosol</td>
</tr>
</tbody>
</table>
The complete sets of RNases responsible for the degradation of rRNA during quality control or starvation-induced rRNA degradation are not fully resolved, but the Deutscher laboratories have identified a few. The RNases that involved rRNA degradation during 6 hours of carbon starvation were RNases II, RNase PH and R (where RNase PH seemed to be involved in accumulation of the larger fragment, and RNase R/ RNase II where involved in the degradation of the pre-fragmented rRNA). The RNases responsible for degradation during rRNA quality control were PNPase and RNase R, and it was not possible to locate the RNase responsible for the pre-fragmentation (Deutscher, 2009). In general, the enzymes responsible for the metabolism of mRNA and tRNA/rRNA appear to be identical (Table 1). Thus, a relevant question is this: Why are rRNA and tRNA much more stable than mRNA during rapid growth?

One factor believed to distinguish between the metabolism of mRNA and tRNA/rRNA was polyadenylation performed by the poly(A)polymerase PAP I and targeting mRNA for rapid turnover (Carpousis et al., 1999; O’Hara et al., 1995). Interestingly, Li et. al. showed that rRNAs are poly-adenylated as well. They therefore proposed that any naked 3’ hydroxyl terminus of RNA is subject to competition between PAP I and one or more exoribonuclease (Li et al., 2002a). Polyadenylation by PAP I will label the RNA for degradation whereas 3’ end maturation by the exoribonuclease covers the RNA to its mature form. The polyadenylation of rRNA also led them to propose that the difference between mRNA and tRNA/rRNA during exponential growth is the level of protein interaction; tRNA and rRNA interact intensively with the proteins of the translational machinery during fast growth, which is not the case for mRNA. On the other hand, during limited growth, the growth rate declines, leaving stable RNA vacant. Therefore, according to the proposal of Li et. al., it is also prone for degradation (Deutscher et al., 2006; Deutscher, 2003).

In addition to the RNases listed in table 1, E.coli also consists of toxin-antitoxin systems that are activated during starvation. Some of these systems also contain endoribonucleases (Christensen et al., 2003; Germain et al., 2015; Maisonneuve et al., 2013).
Paper 1
Title: Transfer RNA is rapidly degraded upon translational arrest in *Escherichia coli*.

Short title: Rapid degradation of tRNA upon translational arrest.

Authors: Sine Lo Svenningsen, Mette Kongstad, Thomas Søndergaard Stenum, Ana J. Muñoz-Gómez*, Michael A. Sørensen.

Department of Biology, University of Copenhagen, Ole Maaløes Vej 5, 2200 Copenhagen N, Denmark.

* Present address: Ajmunoz.gomez@gmail.com

Corresponding author: Associate Professor Michael A. Sørensen, Phone: +45 35323711. Email address: MAS@bio.ku.dk.

Keywords (minimum 3): transfer RNA; *Escherichia coli*; stable RNA degradation; bacterial stress response; amino acid starvation; stringent response;
Abstract

Transfer RNA is an essential component of the protein synthesis machinery. Due to its long half-life compared to messenger RNA, bacterial tRNA has become known as stable RNA. Here, we show that tRNAs become highly unstable as part of *Escherichia coli*’s response to amino acid starvation. Degradation of the majority of the bacterium’s tRNA pool occurs within twenty minutes of the onset of starvation for several individual amino acids, affects both the non-cognate and the cognate tRNA for the amino acid that the cell is starving for, and includes both the charged and uncharged tRNA species. The alarmone ppGpp orchestrates the stringent response to amino acid starvation, and is produced in response to the presence of uncharged tRNA at the ribosomal A site. However, we find that tRNA degradation does not occur as part of the stringent response, as it occurs with similar kinetics in a relaxed mutant. Further, we also observed rapid tRNA degradation in response to another stress, namely treatment with the transcription initiation inhibitor rifampicin, which does not induce the stringent response. We propose a unifying “passive” model for these observations, in which the surplus tRNA is degraded whenever the demand for tRNA in protein synthesis is reduced. Thus, the cellular tRNA pool is a highly regulated entity, which is regulated both at the levels of tRNA synthesis and degradation. We propose that degradation of surplus tRNA is a means to prevent mistranslation in the stressed cell, because it reduces the competition between cognate and near-cognate tRNAs for pairing with the mRNA at the ribosome.

Introduction

Transfer RNAs (tRNA), the adaptor molecules that translate the codons in mRNA to amino acids in proteins, are a fundamental component of the translation machinery and is typically referred to as abundant “house-keeping” RNA. Accordingly, our knowledge about the regulation of tRNA levels in *Escherichia coli* is by and large limited to the growth-rate control of RNA polymerase partitioning, which ensures that the RNAs of the translation machinery are produced in amounts that match the cell’s need for protein synthesis (Hernandez and Bremer, 1993; Maaløe and Kjeldgaard, 1966a; Ryals et al., 1982a). In agreement with the house-keeping role, bacterial tRNA is considered highly stable RNA, which is generally only degraded when compromised in quality (Li et al., 1998; Li et al., 2002b; Mohanty et al., 2012; Neidhardt, 1964).

In eukaryotic cells, however, a different picture is emerging, namely that the cellular tRNA pool is a highly dynamic and tightly regulated entity (recently reviewed in (Kirchner and
Remarkably, measurements of the tRNA pool of hundreds of human cell samples showed that the tRNA composition of proliferating cells differs greatly from that of differentiated cells, and that the cellular tRNA repertoire matches the codon usage signature of proliferation-related or differentiation-related mRNAs, respectively (Gingold et al., 2014). In addition to global changes in the tRNA pool, it also appears that individual tRNAs may be either stabilized, destabilized, or functionally altered depending on the cellular state. In *Saccharomyces cerevisiae*, for example, dynamic methylation of the uridine wobble base of tRNA^ARG(UCU)_ and tRNA^GLU(UUC)_ changes the specificity of these tRNA species to promote translation of mRNAs involved in the DNA damage response (Begley et al., 2007), and different cytotoxic chemicals have been shown to induce specific changes in tRNA modifications, which in several cases appear to be an important part of the cellular stress response, since mutants that lack the corresponding modification enzymes show increased sensitivity to the toxicant (Chan et al., 2010). Together, these and other studies support a general model of dynamic control of both the production, modification, and degradation of tRNAs, which serves to align the supply of aminoacylated tRNAs with the changing demands posed by the transcriptome characteristic of different cellular states or growth conditions.

Imbalance between the supply and demand for particular aminoacylated tRNAs is not only energetically suboptimal, it also results in increased mistranslation, which can be detrimental for the cell. Specifically, starvation for a single amino acid results in reduced levels of the corresponding aminoacylated tRNA, which in turn results in increased incorporation of incorrect amino acids supplied by near-cognate tRNAs at the corresponding “hungry” codons on the mRNA (Parker et al., 1980; Sorensen, 2001). In *E. coli*, this increased error rate upon amino acid starvation is ten-fold less dramatic in wildtype cells than in mutants defective for the stringent response (Hall and Gallant, 1972; O’Farrell, 1978). The stringent response describes the orchestration of large-scale changes in cellular gene expression upon decreases in nutrient availability. The changes in transcription upon amino acid starvation result from direct binding of the alarmone (p)ppGpp to RNA polymerase. The RSH (Rel/SpoT homolog) proteins that synthesize (p)ppGpp appear to be almost universally conserved in bacteria (Atkinson et al., 2011; Mittenhuber, 2001b), although great diversity has been reported with regards to both the signals that activate (p)ppGpp synthesis and the means by which (p)ppGpp exerts its effects in the cell (reviewed in (Boutte and Crosson, 2013)). The major transcriptional effects of the stringent response are decreased expression of genes encoding the translational machinery (including tRNA and
rRNA), as well as a general decrease in mRNA synthesis, except for mRNAs encoding some amino acid biosynthetic enzymes and stress response factors (reviewed in (Magnusson et al., 2005; Potrykus and Cashel, 2008)). The reduced mRNA synthesis rate makes mRNA the limiting component for translation during the stringent response (Sorensen et al., 1994), and has been found to be essential for alleviating misreading at hungry codons, presumably because it reduces the demand for the cognate aminoacylated tRNA, thereby allowing it to remain at a level where it can outcompete near-cognate tRNAs at the hungry codons (Sorensen, 2001; Sorensen et al., 1994). Another means to decrease competition between cognate and near-cognate charged tRNAs at the hungry codons would be to reduce the pools of near-cognate tRNAs. The reduced transcription of tRNA genes imposed by the stringent response upon starvation may to some degree accomplish this goal, but if the existing tRNA pool is long-lived it is unlikely that termination of tRNA synthesis upon amino acid starvation has noteworthy effects on the pools of near-cognate aminoacylated tRNAs in the short term. By contrast, the observed termination of tRNA synthesis during the stringent response can be understood as a means to reduce energetically wasteful synthesis of translation substrates at a time where resources are scarce and the need for de novo protein synthesis limited, rather than a means to reduce mistranslation (Ehrenberg and Kurland, 1984). Indeed, termination of rRNA and tRNA synthesis, which accounts for 85% and 14% of cellular RNA in rapidly growing cells, respectively (Bremer and Dennis, 2008), strongly reduces the cells’ energy requirement. In this work, we investigated the hypothesis that tRNA half-life may be actively decreased under conditions such as amino acid starvation where a reduction of near-cognate tRNA pools would be expected to improve translation accuracy.

Interestingly, it has been reported that different kinds of long-term nutrient limitation eventually results in degradation of the otherwise highly stable rRNA (Jacobson and Gillespie, 1968a; Julien et al., 1967; Maruyama and Mizuno, 1970; McCarthy, 1962; Norris and Koch, 1972; Zundel et al., 2009b). Degradation of rRNA could serve the purpose of recycling their nucleotides as nutrients for the starved cell, and the degradation appears to aid recovery from long-term starvation (Kaplan and Apirion, 1975). On the shorter time-scale of a few hours, Mandelstam & Halvorson showed in 1959 that 24% of ribosome-associated RNA is degraded after four hours of threonine starvation (Mandelstam and Rogers, 1959) and data from the Deutscher laboratory showed degradation of an appreciable amount of rRNA after 3 hours of carbon starvation (Zundel et al., 2009b). This body of experimental results shows that different types of starvation eventually result in degradation of rRNA, and possibly all stable RNA. However, to our knowledge, no degradation mechanism has been identified, and no one
has investigated the short-term kinetics of stable RNA decay. In this work, we show that more than half of the cellular tRNA content is degraded within twenty minutes of a treatment that leads to translational inhibition, such as starvation for an amino acid or deprivation of the mRNA substrate by rifampicin-mediated inhibition of transcription.

**Results**

*The majority of cellular tRNA is rapidly degraded upon amino acid starvation*

To determine tRNA half-life upon amino acid starvation, the histidine auxotroph *E. coli* strain NF915 was grown in minimal medium containing 5 μg/ml histidine, and starved for histidine by filtration and resuspension in the same medium lacking histidine, as described previously (Sorensen, 2001). Aliquots of the cultures were harvested for tRNA preparation immediately before and at 5-80 minutes after starvation, and the levels of selected tRNAs were examined by Northern blot (Figure 1A). Since one of the major effects of amino acid starvation is change in the rates of synthesis, and perhaps degradation, of mRNA, tRNA, and rRNA, there are no cellular RNA species that could serve as an appropriate loading control in these experiments. Similarly, simply normalizing to the total RNA harvested from each sample is ambiguous, given that we are precisely performing the experiment to measure the changes in RNA content per cell upon amino acid starvation. To circumvent these issues, the cell culture aliquots were instead spiked with equal volumes of a culture of *Sulfolobus solfataricus* prior to RNA preparation, and the Northern blots were probed for *S. solfataricus* 5S RNA with a probe that does not cross-hybridize with any *E. coli* RNA. Such an experiment is shown in Figure 1A, and the results are quantified in Figure 1C.
Figure 1: Both cognate and non-cognate tRNA levels decrease rapidly upon histidine starvation. A)
Phosphorimager scan of a Northern blot membrane containing RNA harvested from NF915 (Rel') and probed for
the indicated tRNA species from *E. coli* (top) or *S. solfataricus* 5S RNA (bottom). Starvation was initiated at time
zero and the samples were harvested at the time indicated above each lane (minutes). There are two lanes for each
timepoint. The first lane contains the untreated sample, whereas the second lane contains chemically deacylated
tRNA from the same sample. The same membrane was successively stripped and reprobed for the tRNAs
indicated on the left. The boxed area in the upper blot indicates the part of the blot, which is shown for the
additional tRNAs. B) Same as A), but RNA was harvested from the otherwise isogenic relA2 mutant NF916. C+D)
The intensity of the bands was quantified using ImageJ software, and total tRNA (independent of charging state)
was normalized to *S. solfataricus* 5S RNA from the same sample. Lanes marked Ec and Ss are control lanes
containing RNA from *E. coli* or *S. solfataricus*, respectively. Panel C contains data for strain NF915 and panel D
shows data for strain NF916. The growth medium was MOPS 0.4% Glycerol 50μg/ml Thr, Leu, Arg and 5μg/ml
His and the generation time prior to amino acid starvation was 80 min for both strains.
Figure 1 shows that the majority of cellular tRNA is rapidly degraded upon starvation for histidine. Importantly, both the cognate tRNA^{HisR} (orange squares), as well as tRNA\textsuperscript{argVYZQ}, tRNA\textsuperscript{thrV}, tRNA\textsuperscript{leuZ}, tRNA\textsuperscript{leuPQVT}, and tRNA\textsuperscript{GluTUWW}, decrease in levels within five minutes after the initiation of histidine starvation, and reach a new steady state level of 20-40% of their level in unstarved cells within approximately 40 minutes of starvation. Aminoacylated tRNA can be distinguished from uncharged tRNA on the Northern blot due to the difference in size of the two species. To illustrate this difference, a chemically deacylated aliquot of each sample was loaded on the gel in the lane next to the untreated aliquot from the same sample (Figure 1A). We found that degradation of tRNA occurs independently of the charging status of the tRNA, that is, both charged and uncharged tRNA is rapidly degraded upon amino acid starvation (Figure 1A). Notably, no particular tRNA degradation products (fragments of tRNAs) were detected in any of the Northern blot experiments.

To examine whether this response was specific to histidine starvation, we next tested the stability of a battery of tRNAs upon starvation for leucine, by filtration and resuspension of a culture of NF915 grown with 50 μg/ml leucine into medium lacking leucine. As shown in Figure 2A, the levels of tRNA\textsuperscript{argVYZQ}, tRNA\textsuperscript{leuZ}, tRNA\textsuperscript{leuX}, tRNA\textsuperscript{leuW}, tRNA\textsuperscript{leuU}, tRNA\textsuperscript{leuPQVT}, tRNA\textsuperscript{HisR}, and tRNA\textsuperscript{ThrV} decrease rapidly upon removal of leucine. The leucine starvation procedure completely blocked growth as shown Supplementary Figure S2A. We made sure that the observed loss of tRNA upon starvation was not due to an unintended consequence of the starvation procedure, such as cell lysis, by confirming by colony counts that the number of live cells remained constant for the duration of the experiment (Supplementary Figure S2).
To ensure that the results were independent of the choice of strain background, we also carried out a similar set of experiments with *E. coli* MG1655 under valine-induced starvation for isoleucine (Bonner, 1946; Leavitt and Umbarger, 1961). As in the previous experiments, rapid degradation of tRNA<sub>Arg</sub><sup>YIZQ</sup>, tRNA<sub>Glu</sub><sup>TUVW</sup> and tRNA<sub>His</sub><sup>R</sup> was observed upon starvation for isoleucine (Supplementary Figure S1A, closed symbols). tRNA degradation is specific to the starvation condition, as all the tRNAs accumulate again upon addition of isoleucine to the cultures after 80 minutes of starvation (Supplementary Figure S1A, closed symbols). Finally, we have performed Illumina Deep Sequencing of the transcriptome of *E. coli* MG1655 before

**Figure 2:** Both cognate and non-cognate tRNA levels also decrease rapidly upon leucine starvation. A) Levels of different tRNA species in strain NF915 immediately before and 5, 30 and 60 minutes after starvation for leucine. Cells were grown in MOPS glycerol containing 50 µg/ml threonine, leucine, and arginine, and 5 µg/ml histidine with a generation time of 82 min prior to starvation. RNA was harvested at the indicated time points and subjected to Northern blot analysis, as in Figure 1. B) same as A), but RNA was harvested from the otherwise isogenic relA2 mutant NF916.
and ten minutes after starvation for isoleucine. This experiment was carried out for a different purpose and is not an optimal approach for quantification of tRNAs because modified bases in tRNAs are not faithfully copied by reverse transcriptase (Czech et al., 2010; Puri et al., 2014). However, from this data set, it is evident that all tRNA species with an expression level detectable above background decrease in abundance upon starvation (Supplementary Table S1).

Together, this body of data demonstrates that a hitherto undescribed part of the cellular response to the onset of amino acid starvation is a rapid degradation of the majority of cellular tRNA species, including both the non-cognate and the cognate tRNA for the amino acid that the cell is starving for, and including both aminoacylated and non-aminoacylated tRNA species.

**Rapid tRNA degradation is independent of the (p)ppGpp-mediated stringent response**

The cellular response to amino acid starvation is orchestrated by the stringent response, specifically by RelA-mediated (p)ppGpp production in response to the presence of a deacylated tRNA at the ribosomal A-site (Agirrezabala et al., 2013a; Haseltine and Block, 1973a). Therefore, we tested whether tRNA degradation is part of the stringent response triggered by ppGpp, by measuring tRNA levels in a mutant defective for relA. Figure 1B+D shows the results of histidine starvation experiments equivalent to those shown in Figure 1A+C, but carried out with strain NF916, an otherwise isogenic relA2 mutant of E. coli strain NF915. Similarly, Figure 2B shows a leucine starvation experiment of strain NF916 carried out alongside that of NF915 shown in Figure 2A. It is evident that tRNA degradation initially occurs with similar kinetics in the rel- and rel+ strains, hence activation of the stringent response by RelA cannot be the trigger for activation of tRNA degradation. Upon amino acid starvation, Rel- strains initiate the so-called relaxed response, which includes a drop in the (p)ppGpp pool and an increase in the transcription of rRNA and tRNA genes (Lagosky and Chang, 1981). Presumably, increased tRNA production in the Rel- strain accounts for the difference in tRNA pools between the two strains observed at later time points: All tRNA levels show an upwards trend in the Rel- strain at later time points after starvation, whereas they appear to reach a new steady state in the Rel+ strain (Figure 1). The altered transcriptional response of the starved Rel- strain may also explain why a single tRNA, tRNAleuU, remains present at high levels up to 30 minutes after leucine starvation in the relaxed mutant, but not in wildtype (Fig 2B). The leuU gene is atypically located downstream
of secG, encoding a component of the SecYEG translocase. We assume that transcription of the secG-leuU operon is increased after amino acid starvation in the relaxed mutant, which lacks the ability to down regulate σ70-directed transcription of highly expressed genes, due to reduced levels of (p)ppGpp. This hypothesis is supported by microarray data of secG mRNA, which was found to be down regulated nearly 6-fold in Rel+ cells upon outgrowth in isoleucine-limited medium, but unaltered in the relaxed mutant (Traxler et al., 2008b).

To further investigate the involvement of the (p)ppGpp in tRNA degradation, we also measured degradation of two tRNAs, tRNAArgVYZQ and tRNAThrV, upon starvation for arginine in a strain carrying the spoT1 allele (Laffler and Gallant), which has increased basal ppGpp levels due to reduced ppGpp-hydrolase activity of the bifunctional (p)ppGpp-hydrolase/synthase SpoT (Fiil et al., 1977) (Supplementary figure S3). The pattern of degradation of the tested tRNAs in the SpoT1 strain is similar to that in wildtype and relA2 backgrounds, supporting that tRNA degradation upon amino acid starvation is independent of (p)ppGpp. In many of our experiments, we have also measured the changes in tRNA charging levels during amino acid starvation (See for example Supplementary Figures S1 and S3). The charging levels of those tRNAs accepting other amino acids than the one starved for only increase slightly during amino acid starvation, and do not reach fully charged levels as it could have been anticipated if the tRNA pools remained unchanged (Elf et al., 2003). Our explanation is that earlier models don’t account for the degradation of excess tRNA molecules reported here. We return to this point in the discussion.

One potential consequence of amino acid starvation in E. coli is inactivation of the antitoxin part of toxin-antitoxin modules in a fraction of the cell population, resulting in active toxins, many of which encode RNA endonucleases (Christensen et al., 2003; Germain et al., 2015; Maisonneuve et al., 2013). Thus, we speculated that tRNA degradation could be a consequence of activation of an RNA endonuclease from a toxin-antitoxin module. However, tRNA degradation occurs according to the wildtype pattern in a strain that lacks the 10 type II toxin-antitoxin modules of E. coli (Maisonneuve et al., 2011), showing that tRNA degradation occurs independently of these 10 RNA endonucleases (Supplementary Figure S1, open symbols).

*Inhibition of mRNA synthesis also causes tRNA degradation*
Since tRNA degradation is not dependent on the stringent response, we speculated that the trigger of tRNA degradation might not specifically be amino acid starvation, but rather, tRNA degradation might be triggered by any condition that results in a decreased demand for the tRNAs in protein synthesis. Accumulation of unemployed tRNAs would directly result from the decreased protein synthesis rates of an amino-acid-starved cell, where ribosomes are expected to stall at the “hungry” codons, independent of the (p)ppGpp. If so, other treatments that result in disengaged tRNAs should also lead to rapid tRNA degradation. To test this hypothesis, we measured tRNA levels upon addition of rifampicin, which blocks transcription initiation by binding to RNA polymerase, and therefore quickly leads to reduced protein synthesis due to a lack of mRNA substrate (Hartmann et al., 1967; Reid and Speyer, 1970). Figure 3A shows that, indeed, blockage of transcription initiation with rifampicin results in rapid degradation of tRNA\textsuperscript{TyrTV}, tRNA\textsuperscript{GluTUW}, tRNA\textsuperscript{ArgYZQ} and tRNA\textsuperscript{LysQTVWYZ} in a manner very similar to that of amino acid starvation.

Figure 3: tRNA is rapidly degraded upon transcriptional arrest. Cellular levels of tRNAs were quantified by Northern blot as in Figure 1 and 2. A) 300 µg/ml rifampicin was added to exponentially growing cells at time zero to block transcription initiation. At time 80 minutes, the translation inhibitor chloramphenicol (100 µg/ml) was added to the culture. B) 100 µg/ml chloramphenicol was added to exponentially growing cells at time zero to block translation and induce tRNA transcription. At time 80 minutes, 300 µg/ml rifampicin was added to the
Upon addition of rifampicin, the fraction of tRNAs that are aminoacylated increases (Figure 3C) as we would expect when tRNA deacylation on the translating ribosomes comes to a standstill due to the rifampicin-induced lack of mRNA substrate. We note that charging levels do not reach 100%, suggesting either some residual protein synthesis after addition of rifampicin, or that the charging reactions catalyzed by the aminoacyl tRNA synthases do not proceed to completion. We also note that the level of tRNAs apparently increased for a short period after addition of the translational inhibitor chloramphenicol at 80 min after rifampicin addition (Fig 3A). We do not consider it likely that there is any tRNA synthesis ongoing in these rifampicin- and chloramphenicol-treated cells, but suggest that additional tRNA may be released from a cellular compartment (most likely the ribosomes) upon chloramphenicol treatment, which is not normally recovered by the fast tRNA extraction method used in this study (see Materials & Methods). The increase in tRNA level is concomitant with a minor drop in the fraction of charged tRNA, indicating that the released tRNA is in the uncharged form. Cellular tRNA and rRNA transcription rate can be increased several fold by addition of drugs that block translation, such as chloramphenicol (Dennis, 1976; Schneider et al., 2002). To examine the cellular capacity for tRNA degradation, we first induced tRNA transcription with chloramphenicol for 80 minutes, and then added rifampicin to arrest transcription and initiate tRNA degradation. Figure 3B shows that, as expected, tRNA levels rise upon addition of chloramphenicol to the culture, and reach up to approximately 10-fold the normal level after 80 minutes. Also, the fraction of tRNAs that are aminoacylated increases, consistent with the lack of protein synthesis in the chloramphenicol-treated cells (Figure 3D). However, upon subsequent addition of rifampicin, tRNA levels rapidly plummet, indicating that the effector(s) of tRNA degradation have a large capacity for tRNA turnover. Importantly, the effector(s) of tRNA degradation must already be present in the cells prior to the first addition of antibiotic in these experiments since synthesis of new proteins is inhibited by chloramphenicol, and synthesis of RNA is inhibited by rifampicin.

Discussion
E. coli bacteria carefully match the capacity of their translation apparatus to the demand for protein synthesis posed by their growth rate, which can vary greatly depending on the availability of nutrients. This relation was first noted almost 60 years ago (Schaechter et al., 1958), and depends largely on RNA polymerase partitioning between mRNA synthesis and so-called stable RNA synthesis (tRNA and rRNA). For example, tRNA and rRNA synthesis amounts to 30% of total RNA synthesis at a growth rate of 0.5 doublings per hour, whereas it amounts to 80% of total RNA synthesis at a growth rate of 2.5 doublings per hour in E. coli B/r, and independent of the exponential growth rate, this fraction drops to 24% within 15 minutes of amino acid starvation due to the (p)ppGpp-mediated stringent response (Ryals et al., 1982a).

Here, we report that tRNA is not, in fact, stable. tRNA is rapidly degraded in response to at least two conditions where the demand for tRNAs in protein synthesis is reduced, namely upon amino acid starvation and after blockage of transcription initiation with rifampicin. The degradation of tRNA seems to be independent of the amino acid starved for (Figs. 1, 2, S1, and S3) as well as the strain background (Figs. 1, S1, and S3). Furthermore, all the tRNA we have tested by Northern blot analysis showed similar fast kinetics of degradation upon starvation, and the transcriptome analysis included in Table S1 indicates that all detectable tRNAs are, in fact, unstable during amino acid starvation. In summary, we report here that the match between levels of components of the translational machinery and the demand for protein synthesis is not limited to alterations in the rates of synthesis of said components, but includes also their active degradation, at least in the case of tRNA. Thus, similar to the situation in eukaryotic cells, both the rate of synthesis and the rate of degradation of tRNA are subject to substantial regulation in E. coli.

The degradation of tRNA is measurable within minutes of amino acid starvation (Figures 1A, 2A, S1, and S3), appears to be independent of the (p)ppGpp levels (Figures 1D, 2B, and S3), and can occur when both transcription and translation is inhibited by antibiotics (Figure 3), strongly suggesting that neither the signal(s) nor the effector(s) required for tRNA degradation are synthesized de novo upon starvation. Instead, we favor a passive model, in which tRNA is rapidly degraded when it is not associated with components of the translational apparatus, such as the ribosome, elongation factor Tu, or aminoacyl tRNA synthetases (Figure 4).
In this model, there is no signal per se for tRNA degradation, and the effector(s) is likely to be any of a number of ribonucleases already present in the cell prior to starvation, which would have a lower general affinity for the tRNA than that of the specialized binding partners mentioned above. Further verification of this model and identification of the effector(s) is ongoing work in our laboratory. In essence, this passive model for degradation of tRNA is equivalent to the model suggested for rRNA degradation upon starvation (Zundel et al., 2009b), albeit operating at the faster time scale of minutes rather than hours. Interestingly, a recent report showed that *E. coli* responds to oxidative stress (0.5 mM H$_2$O$_2$) by rapidly degrading its tRNA and strongly decelerating the translation elongation rate (Zhong et al., 2015b). Oxidative stress results in oxidation of cellular RNA and subsequent stalling of ribosomes at 8-oxoG residues in the mRNA (Simms et al., 2014) Therefore, we suggest that application of oxidative stress may resemble the addition of rifampicin reported here, in the sense that both treatments are expected to block translation due to deprivation of functional mRNA substrate. The regulatory pathway that results in tRNA degradation upon oxidative stress has not been identified, but like the case reported here, it was found that tRNA
degradation occurred so rapidly that it is unlikely to be carried out by an effector synthetized
*de novo* in response to the stress condition (Zhong et al., 2015b). On this basis, we argue that
tRNA degradation upon oxidative stress may be another example of the mechanism described
in this work, whereby tRNA is degraded as a result of blockage of translation, in agreement
with the passive model for degradation described above.

Starvation for a single amino acid results in dramatically reduced charging levels of tRNAs
accepting that particular amino acid, whereas models suggest that the charging levels of
tRNAs accepting the remaining 19 amino acids should approximate 100%, because acylated
tRNAs would accumulate in the cells when the rate of tRNA deacylation on the ribosomes is
reduced as a consequence of decreased translation activity (Elf et al., 2003). In our
experiments, we clearly detected a rapid decrease in charging levels of the tRNA accepting the
amino acid subject to starvation (see tRNA<sub>ArgVYZQ</sub> in Supplementary figure S3 and data not
shown), but we only detected a slight increase in tRNA charging levels of the remaining amino
acids upon starvation (see Supplementary figures S1 and S3). Although we cannot exclude that
our tRNA extraction procedure deacylates a fraction of the tRNAs, the results from
experiments where translation is blocked completely by addition of either rifampicin or
chloramphenicol shows that we are able to extract tRNA with high charging levels (>90%) from
those samples using the same procedure (Figure 3). We propose that this apparent
discrepancy is neatly resolved by consideration of the passive model for tRNA degradation
presented in Figure 4. According to this model, only those tRNAs engaged in translation are
protected and excess tRNA, whether acylated or not, is degraded. Thus, when translation
activity is decreased upon starvation, at first due to the deficiency of charged cognate tRNA,
then after a few minutes due to the limitation of mRNA substrate caused by the (p)ppGpp
signal to RNA polymerase (Sorensen et al., 1994), the tRNA pool is decreased to match the
residual translation activity of the starving cell. The tRNA charging levels then remain
unaltered because the remaining tRNAs are engaged in the decharging-recharging cycle of
translation at nearly the same rate as during unlimited growth. In further support of this
model, we note that addition of the lacking amino acid to the starved culture caused a
significant drop in the charging level of tRNAs (see Figure S3), suggesting the tRNA
concentrations during the early nutritional shift-up may be limiting or close to limiting for
protein production.

At first glance, degradation of a fully functional and abundant macromolecular class such as
the tRNAs in response to amino acid starvation may seem contradictory to the needs of the
starving cell, as energy would then need to be spent on replenishing the tRNA pool for continued protein synthesis once the amino acid supply had been restored. We speculate that this energetic cost is out weighted by a beneficial effect of tRNA degradation on the accuracy of translation of the proteins that are synthesized during starvation conditions. In this context, it is important to recall that the cellular response to amino acid starvation is more complex than simply stopping growth. Genes involved in amino acid biosynthesis and stress response genes are upregulated by the stringent response, reflecting a reprioritization of resources from growth to stress adaptability and survival (reviewed in (Magnusson et al., 2005)). Importantly, proteins that are synthesized during amino acid starvation are prone to errors because of competition from near-cognate tRNAs at the codons that code for an amino acid(s) that is in short supply. As described in the introduction, this competition may to some degree be lessened by the halt in tRNA transcription exerted by the stringent response. We argue here that degradation of tRNA upon amino acid starvation serves to rapidly reduce the total pool of cellular tRNA, which, in turn, reduces the probability that a near-cognate tRNA delivers an incorrect amino acid at a “hungry” codon, thereby reducing the error rate in protein synthesis during starvation.

Materials & Methods

Bacterial strains used in the study

The E. coli strains used are all K12 strains.

NF915 and NF916 are CP78 and CP79 (Fiil and Friesen, 1968) cured for λ (N. Fiil). NF915 (RelA + ) genotype: thr leu his argH thi mtl supE44 relA + spoT λ λ - and NF916 (RelA-):isogenic relA2 of NF915.

Δ10 and MG1655 are the TA deleted strain isogenic to the undeleted MG1655 rph1, respectively (Maisonneuve et al., 2011).

MAS282: SpoT1 argA52 lacI q1 lacZ::Tn5 (Sorensen et. al 1994)

MAS827: SØ928: Δ(deoCABD) Δlac thi upp udp ton ((Hove-Jensen et al., 2003) made his- (insertion of miniTn10 near the his locus).

Sulfolobus solfataricus: MAS1008: S. solfataricus P2 strain B (Redder and Garrett, 2006).

Growth conditions and media

E. coli: Cultures were grown exponentially at 37.0° C in MOPS minimal medium (Neidhardt et al., 1974a) supplemented as described for each experiment. All cultures had been growing
exponentially for at least ten generations before being exposed to amino acid starvation or antibiotic treatment.

S. solfataricus: The following final concentrations were used per 500ml of 2x liquid medium stock (modified from Brock et al. (Brock et al., 1972)): 1.3 g (NH₄)₂SO₄, 70 mg CaCl₂·2H₂O, 2.25 mg Na₂B₄O₇·10H₂O, 2.8 mg FeSO₄·7H₂O, 15 µg NaMoO₄·2H₂O, 5 µg CoCl₂·7H₂O, 250 mg MgSO₄·7H₂O, 900 µg MnCl₂·4H₂O, 110 µg ZnSO₄·7H₂O, 25 µg CuCl₂·2H₂O, 15 µg VOSO₄·2H₂O. The pH was adjusted to 3.0-3.5 with a solution of sulphuric acid and autoclaved. Prior to use, the 2x stock was supplemented with 2 g sucrose, 280 mg KH₂PO₄ and autoclaved water to 1 liter. The pH was adjusted to 3.0-3.5 once again. This minimal medium supported a generation time of approx. 16h at 80° C in a shaking water bath.

Starvation
Amino acid starvation for auxotrophic strains was introduced by filtration of the culture and resuspension in medium lacking the amino acid indicated as previously described (Sorensen, 2001) or by addition of 400 µg/ml valine to the culture to introduce an isoleucine starvation (Leavitt and Umbarger, 1961).

tRNA preparation, blotting and hybridization
The preparation, blotting and hybridization were performed as described (Varshney et al., 1991) with modifications (Sorensen, 2001) with the exceptions that a Hybond N+ membrane was used, and that 10% S. solfataricus cells (measured as OD₄₃₆) was added as TCA-stopped culture to each sample before the first centrifugation. None of the probes we have used have shown any cross reactivity with RNA from the other species on the blots.

Probes used for Northern blots:
tRNA ArgVYZQ 5’-TCCGACCGCTCGGTTCGTAGC; tRNA GluTUWV 5’-CTCGTACCGCCTGAAAGGG; tRNA HisR 5’-CAGCACAATGGAACTCAAATCC; tRNA LeuPQVT 5’-GTAAGGACACTAACACCTGAAGC; tRNA LeuU 5’-TATTTGGGCActCCACCTCAAGG; tRNA LeuW 5’-CTTGCGGGCCCGAGAACTAAATC; tRNA LeuX 5’-TATTTCTACGGTTGATTTTGAA; tRNA LeuZ 5’-AAAATCCCTGCGGCGCCAGAACCTAAATC; tRNA LysQTVWY 5’-TGCGACCAATTTGATTTTGAA; tRNA LeuZ 5’-AAAATCCCTGCGGCGCCAGAACCTAAATC; tRNA LysQTVWY 5’-TGCGACCAATTTGATTTTGAA; tRNA ThrV 5’-TGGGACCTACCTCACTCAGCT; tRNA TyrTV 5’-TCGAACCTCAGCTGATGA; S. sol 5S RNA83-105 5’-GGGGCTGGGGCTGCTGAGTT; S. sol. 5S RNA62-81 5’-CAGGCCCCAATACTGAG.
Normalization of tRNA levels
Each blot was probed several times for specific tRNAs as well as the S. solfataricus 5S RNA. We measured the radioactivity present in specific bands using a phosphorimager scanner as described (Sorensen, 2001). We mixed the same amount of S. solfataricus cells into each sample of E. coli culture before the first step of RNA preparation. Therefore, we could use the signal from the S. solfataricus 5S RNA as an internal standard for gel loading and recovery of total RNA in each sample. The counts found in tRNA bands were divided by the counts found in S. solfataricus 5S RNA in the same lane and this ratio was normalized to the ratio found in the sample that was harvested before starvation.

Author contributions

The authors thank Marit Warrer for excellent technical assistance and S. Shiraz for help with the bioinformatic analysis of data in Table S1. SLS and MAS wrote the manuscript; SLS, MAS and MK designed experiments; MAS, MK, TS and AG performed experiments; MAS designed research.

Funding

We acknowledge research funded by the Danish council for Independent Research | Natural Sciences (1323-00343B) (SLS), the Lundbeck Foundation R108-A10583 (MAS) and the Danish National Research Foundation (DNRF120) (SLS; MAS).

References


Supplementary

Supplementary Table 1: Relative tRNA levels before and after Ile starvation found by transcriptome analysis by Illumina sequencing of total RNA. The values in column 2 and 3 are the average number of reads mapping to the genes in column 1. We note that tRNAs are highly modified RNA molecules and therefore underrepresented in high throughput analysis like this (Zhong 2015). We doubt that the relative number of reads between the different tRNA species may represent actual differences in concentrations but rather different difficulties for the reverse transcriptase to read through the modified bases without causing mismatch in the mapping analysis. However, for the same tRNA before and after starvation, we assume that the modification pattern is the same and that the relative difference in reads...
indicates relative differences in concentration. The average number of reads mapping to \textit{rpoS} and \textit{thrA} open reading frames were included in the analysis as positive controls.

<table>
<thead>
<tr>
<th>tRNA</th>
<th>st-st growth</th>
<th>10min starvation.</th>
<th>St-st/10min</th>
<th>anti-codon</th>
</tr>
</thead>
<tbody>
<tr>
<td>lysT</td>
<td>2367</td>
<td>49</td>
<td>48.31</td>
<td>UUU</td>
</tr>
<tr>
<td>valX</td>
<td>1941</td>
<td>41</td>
<td>46.84</td>
<td>UAC</td>
</tr>
<tr>
<td>valU</td>
<td>1423</td>
<td>36</td>
<td>39.24</td>
<td>UAC</td>
</tr>
<tr>
<td>metW</td>
<td>2771</td>
<td>75</td>
<td>36.71</td>
<td>CAU</td>
</tr>
<tr>
<td>valT</td>
<td>1300</td>
<td>37</td>
<td>35.14</td>
<td>UAC</td>
</tr>
<tr>
<td>metT</td>
<td>754</td>
<td>23</td>
<td>32.78</td>
<td>CAU</td>
</tr>
<tr>
<td>leuW</td>
<td>1353</td>
<td>44</td>
<td>30.75</td>
<td>CAA</td>
</tr>
<tr>
<td>thrV</td>
<td>853</td>
<td>30</td>
<td>28.11</td>
<td>GGU</td>
</tr>
<tr>
<td>metZ</td>
<td>1880</td>
<td>73</td>
<td>25.66</td>
<td>CAU</td>
</tr>
<tr>
<td>argV</td>
<td>571</td>
<td>23</td>
<td>24.89</td>
<td>ACG</td>
</tr>
<tr>
<td>valY</td>
<td>988</td>
<td>40</td>
<td>24.72</td>
<td>UAC</td>
</tr>
<tr>
<td>glnU</td>
<td>1199</td>
<td>53</td>
<td>22.62</td>
<td>UUG</td>
</tr>
<tr>
<td>lysY</td>
<td>600</td>
<td>27</td>
<td>22.22</td>
<td>UUU</td>
</tr>
<tr>
<td>metV</td>
<td>1174</td>
<td>57</td>
<td>20.60</td>
<td>CAU</td>
</tr>
<tr>
<td>lysW</td>
<td>885</td>
<td>43</td>
<td>20.58</td>
<td>UUU</td>
</tr>
<tr>
<td>argY</td>
<td>347</td>
<td>17</td>
<td>20.39</td>
<td>ACG</td>
</tr>
<tr>
<td>glnW</td>
<td>1036</td>
<td>51</td>
<td>20.31</td>
<td>UUG</td>
</tr>
<tr>
<td>hisR</td>
<td>1013</td>
<td>52</td>
<td>19.56</td>
<td>GUG</td>
</tr>
<tr>
<td>metU</td>
<td>585</td>
<td>30</td>
<td>19.50</td>
<td>CAU</td>
</tr>
<tr>
<td>valZ</td>
<td>938</td>
<td>50</td>
<td>18.76</td>
<td>UAC</td>
</tr>
<tr>
<td>lysZ</td>
<td>444</td>
<td>24</td>
<td>18.50</td>
<td>UUU</td>
</tr>
<tr>
<td>argX</td>
<td>730</td>
<td>41</td>
<td>17.94</td>
<td>CCG</td>
</tr>
<tr>
<td>lysV</td>
<td>337</td>
<td>23</td>
<td>14.69</td>
<td>UUU</td>
</tr>
<tr>
<td>thrU</td>
<td>696</td>
<td>52</td>
<td>13.44</td>
<td>UGU</td>
</tr>
<tr>
<td>lysQ</td>
<td>261</td>
<td>20</td>
<td>13.05</td>
<td>UUU</td>
</tr>
<tr>
<td>argZ</td>
<td>319</td>
<td>25</td>
<td>12.68</td>
<td>ACG</td>
</tr>
<tr>
<td>leuZ</td>
<td>344</td>
<td>27</td>
<td>12.56</td>
<td>GAG</td>
</tr>
<tr>
<td>ileT</td>
<td>437</td>
<td>36</td>
<td>12.05</td>
<td>GAU</td>
</tr>
<tr>
<td>cysT</td>
<td>554</td>
<td>47</td>
<td>11.70</td>
<td>GCA</td>
</tr>
<tr>
<td>ileV</td>
<td>413</td>
<td>36</td>
<td>11.47</td>
<td>GAU</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>----</td>
<td>----</td>
<td>------</td>
<td>-----</td>
</tr>
<tr>
<td>thrT</td>
<td>182</td>
<td>17</td>
<td>10.69</td>
<td>GGU</td>
</tr>
<tr>
<td>trpT</td>
<td>339</td>
<td>32</td>
<td>10.65</td>
<td>CCA</td>
</tr>
<tr>
<td>ileU</td>
<td>397</td>
<td>39</td>
<td>10.12</td>
<td>GAU</td>
</tr>
<tr>
<td>leuT</td>
<td>307</td>
<td>30</td>
<td>10.12</td>
<td>CAG</td>
</tr>
<tr>
<td>glyT</td>
<td>267</td>
<td>29</td>
<td>9.25</td>
<td>UCC</td>
</tr>
<tr>
<td>tyrU</td>
<td>504</td>
<td>55</td>
<td>9.20</td>
<td>GUA</td>
</tr>
<tr>
<td>aspT</td>
<td>342</td>
<td>38</td>
<td>8.89</td>
<td>GUC</td>
</tr>
<tr>
<td>argQ</td>
<td>236</td>
<td>27</td>
<td>8.86</td>
<td>ACG</td>
</tr>
<tr>
<td>glnV</td>
<td>469</td>
<td>53</td>
<td>8.85</td>
<td>CUG</td>
</tr>
<tr>
<td>serV</td>
<td>663</td>
<td>78</td>
<td>8.45</td>
<td>GCU</td>
</tr>
<tr>
<td>glnX</td>
<td>354</td>
<td>47</td>
<td>7.53</td>
<td>CUG</td>
</tr>
<tr>
<td>aspU</td>
<td>269</td>
<td>37</td>
<td>7.27</td>
<td>GUC</td>
</tr>
<tr>
<td>leuU</td>
<td>233</td>
<td>33</td>
<td>7.16</td>
<td>UAG</td>
</tr>
<tr>
<td>glyW</td>
<td>311</td>
<td>45</td>
<td>6.89</td>
<td>GCC</td>
</tr>
<tr>
<td>serT</td>
<td>90</td>
<td>15</td>
<td>6.00</td>
<td>UGA</td>
</tr>
<tr>
<td>gltT</td>
<td>791</td>
<td>138</td>
<td>5.72</td>
<td>UUC</td>
</tr>
<tr>
<td>valV</td>
<td>72</td>
<td>14</td>
<td>5.12</td>
<td>GAC</td>
</tr>
<tr>
<td>alaV</td>
<td>182</td>
<td>38</td>
<td>4.79</td>
<td>UGC</td>
</tr>
<tr>
<td>alaT</td>
<td>170</td>
<td>36</td>
<td>4.79</td>
<td>UGC</td>
</tr>
<tr>
<td>glyV</td>
<td>399</td>
<td>84</td>
<td>4.77</td>
<td>GCC</td>
</tr>
<tr>
<td>gltW</td>
<td>909</td>
<td>195</td>
<td>4.67</td>
<td>UUC</td>
</tr>
<tr>
<td>glyX</td>
<td>286</td>
<td>75</td>
<td>3.83</td>
<td>GCC</td>
</tr>
<tr>
<td>alaW</td>
<td>346</td>
<td>93</td>
<td>3.74</td>
<td>GGC</td>
</tr>
<tr>
<td>tyrV</td>
<td>137</td>
<td>38</td>
<td>3.63</td>
<td>GUA</td>
</tr>
<tr>
<td>alaU</td>
<td>183</td>
<td>51</td>
<td>3.58</td>
<td>UGC</td>
</tr>
<tr>
<td>tyrT</td>
<td>119</td>
<td>33</td>
<td>3.57</td>
<td>GUA</td>
</tr>
<tr>
<td>gltU</td>
<td>483</td>
<td>135</td>
<td>3.57</td>
<td>UUC</td>
</tr>
<tr>
<td>valW</td>
<td>39</td>
<td>11</td>
<td>3.51</td>
<td>GAC</td>
</tr>
<tr>
<td>metY</td>
<td>205</td>
<td>58</td>
<td>3.51</td>
<td>CAU</td>
</tr>
<tr>
<td>glyY</td>
<td>163</td>
<td>47</td>
<td>3.50</td>
<td>GCC</td>
</tr>
<tr>
<td>leuP</td>
<td>78</td>
<td>23</td>
<td>3.40</td>
<td>CAG</td>
</tr>
<tr>
<td>leuQ</td>
<td>99</td>
<td>30</td>
<td>3.34</td>
<td>CAG</td>
</tr>
<tr>
<td>aspV</td>
<td>78</td>
<td>24</td>
<td>3.25</td>
<td>GUC</td>
</tr>
<tr>
<td>gltV</td>
<td>509</td>
<td>160</td>
<td>3.18</td>
<td>UUC</td>
</tr>
<tr>
<td>Codon</td>
<td>Frequency</td>
<td>Position</td>
<td>Distance</td>
<td>Sense</td>
</tr>
<tr>
<td>-------</td>
<td>-----------</td>
<td>----------</td>
<td>----------</td>
<td>-------</td>
</tr>
<tr>
<td>argU</td>
<td>36</td>
<td>14</td>
<td>2.57</td>
<td>UCU</td>
</tr>
<tr>
<td>leuX</td>
<td>34</td>
<td>14</td>
<td>2.42</td>
<td>UAA</td>
</tr>
<tr>
<td>proM</td>
<td>41</td>
<td>17</td>
<td>2.41</td>
<td>UGG</td>
</tr>
<tr>
<td>alaX</td>
<td>221</td>
<td>93</td>
<td>2.39</td>
<td>GGC</td>
</tr>
<tr>
<td>pheV</td>
<td>42</td>
<td>19</td>
<td>2.27</td>
<td>GAA</td>
</tr>
<tr>
<td>asnT</td>
<td>54</td>
<td>26</td>
<td>2.08</td>
<td>GUU</td>
</tr>
<tr>
<td>proL</td>
<td>39</td>
<td>21</td>
<td>1.88</td>
<td>GGG</td>
</tr>
<tr>
<td>leuV</td>
<td>54</td>
<td>30</td>
<td>1.82</td>
<td>CAG</td>
</tr>
<tr>
<td>glyU</td>
<td>35</td>
<td>22</td>
<td>1.58</td>
<td>CCC</td>
</tr>
<tr>
<td>asnV</td>
<td>46</td>
<td>30</td>
<td>1.55</td>
<td>GUU</td>
</tr>
<tr>
<td>serW</td>
<td>38</td>
<td>25</td>
<td>1.52</td>
<td>GGA</td>
</tr>
<tr>
<td>proK</td>
<td>27</td>
<td>19</td>
<td>1.40</td>
<td>CGG</td>
</tr>
<tr>
<td>asnW</td>
<td>85</td>
<td>62</td>
<td>1.37</td>
<td>GUU</td>
</tr>
<tr>
<td>thrW</td>
<td>42</td>
<td>31</td>
<td>1.35</td>
<td>CGU</td>
</tr>
<tr>
<td>serU</td>
<td>21</td>
<td>16</td>
<td>1.35</td>
<td>CGA</td>
</tr>
<tr>
<td>asnU</td>
<td>66</td>
<td>50</td>
<td>1.31</td>
<td>GUU</td>
</tr>
<tr>
<td>pheU</td>
<td>39</td>
<td>32</td>
<td>1.23</td>
<td>GAA</td>
</tr>
<tr>
<td>serX</td>
<td>32</td>
<td>28</td>
<td>1.14</td>
<td>GGA</td>
</tr>
<tr>
<td>ileX</td>
<td>8</td>
<td>9</td>
<td>0.90</td>
<td>CAU</td>
</tr>
<tr>
<td>argW</td>
<td>14</td>
<td>19</td>
<td>0.73</td>
<td>CCU</td>
</tr>
<tr>
<td>selC</td>
<td>17</td>
<td>24</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td>ileY</td>
<td>2</td>
<td>7</td>
<td>0.30</td>
<td>GAU</td>
</tr>
</tbody>
</table>

| ThrA | 168 | 399 | 0.42 |
| RpoS | 204 | 500 | 0.41 |
| RrlA | 214170 | 215184 | 1.00 |
Supplementary Figure S1: tRNA degradation in MG1655 and delta-10 TA after isoleucine starvation.

Strains were grown in minimal MOPS, 0.2 % glucose 10 µg/ml Uracil and generation times were 54 min for both strains, shifted to 240 min during starvation. Starvation was introduced by addition of 400 µg/ml valine. We note that His-tRNA charging levels were underestimated due to decharging during electrophoresis (Sørensen et al 2001)

Open symbols Delta 10 Closed symbols WT (MG1655) Grown in MOPS Glucose 10µg/ml Ura
Gt=54’ and 240’ during starvation
Supplementary Figure S2: Colony counts and absorbance measurements (OD436) of Rel (NF915) and RelA (NF916) strains from 0-200 minutes after leucine starvation.
Supplementary Figure S3: tRNA degradation in the spoT mutant

Material and Methods for Supplementaries:
The RNA used for data in Table 1 was extracted from a MG1655 derivative containing a plasmid carrying an IPTG inducible birA gene (pBirA); no IPTG was added. The strain contained also a C-terminal tagged Hfq protein. The tag compared to the untagged wt did not influence the growth rate. The cells were grown in MOPS 0.2% Glucose medium supplemented with 5 μg/ml of biotin. Isoleucine starvation was introduced by addition of 400 μg/ml valine. Cells were harvested right before and 10 min after valine addition. Large volumes of cultures were harvested by cooling to 0° C in ice-water, centrifugation and wash in 0° C biotin free medium, centrifugation and resuspension in 2 ml of lysis buffer, followed by repeated rounds of sonic treatment. At this point an aliquot of each sample was phenol extracted and stored as total RNA at -80°C.

The Illumina sequencing was performed by GATC BIOTECH (http://www.gatc-biotech.com) on total RNA samples. The cDNA was random primed reverse transcripts and the second strand
DNA synthesis was also random primed. More than 2 Billion base pairs were sequenced from each sample.

We used the free software program Artemis release 16.0.0 from the Sanger Institute for our analysis. The number of reads mapped to rrlA (23S gene of ribosomal operon A) in each sample was used to normalize the amounts of RNA in each sample. The actual number of reads in the sample “10 min starvation” was multiplied by 0.74 to give the numbers in Table 1.

END OF PAPER 1
1.2 Discussion

1.2.1 Normalization

Normalization is often overlooked when planning an experiment, but when the experiment includes a comparison of data obtained in two different conditions or from two types of tissue, normalization cannot be ignored.

Our data is mainly obtained by northern where the aim is to calculate the number of transcripts per cell, allowing for comparison between two conditions (or tissues). A common strategy in northern normalization is loading equal mass of RNA into each well when loading the gel (and loading several wells with the same sample as loading control). If it can be assumed that the mass of the cell is constant during the time of the experiment, this is a simple and useful method. But as discussed in section 1.1.1.3, this assumption is far from valid at all times since RNA/protein and DNA/RNA change with growth rates and composition of the media (Dennis and Bremer, 1974; Kubitschek, 1969). Because we study tRNA levels in two conditions (interpreted as two media), we cannot use this method. Another common strategy is normalizing to the level of an internally expressed gene, assuming the transcript expressed by the gene is present at a constant level. But as we question the stability of ‘stable’ RNA normally used as an internal control, this method does not work either.

Our solution to the normalization problem has been to use an external cell as a ‘spike in’ and normalize to a transcript present in this cell. The cell used for normalization could be another type of bacteria like Sulfolobus solfataricus, but the ‘spike in’ cell could also be an E.coli expressing some transcript downstream of an externally controlled promoter. There are three criteria for a good ‘spike in’ cell:

1. The transcript used in the ‘spike in’ cell, should be abundantly expressed; this implies that only a small fraction of ‘spike in’ cells need to be added relative to the number of the sample cells.

2. The probe used to quantify the expression of the transcript from the ‘spike in’ cell should not cross hybridize with any other transcript from within the sample cell or ‘spike in’ cell.
3. The size of the transcript from the ‘spike in’ cell should be suitable as compared to the size of the transcript of interest in the sample cells, but the signal should not overlap.

Such a ‘spike in’ cell can be used to monitor cell opening, for processing errors between samples, and as a loading control. For this method to be reliable, the cell numbers of the sample cells must not change during the time of the experiment. Figure S2 in Paper 1 clearly shows that this criterion is fully filled in our setup.

In praxis, we added the same volume of the ‘spike in’ cell to all the withdrawn samples after the experiment was performed, and then we processed the sample and the ‘spike in’ cells together. By doing so, there are still three minor errors: one is a pipetting error when transferring the ‘spike in’ cells to the samples, the second is the homogeneity of the culture containing the ‘spike in’ cells, and the third is the cell-opening efficiency of the sample cell. After all, the cell-opening efficiency of the sample cell might change during the time of the experiment. The two first errors can be lowered by being careful at the step of adding the normalization cell, and the third is compensated for by choosing an appropriate method for cell opening. The last possibility for proper normalization across two conditions is additions of ‘spike in RNA’ after the sample cells are harvested and opened. This strategy does not account for the cell-opening process and the first handling but is otherwise the same as using ‘spike in’ cells.

1.2.2 RNases involved

The trigger for ribosome degradation during six hours of carbon starvation showed to be an increase in the free ribosome subunits, 30S and 50S, that exposed the interaction surface of the two subunits where large part of the rRNA is located. The increased formation of subunits was generated by a reduction of translation activity in the carbon starved cells (Zundel et al., 2009a). In our model, ‘the passive model of tRNA degradation’, we propose that the reduction in translational activity, due to the amino acid starvation, generates vacant tRNA, and vacant tRNAs are prone to degradation. The main difference between the two models is the response time. We see tRNA to be down regulated to 60% after only ~25 minutes where the ribosomal RNA only reaches such low levels after 24 hours. One reason for this difference in rate of degradation could be the formation of the 100s particles since the induction of (p)ppGpp upregulates expression of rfm (Izutsu et al., 2001) and RMFs are responsible for the formation
of the 100s particle, keeping at least a fraction of the unoccupied ribosomes unavailable for RNase degradation.

However, both our and Zundel’s results suggests that the effectors of the degradation are present before starvation is triggered. The most obvious candidate for mediating the tRNA degradation is one of the several RNases in *E. coli*. (See Table 1 in Section 1.1.4 for a selection.) Of particular interest is RNase E, the endoribonuclease of the degradosome. Apart from RNase E, the degradosome includes the helicase RhlB and the exonuclease PNPase (Carpousis, 2007; Mackie, 2013). The degradosome is believed to be an important component of mRNA decay, but fragments of rRNA have also been found associated; besides, the degradosome can degrade rRNA *in vitro* (Bessarab et al., 1998). Meanwhile, just deleting RNase E to test the effect of the degradosome is not possible since inactivating RNase E leaves the cell inviable (Liang and Deutscher, 2013). To overcome this problem, temperature-sensitive RNase E mutant has been developed, allowing researchers to test if this mutant is implicated in the tRNA degradation during starvation.

Another problem with locating the RNase(s) responsible for the tRNA degradation is RNase redundancy. If one is deleted, other RNases take over (Basturea et al., 2011; Belasco and Higgins, 1988; Houseley and Tollervey, 2009; Sulthana and Deutscher, 2013). Therefore, several RNases were deleted in order to locate the RNases responsible for rRNA degradation during the six hours of carbon starvation (Basturea et al., 2011; Liang and Deutscher, 2013). Another interesting RNase to test could be RNase I. Davis *et. al.* (Davis et al., 1986) found tRNA to be stable during 75 hours of phosphate starvation by using an RNase I mutant as background strain. RNase I is an unspecific endonuclease mainly located in the periplasm of *E. coli*, but minor numbers of RNase I also seems to be located in the cytoplasm in some inactivated form. The mechanism for activation and deactivation of the cytoplasmic-contended RNase I is unknown (Deutscher, 2009). RNase I mutants are common as background strains when studying RNA maturation or degradation. This happens probably to avoid degradation when cells are disrupted, but in the experiments of David *et. al.*, rRNA was degraded during the time of the experiment but not the tRNA. This means that the RNase(s) responsible for tRNA regulation during amino acid starvation could actually be the RNase I.

**1.2.3 Our hypothesis in context**

In this section, I have tried to include our result in the context of adaptation during amino acid starvation as described by the different
models in the Introduction and in Paper 1. Our results are relevant at the time span, just after induction of starvation, and until the cell has either synthesized the amino acid starved for (Traxler et al., 2008a) or by other means adapted to the new condition. Our model explains how the cell avoids a devastating error in the proteins expressed after starvation is induced. Fig. 4 is a qualitative illustration of the tRNA, mRNA, and (p)ppGpp levels per cell from before the system was perturbed and through the period of adaptation. It also depicts the different events associated with the perturbation as well.

The set out is steady-state growth and an average constant amount of mRNA, rRNA, tRNA, per cell, reflecting the composition of the growth media (Dennis and Bremer, 1974; Maaløe and Kjeldgaard, 1966b; Shehata and Marr, 1971). The sudden depletion of one amino acid increases the possibility of incorporating an uncharged tRNA at the A-site of the translating ribosomes (marked by 1 in Fig.4). The uncharged tRNA at the A-site is possibly directly interacting with the ribosomal L11 and the RelA (Agirrezabala et al., 2013b) thereby activating the stringent response. The ‘hopping’ model (marked by 2 in Fig.4) now describes how relatively few RelA are able to quickly increase the (p)ppGpp level by ‘hopping’ from ribosome to ribosome while catalyzing the production of (p)ppGpp (English et al., 2011). The increased level of (p)ppGpp decreases the RNA chain growth rate (marked by 3 in Fig.4) (Sørensen et al., 1994; Vogel et al., 1992) probably by direct interaction between the RNAP and the (p)ppGpp (Liu et al., 2015). This process leads to a decrease in free RNAP and a decrease in mRNA levels (marked by 4 in Fig.4), affecting the mRNA expressed from gene with promoters that are hard to saturate the most (Jensen and Pedersen, 1990), in Fig.4 these mRNA are named ‘mRNA hard’. At times, just after depleting one amino acid, the competition between the depleted tRNA and the near cognate tRNA must be severe, implying increased frequency in the error of translation(Kramer and Farabaugh, 2007; Sørensen et al., 1994). The depletion in one tRNA lowers the activity of translation and resembles the general situation of the rel‘ mutant during starvation at first. Within 2 to 4 minutes, the mRNAs pool is reduced (Kaberdin et al., 2011), resulting in the mRNA pool being the limiting factor for translation activity. This then leaves a fraction of the tRNA in vacancy and prone to degradation (marked by 5 in Fig.4). As soon as the vacant tRNAs are degraded, the competition between the depleted tRNA and the near-cognate tRNA is reduced, and so is the error frequency of translation (Sørensen et al., 1994). Thereby it is ensured that as little energy as possible is wasted in non-functional proteins during the period of adaptation. The level of mRNA expressed from genes with promoters that are easy to saturate (in Fig.4 these are named ‘mRNA easy’), is by this process upregulated
relatively to the level of mRNA expressed from genes with promoters hard to saturate (marked by 6 in Fig.4). The increased in (p)ppGpp also increases the expression of rfm; thus, there is a possibility that 100S formation is increased (Yoshida and Wada, 2014) (marked by 7 in Fig.4).

Figure 4. Quantitative overview of the (p)ppGpp, tRNA, mRNA levels per cell, from the point of starvation until adaptation is completed. See text in Section 1.2.3 for details.

1.3 Conclusion

In conclusion of the first part, we were able to demonstrate that tRNA is rapidly degraded after the induction of amino acid starvation. This was accomplished by using the developed normalization strategy, proved to be suitable since there was no cell division during the time of the experiment.

The initial tRNA degradation occurs with similar kinetics in the rel+ and rel− strains; thus, activation of the stringent response cannot trigger the activation of the degradation. Upon rifampicin treatment, the degradation pattern resembles that of amino acid starvation, leading us to propose that any condition that results in a decreased demand of tRNA due to a reduction in protein synthesis that activates the tRNA degradation response. The proposed model, ‘the passive model of tRNA degradation’, is similar to the model proposed by Zundel et. al. (Zundel et al., 2009a), and the effectors of the degradation are probably one or more of the RNases already present before starvation. Thus we have shed light on a possible general response to decreased transcription activity, namely tRNA degradation, resulting in a new
tRNA plateau matching the new transcriptional activity. This leads to a reduced error frequency in translation and a savings of valuable energy in the adaptation phase of bacteria *E. coli*.

### 1.4 Future perspectives

*Below, the most important future experiments are briefly outlined. First of all, knowing who is responsible for the observed tRNA degradation seems important. These are ongoing projects in our laboratory. At the moment, we have two approaches.*

1. **The direct method:** Locate the RNases responsible for the tRNA degradation by deleting/inactivating one or several RNases; alternatively overexpress an inhibitor of the RNases tested; then repeat the amino acid starvation experiment for the wt and the mutant strain; if the life time of the tRNA is prolonged in the mutant, it is implicated.

2. **The indirect method by selection:** Locate the effectors of the tRNA degradation by using a lambda phage with a stop codon mutation in the CI gene and a plasmid expressing a nonsense suppressor tRNA under control of a promoter than can be externally induced; after amino acid starvation and lambda phage attack, only the cells with stable tRNA (and nonsense suppressor tRNA) will have phages that are able to enter the lysogenic cycle. Thereby only the cells with stable tRNA will have a possibility of surviving the lambda phage attack.

   *Secondly it seems important to test further to see if it is true that the tRNA degradation is coupled to reduction in translational activity. Therefore, two other experiments might be interesting to perform.*

3. **Test** to see if the proposed model is correct by starving the cells for other nutrients, or expose the cells to other stress conditions, leading to a reduction in translation; then record the tRNA (and the charging) level through the perturbation; this has already been done by Zhong *et al.* who recorded tRNA levels after exposing the cells to oxidative stress, thus supporting our model; however, the procedure could be repeated for other nutrients or stress factors as well (Zhong et al., 2015a).

4. **Test** if there is a correlation between translation activity and the tRNA level both in the adaptation phase and after. Finding such relation would confirm our
hypothesis; reduction in translational activity leads to degradation of tRNA. Thus raising the possibility that the ratio between the rRNA/(tRNA + rRNA), found to be constant at the different grow rates (Dennis, 1972; Dennis and Bremer, 1974), still is constant during the adaptation phase, if only considering the total translational activity per cell and not the numbers of ribosomes. This could be accomplished by radioactively labeling of the cellular proteins as performed by O'Farrell (O'Farrell, 1978), parallel to withdrawing the samples for quantification of the tRNA level as in our ‘standard’ experiment.

Finally, it would be very interesting to know if the found mechanism of tRNA degradation is conserved in other species.

5. To accomplish this goal, repeat the amino acid starvation experiment for other species by starving other cultivable bacteria, and record their tRNA levels before and after the starvation.
2 Part

2.1 Introduction: TLR and Hfq

Because the lifetime of tRNA after starvation showed to be prolonged in a MG1655 Δhfq (unpublished), we hypothesized that the tRNA regulation, apart from Hfq, also involved transcripts from a family of intergenic repeats, namely the TLRs. Therefore, we constructed an MG1655 mutant where all TLRs were deleted. Unfortunately, the experiments showed that TLRs were not involved in tRNA regulation. Instead, we found 3 transcripts from the TLR to be enriched at RNA chaperone Hfq. Thus, the TLRs and Hfq are the subject of Part 2.

2.1.1 The family of tRNA-linked repeats

The family of TLR is made of intergenic repeats located in operons of stable RNA. We came across their existence since the majority of the TLRs are located in operons of tRNA.

Based on homology, the TLRs can be grouped in two groups: TLRA and TLRB. (See Fig. 5A and 5B for clustalX2 alignment of the TLRA and TLRB, respectively.) The TLRA contains homology at the middle section and the 3’ end, and the TLRB contains the homology at the 5’ end. The locARNA structure of TLRA (see Fig. 5C) is equipped with two small stems marked with I and II. Stem I displays complete structural and sequence homology, but stem II only partial. The locARNA structure of TLRB (see in Fig. 5D) has a long stem marked with I containing almost complete structural and sequence homology, but as the TLRB only consists of 5 TLRs, the level of homology is not remarkable. (For an interpretation of the color code, see the legend and lower corner of Fig. 5D).

The discovery of the first TLRs was made in 1978 by J. Egan and A. Landy (Egan and Landy, 1978). These TLRs were known as the 3.14 repeat and were located in the operon of tyr tRNA. Egan and Landy saw that that each repeat had 19 nt of sequence identity to the 3’end of the mature tyr-tRNA. (See Fig. 6 framed areas.)
The family of TLR grew during 80’ and the 90’ and TLRs were discovered in operon of tRNA; proK (Kuchino et al., 1985), lysT, lysW, lysY, lysZ, and lysQ (Yoshimura et al., 1984) and serV, argQ, argV, argY, argZ, and serX (Komine et al., 1990).

Finally, in 1999, three TLRs located downstream the rrfA, rrfB, and rrfF were described by K. E. Rudd (Rudd, 1999), and so far, the *E.coli* TLR family consists of 22 TLRs.

According to Ecogene (Zhou and Rudd, 2013), there are two naming conventions for the TLR:

1. The first letter in the name is ‘r’ for repeat; the second letter is the first letter of the tRNA or 5S rRNA gene immediately upstream from the TLR; and the last capitalized letter is the fourth letter of the linked tRNA or 5S rRNA gene.
2. TLRs are subdivided into two classes based on their homology: TLRA and TLRB (Fig. 4). These subgroups are then numbered clockwise round the *E.coli* chromosome.

The 18-19 nt identity to the 3’end of the mature tRNA is the hallmark for all TLRs belonging to the tRNA operons (Komine et al., 1990), leading Rudd to suggest that the TLRs may be remnants of cryptic prophage or plasmid integration and excision (Rudd, 1999). Only one transcript, the *rtT* (TLRA9), located at the 3.14 repeat, has been discovered. This transcript was mapped twice with two different results: A transcript of 171 nt mapped by Bosl and Kersten (Bosl and Kersten, 1991) and a transcript of 129 nt found by Li and Deutscher (Li and Deutscher, 2002).

**Figure 5: ClustalX 2 alignment of the TLRs.** Based on homology, the TLRs can be grouped in two: A) TLRA, containing 17 TLRs and B) TLRB, containing 5 TLRs. The identity of the nucleotides is marked by color code; thymine is green, cytosine is blue, guanidine is orange, and adenosine is red. There is a star above marking complete homology. Above the alignment, gray bars mark two stems, I and II in Fig. 5A and one stem, I in 5B. These stems are predicted by the locARNA structures at Fig. 5C and 5D. The base pairs of the locARNA structures are colored according to both the level of structure and sequence conservation; red means that the base pair of the predicted structure is found to consist of only one type of base pair; yellow means that the base pair of the predicted structure is found to consist of two types of compatible base pairs, etc. (The color code is inserted at the lower corner of 5D.) The saturation level of the colors decreases according to the number of incompatible base pairs formed to obtain the predicted structure. Full saturation means all bases in the alignment do base pairs with a compatible base. The compatible type of base pairs are C-G, G-C, A-U, U-A, G-U, and U-G. A. (Amit et al., 2012; Will et al., 2012).
In an attempt to find the function of the 3.14 repeat, *in vitro* and *vivo* studies were performed and revealed that the TLRs located downstream from the tyr tRNA were working as rho-dependent terminators (Kupper et al., 1978; Madden and Landy, 1989).

Since we propose that at least three of the TLRs function as regulatory RNA, a brief description of regulatory RNAs is given in the next section.

![Fig 5](image.png)

**Fig. 5**: Location of the TLRs downstream tRNA\(^{\text{tyrT}}\) and tRNA\(^{\text{tyrV}}\). At the right side of the figure, the genes encoded by the sequences are written. There are two tRNAs: tRNA\(^{\text{tyrT}}\) and tRNA\(^{\text{tyrV}}\) marked by a gray area below the letters. There are also four TLRs: rt\(T\) (TLR9), rt\(V\)1 (TLR8), rt\(V\)2 (TLR7), and the rt\(V\)3 (TLR6). The frames mark the 19 nt of sequence identity between the 3’ end of the mature tRNA and the TLRs. This identity implies that most of the annotated TLRs contain the 3’ end CCA, which is also characteristic of tRNA.

### 2.1.2 Regulatory RNA in general

The total RNOME of a bacteria cell can roughly be divided into three classes of RNAs: the RNAs that are translated into protein (the mRNA), the RNA used in the translation machinery (the rRNA/tRNA), and the RNA with other regulatory functions than just mentioned, here abbreviated as regRNA. Until today (2016.04.25), there are 108 regRNAs from *E. coli* K-12 registered in the Bacterial Small-Regulatory-RNA-Database (Li L., 2013).

The function of regRNA in bacteria are as post-transcriptional regulators involved in stress response and virulence (as reviewed in (Sobrero and Valverde, 2012; Vogel and Luisi, 2011), quorum sensing (Svenningsen et al., 2008), and carbon metabolism (Vanderpool, 2007; Wadler and Vanderpool, 2007), but also as RNA-preventing transcriptional noise from the other regRNAs (Lalaouna et al., 2015). Regulation can be performed by cis-encoded antisense
RNAs (asRNA), where base pairing between the mRNA target and the asRNA are fully complementary, or by trans-encoded RNAs (sRNA), where the sRNA is only partially complementary to the target. From the cellular point of view, regulation by RNA must be favorable since; it provides the cell with an additional level of regulation, the production of RNA has a low metabolic cost compared to the cost of producing proteins, and it allows for fast response due to; the short production time (Vogel and Luisi, 2011).

2.1.3 The chaperone Hfq

The regRNAs are largely able to efficiently perform their task only with the help of the chaperone protein Hfq. Hfq was first discovered as the host factor for the phage Qβ (de Fernandez et al., 1972) but has in the last two decades been known for facilitating sRNA mRNA interaction and fine tuning post-transcriptional expression during stress. Recently, evidence of asRNA, DNA interactions and sponge-like-function has also emerged (Lalaouna et al., 2015; Opdyke et al., 2004; Ross et al., 2013; Updegrove et al., 2010).

The inactivation of Hfq causes a broad pleiotropic phenotype such as a decrease in growth rate, impaired stress response, a loss of virulence, and altered metabolic regulation (Chao and Vogel, 2010; Kendall et al., 2011; Tsui et al., 1994). In the following section, the term Hfq6 is used when referring to the Hfq hexamer and when referring to the monomer Hfq.

2.1.3.1 Conservation of Hfq

Hfq from Escherichia coli is only 102 amino acids long, and the N-terminal is highly conserved in α/β/γ-proteobacteria in low GC-gram-positive bacteria, and also, but to a lesser extent, in Thermotogales Aquificales (Sun et al., 2002). The N-terminal of Hfq contains the Sm1 motive, found in the Sm proteins; this is an essential component of the Eukaryotes and Archea spliceosome. Thermotogales Aquificales are proposed to be of the early divergent, suggesting that Hfq is an ancestral Sm protein (Achenbach-Richter et al., 1987; Deckert et al., 1998). The last ~20 amino acids of the extended C-terminal found in E.coli, is only conserved in 7 closer related spices (Sun et al., 2002).

2.1.3.2 Structure of Hfq

There are an increasing number of reports stating the importance of the Hfq C-terminal, and in 2011, the full length of Hfq form E.coli was successfully crystalized. (See Fig. 7A-D.) (Beich-
Frandsen et al., 2011b). The structure revealed that Hfq is packed as hexamers, resembling a toroidal doughnut with a doughnut ring diameter of 65 Å, a thickness of 28 Å, and with the central poor of 11 Å. One side of the torus is termed the proximal face (Fig. 7A) and the other side the distal face (Fig. 7D). A third surface, termed the lateral or the rim face (Fig. 7C), connects the proximal and distal face. The core (residue 7-65) of the full length of Hfq, seen in Fig. 7A-D, is similar to the structure solved by Sauter et al. eight years previously (Sauter et al., 2003), but Sauter et al. only solved the structure of the core segment, and the location of the C-terminal has therefore remained unknown. In Fig. 7B and 7C, the location of the C-terminal is seen, clinging along the rim and terminating in the solvent on the proximal face. (See green arrow, Fig 7C.)

Since the crystallization of proteins often implies using buffers that are distinct from the biological relevant, using a much higher protein concentration than in the cell (Whitford, 2005), it is standard procedure to verify that the crystalline structure is biologically relevant. Therefore Beich-Frandsen et al. resolved the full length Hfq structure in a more biologically relevant buffer by using a combination of synchrotron radiation circular dichroism (SRCD), nuclear magnetic resonance (NMR), and small angle X-ray scattering (SAXS). From this data, they created the superposition, as seen in Fig. 7E, containing the ab initio model based on the SAXS data, the coordinates from the Hfq core structure solved by Sauter et al. (Sauter et al., 2003), and three rigid body simulations. (See legend, Fig. 7, for symbols of the different models.)

As Fig. 7E shows, the different models are not completely superimposed in the area of the C-terminal. Even so, all models agree on the C-terminal extending laterally out from the core. The authors suggest that the difference between the SAXS ab initio model and the crystalline form of Hfq are due to the crystal being in a lower energy state (crystal was formed at 5 °C), as compared to the structure obtained by SAXS acquired at 37°C.
7A-D) Images of the *E. coli* Hfq crystal structure: The structure is solved by Beich-Frandsen et al. to a resolution of 2.85 Å. The crystal structure is seen from different sides: A) the proximal face, B) the tilted side view, C) the lateral face, and D) the distal face. The pink-colored areas are the alpha helix, the yellow areas are the beta barrels, and white strands are the disordered areas including the C-terminal. The N terminal is marked by a blue arrow, and the C-terminal with a green arrow (the arrows are inserted by the author). E) Hfq superposition of the PDB structure 1HK9, is represented as dark-gray ribbons, cylinders, and spheres; the Hfq SAXS *ab initio* model is based on the full length *E. coli* Hfq in solution represented as light-gray/blue spheres; and three rigid body models represented as red, green, and gold spheres. The 7E upper panel is the lateral face; the 7E lower panel is the proximal face. Figure 7A-D are adapted from the JSmol viewer at the Protein Data Bank using accession number 3QHS. Figure 7E is adapted from (Beich-Frandsen et al., 2011a).

2.1.3.3 Properties of Hfq

The ‘canonical function’ of Hfq is linked to the properties of the proximal and the distal faces. By crystallization, it was shown that RNA, bound to Hfq, is located in circular groves at the distal and the proximal face, but the two faces display different binding capabilities. 

The proximal surface generally binds A/U-rich single-stranded regions, located adjacent to a stem-loop structure, which are motives typically found in sRNAs (Mikulecky, 2004; Schumacher et al., 2002). The distal face seems to prefer the ARN or the AAN repeat often found in mRNAs (where A is an adenine, R is a purine, and N can be any nucleotide) (Link et al., 2009; Robinson et al., 2014; Schu et al., 2015b). The binding capacity of the lateral face and C-terminal is less explored, but the lateral face may be important for binding internal U-rich sequences of RNAs (Sauer et al., 2012; Schu et al., 2015b), and the C-terminal has been
shown to interact with mRNA (Robinson et al., 2014). It might also play an important role in DNA interaction. The function of the DNA interaction is still unknown, but DNA encoding membrane proteins were overrepresented in DNA co-precipitated with an N-terminal tagged Hfq6, and Hfq lacking the C-terminal was not able to interact with DNA (Updegrove et al., 2010).

Canonically, the function of Hfq is promoting interaction between the regRNA and the mRNA target (duplex RNA) leading to the inhibition of translation (Fig. 8A), an increase in translation (Fig. 8B), or a coupled degradation of the duplex RNA (Fig. 8D, E). Interaction between Hfq6 and the regRNA can prolong the regRNA lifetime (Fig. 8C), and interaction between mRNA and Hfq6 can stimulate mRNA polyadenylation. (See also legend Fig. 8.) (Vogel and Luisi, 2011).

Figure 8: The widely accepted model of the canonical Hfq function. A) Inhibition of translation: 30S and the complex Hfq6 – regRNA compete for the ribosomal binding site (RBS), leading to inhibition of translation and down regulation if the Hfq6–regRNA is the first binding. B) Activation of translation: Hfq promotes unfolding of the mRNA secondary structure, otherwise sequestered, leading to transcription. C) RNase E and Hfq6 compete for the RNA: If the RNA binds Hfq6, the lifetime of the RNA is prolonged. D) Hfq6–regRNA promotes coupled degradation of the duplex RNA. E) Stimulation of polyadenylation: mRNA interaction with
Hfq6 stimulates polyadenylation of mRNA, leading to exoribonuclease degradation. The figure is adapted from (Vogel and Luisi, 2011).

The regRNAs that interact with Hfq6 have been categorized into two classes based on their stability in different Hfq6 mutants after overexpressing the regRNA with and without addition of rifampicin (Schu et al., 2015b). The two underlying assumptions in the analysis include, first, that interaction with Hfq6 prolongs the stability of the regRNA. Thus, if the interaction is disrupted because of site specific Hfq6 mutation, and the regRNA displays instability, the regRNA is interacting with the site of Hfq6 where the mutation is located. Secondly, pairing between target mRNAs and the regRNA leads to a coupled displacement of the duplex RNAs making the RNA complex prone to degradation. Thus disruption of the mRNA binding site at Hfq6 prolongs the stability of the cognate regRNA. Class 1 regRNAs involved RNA using the proximal and the lateral face for binding and interacting with cognate mRNA with distal face binding motives. Upon Hfq6 detachment, the duplex RNA in Class 1 showed coupled RNA degradation.

The regRNA of Class 2 uses the proximal and distal faces for binding, and their cognate mRNA has lateral face-binding motives. In Class 2, coupled duplex RNA degradation upon detachment from Hfq6, is not always seen.

Recently, evidence for Hfq assisting the transcript 3’ETS\textsuperscript{LeuZ} in a sponge-like function, sponging up transcriptional noise from the sRNA, RyhB, and RybB, has appeared. The transcript 3’ETS\textsuperscript{LeuZ} is located at the 3’ end of the glyW-cysT-leuZ polycistronic tRNA transcript, linking the expression of 3’ETS\textsuperscript{LeuZ} to growth instead of stress (Lalaouna et al., 2015). Four other regRNAs, having sponge-like functions, have formerly been reported in E.coli, but not all need assistance by Hfq (Figueroa-Bossi et al., 2009; Miyakoshi et al., 2015; Overgaard et al., 2009; Tree et al., 2014)

2.1.3.4 Cellular concentration and location of Hfq

For Hfq-mediated regulation to happen, the regRNA, the target RNA, and Hfq6 must meet. Therefore, the affinity, the cellular or local concentration, and the conformation of the component are important. Recordings of the in vitro polymerization of Hfq, in potassium phosphate buffer pH 7.5, showed that the monomer-to-hexamer transition of Hfq occurred at 0.8 μM Hfq; at this concentration, half of Hfq was hexameric. The concentration where half of
the Hfq was bound in higher multimers was at 4.9 µM Hfq, and the conformation with the highest affinity towards RNA was the hexameric (Panja and Woodson, 2012).

Since half of Hfq apparently is polymerized into higher-order structures at 4.9 µM Hfq, and this conformation seems to have lower affinity towards RNA, the cellular concentration of Hfq is of high interest. Thus, the numbers relevant for calculating the cellular concentration of Hfq in *E. coli*, is shown in Table 2. The cellular amount of Hfq has been reported for *E. coli* during exponential growth in YT, LB, M9, and for the stationary phase in LB.

The numbers of Hfq in YT (first row) contain only a fraction of the Hfq remaining in the supernatant after 1 hour of centrifugation at 13.000 g. Accordingly, the number may not contain the total Hfq but only a fraction (Carmichael et al., 1975). The estimation of Hfq in LB/0.2 % glucose (last row) was reported to be the cytosolic fraction of Hfq (Ishihama et al., 2008). Therefore, in neglecting the first and the last rows, it seems as if Hfq is stably present during the exponential growth and stationary phases at a concentration of ~20 µM.

If the transition between the monomer, hexamer, and polymer conformations of Hfq *in vivo*, resembles the *in vitro* transition measured by Panja and Woodson, an Hfq concentration of ~20 µM suggests Hfq to be assembled into higher ordered complexes *in vivo*.

At the quaternary Hfq6, the structure displayed in Fig. 8E, the C-terminal of Hfq is extending out from the core of the Hfq6, making it prone to interaction with other Hfq6 or other cellular components. Assembling of the Hfq into higher-order fibrillary structures has recently been confirmed (Fortas et al., 2015).

<table>
<thead>
<tr>
<th>media</th>
<th>Hfq per cell</th>
<th>In molar</th>
<th>Mean volume cell 10^-15</th>
<th>Hfq uM</th>
</tr>
</thead>
<tbody>
<tr>
<td>YT (exponential)</td>
<td>2500 (1)</td>
<td>4*10^-21</td>
<td>4 (vol. from LB glucose) (4) #</td>
<td>1</td>
</tr>
<tr>
<td>M9/0.4%glu (exponential)</td>
<td>45000 (2)</td>
<td>7.5*10^-20</td>
<td>3.2 (4)</td>
<td>23</td>
</tr>
<tr>
<td>LB (exponential)</td>
<td>55000 (3)</td>
<td>9*10^-20</td>
<td>4.4 (4)</td>
<td>20</td>
</tr>
<tr>
<td>LB (stationary phase)</td>
<td>18000 (3)</td>
<td>3*10^-20</td>
<td>1.5 (4)</td>
<td>20</td>
</tr>
<tr>
<td>LB/0.2% glu (exponential)</td>
<td>5800 cytosolic (5)</td>
<td>1* 10^-20</td>
<td>4 (4)</td>
<td>2.5</td>
</tr>
</tbody>
</table>

Table 2: Numbers relevant for calculating the cellular concentration of Hfq6. The medium for the experiment is written in the first row, and the condition is noted in the parenthesis. The numbers in the parenthesis mark the article from which the values are adapted: 1 (Carmichael et al., 1975) and 2 (Kajitani et al., 1994), in the article, the numbers are given in the range 5000-10,000 and the concentration is calculated from the mean (7500), 3 (Ali Azam et al., 1999), 4 (Volkmer and Heinemann, 2011) the cell volume marked # is taken from cells grown in LB 0.4% glucose, 5 cytosolic Hfq grown in LB 0.2 % glucose (Ishihama et al., 2008).
The Hfq fibrillary structures were repeatedly localized at an angle of 45° degrees along the axis of the *E.coli* cell, and the structures seemed to coil around close to the periphery of the cytoplasm, suggesting inner-membrane interaction (Fortas et al., 2015). Hfq seemed to be colocalized in this repetitive pattern together with RNase E, RhlB, RNaseIII, PAPI, RnaA, and RraB (Taghbalout and Rothfield, 2008; Taghbalout et al., 2014) where RraA and RraB are protein inhibitors of RNaseE (Zhou et al., 2009). Hfq localization close to the inner membrane, but also in the cytosol has been reported by Diestra et al. (Diestra et al., 2009). The experiment conducted by Taghbalout et al. was conducted by using immunofluorescence, a method that perforates the cell membrane, raising the possibility that the cellular RNA will be decreased, or that the Hfq in general will be in a non-biological relevant state. The other experiment carried out by Diestra et al. was performed by using an electron microscope, were a tag of ~5 kDa was fused to Hfq. Moreover, the treatment to obtain the images was harsh. Therefore both reports could simply be artifacts from the experimental setup; however the results of the experiments could be valid.

2.1.3.5 RNA cycling at Hfq

There is experimental evidence for Hfq to be saturated with RNA *in vivo* (Hussein and Lim, 2011; Moon and Gottesman, 2011), but the low dissociating rates between Hfq<sub>6</sub> and the cognate duplex RNA measured *in vitro* (Fender et al., 2010; Olejniczak, 2011; Salim and Feig, 2010), suggest long residential time of the duplex RNA bound Hfq<sub>6</sub>, and therefore, a long time before a new RNA can access Hfq<sub>6</sub>. This does not fit well with the findings that at sudden stress, the Hfq-mediated stress response is activated within 1 to 2 minutes (Massé et al., 2003; Papenfort et al., 2006).

As described in Section 2.1.3.3, Lalaouna et al. subdivide the RNA interacting with Hfq<sub>6</sub> into two classes. Class 1 was assumed to display a coupled degradation of the duplex RNA upon parring at Hfq<sub>6</sub> (Lalaouna et al., 2015), possibly mediated by RNase E (Massé et al., 2003). The coupled duplex RNA degradation is one model explaining how new RNA gain access to Hfq<sub>6</sub>. But what about the Class 1 RNA that do not meet their cognate RNA when bound Hfq<sub>6</sub>? Will it reside at Hfq<sub>6</sub> according to its particular dissociation constant? Moreover, the Class 2 RNA did not always display coupled degradation. How are these RNAs displaced from Hfq<sub>6</sub>? In the ‘active RNA-cycling model’ (Fender et al., 2010), the duplex RNA or single RNA is relived from Hfq<sub>6</sub> by the competition from unbound RNAs. In this model, the RNA interacting with Hfq<sub>6</sub> still has side preference; RNA with the AAN motive prefers the distal face, but the RNA
binds other faces if the preferred side is occupied. Therefore, most RNAs are able to displace an already-bound RNA from Hfq₆ and also duplex RNA in complex with Hfq₆. The competing power of an RNA displacing an already bound RNA is not solely a matter of the affinity, but may also depend on the specific interaction between Hfq₆ and that particular RNA, displacing that particular RNA that is already bound.

The active cycling model seems appealing since the displacement of duplex RNA from Hfq₆ is linked to cellular demands; a sudden stress induction of RNA will displace already bound RNA when needed, and the more urgent (the more RNA induced), the faster the displacement. But which model best describes the Hfq₆ mediated RNA regulation has not yet been solved. The development of models is complicated by the lack of basic knowledge as such: What is the concentration of Hfq in the cell at different conditions? Where is Hfq located in the cell and in which configuration? Is Hfq in fibril conformation biological relevant? Or will the cellular RNA prevent such formation? What is the equilibrium among the fibrillary, the hexametric, the monomers, and the other mers of Hfq, at different conditions (also in presence of RNA)? Are the fibril-binding properties in vivo different than the hexametric as found by Panja and Woodson in vitro (Panja and Woodson, 2012)?
Paper 2
Title: Transcripts from the family of tRNA-linked repeats interact with Hfq

Short title: new RNA is stabilized by Hfq

Authors: Mette Kongstad¹, Thomas Søndergaard Stenum¹, Birgitte H. Kallipolitis², Sine Lo Svenningsen¹, Michael A. Sørensen¹.

¹Department of Biology, University of Copenhagen, Ole Maaløes Vej 5, 2200 Copenhagen N, Denmark.
²Department of Biochemistry and Molecular Biology, Southern Danish University, Campusvej 55, 5230 Odense, Denmark

Corresponding author: Associate Professor Michael A. Sørensen, Phone: +45 35323711. Email address: MAS@bio.ku.dk.

Keywords (minimum 3): tRNA-linked-repeats; Hfq; Escherichia coli; rrn operon

Abstract

The family of tRNA-linked repeats (TLRs) in Escherichia coli consists of 22 intergenic repeats, annotated at 68-181 bp in length, and located in operons encoding components of the translation machinery. Only RNA from the rtT area has previously been detected, and no other function than as Rho-dependent terminator has been described for any of the TLRs (Bosl and Kersten, 1991; Li and Deutscher, 2002). We report here that the TLRs RrA, RrB and RrF, which are transcribed from the rrn operons encoding ribosomal RNAs, accumulate in E. coli during exponential growth, and that a processed RrA, RrB, RrF transcript is enriched at the RNA chaperone Hfq. Both ends of the Hfq-associated transcript were mapped by circular RACE and specific binding between the transcript and Hfq was confirmed by electrophoretic mobility shift assays and structure probing. The structure probing showed an Hfq-dependent restructuring of the RrB transcript consistent with a role for Hfq in mediating a change from the suboptimal structure dictated by co-transcriptional, sequential folding to the thermodynamically favorable structure which minimizes the free energy of the molecule. Finally, the conservation of both sequence and structure of the RrA, RrB, RrF transcripts among members of the Enterobacteriaceae and the fact that interaction with Hfq stabilizes the transcripts, suggest a conserved regulatory role for these RNAs.
Introduction

The family of tRNA-linked repeats

Small RNAs (sRNA) have been established as important regulators of bacterial virulence and stress responses (Sobrero and Valverde, 2012; Storz et al., 2011; Vogel and Luisi, 2011). Due to our long-standing interest in tRNA and post-transcriptional regulation, we took notice of a number of defined transcripts from the intergenic regions of some tRNA operons in the transcriptomic data published by Raghavan et al. (Raghavan et al., 2011), which corresponded to members of the annotated family of tRNA-linked repeats (TLR). The TLRs is a class of 22 intergenic repeats in Escherichia coli only described as Rho-dependent terminators (Bosl and Kersten, 1991; Egan and Landy, 1978; Komine et al., 1990; Kuchino et al., 1985; Kupper et al., 1978; Madden and Landy, 1989; Rudd, 1999). The first TLRs were discovered in 1978 (Egan and Landy, 1978). These TLRs were known as the 3.14 repeat and were located in the operon of tyrTV tRNA. Egan and Landy noticed that each repeat contained 19 nt of sequence identity to the 3’ end of the mature tyr-tRNA encoded immediately upstream. The 18-19 nt identity to the 3’ end of a mature tRNA is the hallmark of all TLRs situated in tRNA operons (Komine et al., 1990), and this hallmark led to the suggestion that the TLRs may be remnants of cryptic prophage or plasmid integration and excision events (Rudd, 1999). However, homologous TLR sequences are found in other members of the Enterobacteriaceae, potentially indicating a conserved role for the sequences (see Results).

The TLR naming convention used in this article is adapted from Ecogene (Zhou and Rudd, 2013). The first letter is “r” for repeat, the second letter is the first letter of the tRNA or 5S rRNA encoded upstream of the TLR and the last letter is the fourth letter of the tRNA or 5S rRNA gene, capitalized. Using this nomenclature the TLR located downstream of the 5S gene rrfB is rrB, and in case it is expressed the transcript is named RrB.

The only TLR transcript that has been detected previously was a transcript corresponding to rtT from the 3.14 repeat, although two different results have been reported for the location of the 5’ and 3’ ends of the transcript (Bosl and Kersten, 1991; Li and Deutscher, 2002).

Three of the annotated TLRs (rrA, rrB, rrF) are located downstream of the genes rrfA, rrfB, and rrfF encoding ribosomal 5S RNA, and differ only by a single nt. We collectively name these TLRs rrABF. The rrABF are sandwiched between the two Rho-independent terminators T1 and T2 (Fig. 1a). The efficiency of terminator T1 and T2 has been measured to 87% and 100%,
respectively (Orosz et al., 1991). Therefore transcripts that depend on T1 read-through are expected to be transcribed at a level around 13% of the transcription of *rrfA*, *rrfB*, *rrfF*. The promotors of the ribosomal RNA operons are among the strongest in *E. coli* and their activity is closely linked to the growth rate of the cell (Schneider et al., 2003). Accordingly, if RrABF do not have transcriptional start sites of their own, their expression is expected to be at the highest levels during rapid growth. Examination of the RNA sequencing data published by Raghavan and co-workers (Raghavan et al., 2011) confirms the existence of transcripts from the *rrABF* area although these transcripts appear to terminate 19 nt before the annotated 3’ ends (Fig. 1a). If the transcription of RrABF was regulated independently of ribosomal RNA, an rRNA independent transcriptional start site would be located upstream of the *rrABF*. Thomason et. al. have sequenced the 5’ ends of RNAs containing three 5’-end phosphate groups in a search for transcriptional start sites in *E. coli*. The RNA was harvested under three different conditions: stationary phase in LB medium, and exponential growth in LB and M63 media (Thomason et al., 2015). They found transcriptional start sites located at the first nt of the 5S gene in the operons of several rRNA genes under all three growth conditions (Fig. 1a). In addition, they mapped a transcriptional start site downstream of the 5S gene in the terminator T1 sequence just upstream of the annotated RrA during exponential growth in M63 media (Fig. 1a). The presence of a putative transcription start site just upstream of RrA opens the possibility for rRNA-independent expression of the TLRs of the *rrn* operons.

#### Figure 1. a: Sequence of the *rrfB* gene and its downstream region. The mature 5S RNA is indicated in the sequence in italics (1-120). Terminator T1 and T2 are indicated by underscoring the palindromic parts of their sequences. Between T1 and T2 is the annotated sequence of *rrB* indicated by bold letters (180-283) and partly overlapping is the RNA sequence we found bound to Hfq indicated by italics (163-276; *rrB*<sub>hom</sub>). The red sequences (108-125, 276-293) is a direct repeat. Each repeat sequence can base pair with the sequence in blue (1-10; see Fig.
One important player in sRNA-mediated regulation is the abundant and highly conserved 11 kDa Lsm-like protein Hfq (Mikulecky, 2004; Sobrero and Valverde, 2012). Hfq mediates sRNA-mRNA interaction by facilitating short and imperfect base-pairing between the sRNA and the mRNA target, leading to post-transcriptional regulation of protein expression (Ikeda et al., 2011; Sledjeski et al., 2001; Vogel and Luisi, 2011). Hfq interaction with antisense RNA (asRNA) and DNA has also been reported (Opdyke et al., 2004; Ross et al., 2013; Updegrove et al., 2010). Of particular interest for this study, Hfq has recently been shown to be implicated in the regulation of several sRNAs by facilitating their interaction with external or internal transcribed spacers of tRNA operons in *E. coli* (Lalaouna et al., 2015).

During exponential growth, Hfq is highly expressed. The prevalent conformation of Hfq is hexameric (Hfq₆), forming a toroidal doughnut-like structure (Kajitani et al., 1994; Panja and Woodson, 2012; Sobrero and Valverde, 2012). One side of the torus is termed the proximal face, the other side the distal face. (Sobrero and Valverde, 2012). The proximal face generally binds A/U-rich single-stranded regions located adjacent to a stem-loop structure, motifs typically found in sRNAs (Mikulecky, 2004). By contrast, the distal face seems to preferentially bind the ARN or AAN repeat often found in mRNAs (where A is an adenine, R is a purine and N can be any nucleotide (Link et al., 2009; Robinson et al., 2014; Schu et al., 2015b)). The third surface termed the lateral or the rim face connects the proximal and distal face. The lateral face may be important for binding internal U-rich sequences of RNAs (Sauer et al., 2012; Schu et al., 2015b). In *E. coli*, the unstructured and highly flexible C-terminal tail of Hfq₆ has been shown to interact with mRNA (Robinson et al., 2014) but might also play an important role in Hfq-DNA interaction (Updegrove et al., 2010).

Here we focus on the TLR sequences located downstream of the genes encoding 5S rRNA. We show that indeed, TLR sequences are expressed, processed to defined lengths, and highly enriched in the RNA that binds to Hfq. In many respects, these transcripts behave like our reference Hfq-dependent sRNA OxyS. We also find that the TLRs are conserved among several Enterobacteria with respect to both sequence and structure, and we therefore suggest that they function as bona fide sRNAs, most probably working in an Hfq-dependent manner.
Results

Transcripts from the *rrABF* regions are enriched at Hfq

To determine whether transcripts from the *rrABF* regions could be detected in *E. coli*, we harvested total RNA from exponentially growing wildtype cells (wt) and subjected the RNA to Northern blot analysis using a probe specific for RrABF (Figure 2a, lane 1). One distinct band was clearly visible on the Northern blot, indicating that, indeed, a transcript from one or more of the *rrABF* regions accumulates in *E. coli* during exponential growth (black arrow). Since sRNAs are often found to act in concert with Hfq, we investigated whether the TLR transcript could be detected in complex with Hfq. Chromosomally encoded Hfq was tagged at the C-terminus with a biotinylation sequence (Beckett et al., 1999), and it was confirmed that the tagged Hfq protein stimulated sRNA-mediated repression of an mRNA target to the same extent as wildtype Hfq (Supplementary Figure S1). Next, we confirmed that the RrABF transcript was present in total RNA from wt cells containing the tagged Hfq (Hfqbio) (Fig 2a, lane 2).

Since *rrABF* is expected to be co-transcribed with the *rrn* genes, which are downregulated during stringent response, we checked whether the transcript was present after amino acid starvation by harvesting total RNA from cells exposed to 10 minutes of valine-induced isoleucine starvation (Leavitt and Umbarger, 1962) (Fig 2a, lane 3). No band could be detected in the sample of total RNA from starved cells, indicating that the transcript is short lived and that transcription of RrABF is downregulated during amino acid starvation like transcription of the ribosomal RNA. Another blot, containing more RNA (15 ug of total RNA per lane) also didn’t reveal a RrABF-specific band in total RNA from the starved cells, confirming this observation (data not shown).

Interestingly, when we harvested RNA co-precipitated with Hfq from exponentially growing cells containing Hfqbio, two additional bands were detected by the RrABF -specific probe (Fig. 2a, lane 5). These bands were also undetectable after 10 minutes of isoleucine starvation (Fig. 2a, lane 6). The longest enriched band in lane 5 (open arrow) corresponds to the full-length transcript containing 5S (see also Supplementary Fig. S2, lane 5). The shortest enriched band (gray arrow) corresponds to a processed version of RrABF, which we will refer to as RrABF_short. The RrABF_short -enrichment at Hfq was confirmed in three independent experiments (Fig 2a and data not shown). No RrABF-specific band were detected in a control precipitation experiment from a strain with untagged Hfq (Fig 2a, lane 4), confirming that RrABF is not
unspecifically enriched by the precipitation procedure. Further, it was controlled that intact RNA was present in all lanes by probing for tRNA^{His} (Fig. 2b).

These results open the possibility that the RrABF_{short} transcript may have an sRNA-like function, since it is highly enriched on Hfq. Alternatively, the Hfq-enrichment of specific RNA species could simply be an indication of Hfq involvement in the degradation of the TLR transcripts by the Hfq-associated degradosome (Carpousis, 2007; Ikeda et al., 2011), maybe as part of pre-rRNA processing.

Figure 2: A transcript from the rrABF region can be detected in exponentially growing E. coli and a processed transcript RrABF_{short} is enriched on Hfq. a: Northern blot of a denaturing 10% acrylamide-urea gel containing RNA from samples as indicated above each lane. The probe (Extension rrB #1) recognizes RrABF (see Figure 5 for location of probe complementarity and Materials and Methods for probe sequence). Strains are wt (MG1655) or its isogenic derivative tagged with a biotinylation site in the C-term of Hfq (Hfqbio). All strains contain uninduced pBirA. RNA was harvested during exponential growth in minimal glucose medium or 10 min into Ile starvation as indicated above each lane. The black arrow indicates the RNA species that is not enriched at Hfq, while gray arrows (open and closed) mark Hfq-enriched RNA species. The length of 5S RNA and tRNA^{His} are indicated on the left for comparison. These size markers were obtained by re-probing the membrane for these RNA species. In same
membrane probed for tRNA^His demonstrates that intact RNA was loaded in all the indicated lanes. We note that the same volume of RNA, rather than the same amount of RNA, was loaded in each lane. This way, the RNA loaded in each lane corresponds roughly to RNA extracted from the same number of cells. Naturally, the RNA concentration in samples containing only Hfq-co-precipitated RNA (lanes 5,6) were much lower than in samples containing total cellular RNA (lanes 1-3), and the RNA concentration in the control precipitation with untagged Hfq was lower still (lane 4).

**Hfq enhances the stability of RrABF<sup>short</sup>**

A hallmark of many regulatory sRNAs in *E. coli* is that they are stabilized upon interaction with Hfq<sub>6</sub>. That is, if the sRNA-Hfq interaction is disrupted, the sRNA becomes less stable (Schu et al., 2015a; Viegas et al., 2011). If, on the other hand, an RNA-Hfq<sub>6</sub> interaction is due to Hfq<sub>6</sub>-mediated degradation of the RNA, then the RNA would be expected to be more stable in a Δhfq strain than in the wildtype. Thus, we compared the stability of RrABF transcripts in MG1655 and in an isogenic Δhfq derivative. Figure 3 shows the fraction of RrABF<sup>short</sup>, and the longer transcripts RrABF 1 and RrABF 2 (see Figure S3 for the primary data and the location of bands corresponding to each transcript) remaining in total RNA from cells harvested at the indicated times before and after addition of rifampicin, which blocks transcription initiation and therefore halts the synthesis of RrABF.
Figure 3 shows that RrABF\textsuperscript{short} is more stable in an hfq mutant, indicating that the interaction of RrABF\textsuperscript{short} with Hfq protects RrABF\textsuperscript{short} from degradation. This finding supports the hypothesis that interaction of RrABF\textsuperscript{short} with Hfq is not due to Hfq\textsuperscript{6}-mediated degradation of RrABF\textsuperscript{short} RNA, but rather that Hfq participates in stabilizing
RrABF\textsuperscript{short}, as is the case for other regulatory sRNAs (Schu et al., 2015a). We did not detect any difference in the halflives of RrABF1 or RrABF2 between the Δhfq strain and the wildtype strain, underlining that Hfq specifically affects the enriched RrABF\textsuperscript{short} transcript. The RrABF\textsuperscript{short}-Hfq interaction could also be due to a role for Hfq in processing the upstream 5S RNA from the \textit{rrnB} transcript. If so, we would expect that unprocessed fragments of 5S should accumulate in the Δhfq strain, at least at the three time points before rifampicin is added. We used a probe recognizing 5S RNA to look for such unprocessed fragments, but could not detect any (data not shown).

Together, these experiments support that the RrABF\textsuperscript{short}-Hfq interaction does not result from a role of Hfq in 5S processing or RrABF degradation.

**Mapping the ends of the RrABF transcripts.**

To further characterize the Hfq-RrABF interaction, we performed a series of electrophoretic mobility shift assays (EMSA) and structure probing experiments. In order to make an \textit{in vitro} transcript of RrABF\textsuperscript{short} for use in these experiments, it was first necessary to determine the precise 3’ and 5’ ends of RrABF\textsuperscript{short}. We first sought to map the most abundant transcript from total RNA using the S1 nuclease protection assay (Berk and Sharp, 1977). Briefly, a [\textgamma-\textsuperscript{32}P]ATP end-labeled DNA oligo antisense to the TLR area of interest was hybridized to total RNA and single stranded overhangs were removed by addition of S1 nuclease. The resulting fragments were visualized on a denaturing polyacrylamide gel. However, in our hands this assay was not sensitive enough to detect the native cellular levels of RrABF. Instead, total RNA from strains harboring plasmids overexpressing the 5S gene \textit{rrnB} along with \textit{rrB} was used (Figure 4abc). In this case, a 5’end was detected 3 nt downstream of the mature 3’end of the 5S transcript. This cleavage site is already described as the initial RNase E processing site of the pre-5S (Li and Deutscher, 1995; Roy et al., 1983). A 3’end was detected 13 nt downstream of the annotated 3’end. This 3’end was verified using two different probes. These results predict that the most abundant RrABF transcript in total RNA has a length of 173 nt which is in accordance with our observations from Northern blots, using probes specific for RrABF (See Figure 2a black arrow and Fig. S2). We also carried out S1 analysis with probes antisense to the area between the two observed ends to detect any alternative transcript ends, but could not detect any (Fig. 4d and data not shown).

As seen in Figure 2a, the RrABF\textsuperscript{short} fragment which is enriched on Hfq is shorter than the most abundant fragment found in the samples with total RNA. To determine the ends of the Hfq-enriched RrABF\textsuperscript{short} fragment we used the method of circular RACE (McGrath, 2011). Briefly,
RNA purified from the Hfq co-precipitation experiment was circularized using RNA ligase, reverse transcribed using random hexameric primers, PCR amplified twice using nested sets of rrABF-specific primers, and subjected to deep sequencing on the Illumina platform. Identification of the site of circularization should then reveal both ends of the transcript. Circular RACE yielded a reasonably homogeneous set of sequences, of which the ten most abundant are visualized in Figure 4e. It is evident that the majority of reads predicts the Hfq-associated RrABF\textsuperscript{short} transcript to be 115 nt in length, which is in accordance with our observations from Northern blots, both in terms of transcript length and in terms of location of the ends of the transcript (Fig. 1, 2a & S2).
Figure 4: Determination of 5’ and 3’ ends of RrB by S1 nuclease analysis and of RrB\textsuperscript{short} by circular RACE. In
the S1 nuclease mapping analysis (panels a, b, and c) the \textsuperscript{32}P-labelled probe DNA was visualized by
autoradiography of 10% poly-acrylamide sequencing gels. a: mapping of the 5’-end using probe rrb68-8upstrm
(showed in d) and total RNA from IPTG induced TSS23 (MG1655 + pTSS2) over-expressing truncated \textsuperscript{7}rfB and
rrB. The S1-protected fragment (lane 4) is 13 nt shorter than the untreated probe (lane 1). The exact number of
nucleotides removed was determined by loading two labelled oligos (lane 2 and 3, probe rrb68-8upstrrn truncated
by 14 and 18 nt respectively). b: probe rrb3\textsuperscript{3L}-map40-20 was hybridized to total RNA from IPTG induced
TSS22 (△22TLR + pTSS1) over expressing rrfB and rrB. The S1 protected band was shortened by 7 nt (lane 3)
compared to the untreated probe (lane 1). Lane 2 shows a similar experiment using total RNA from the IPTG
induced strain TSS20 (△22TLR harbouring empty vector). c: Probe rrb3\textsuperscript{3L}-map30-35 was hybridized to total
RNA from TSS21 (MG1655 + pTSS1) over expressing rrfB and rrB. The S1 protected band was shortened by 22
nt (lane 3) compared to the untreated probe (lane 1). Lane 2 shows a similar experiment using total RNA from the
strain TSS20 (△22TLR harbouring empty plasmid). d: Map of the probes used for the experiments shown in a,
b, and c. Top line represents the genomic map; rrfB is shown in red, the intergenic sequences in grey and the
annotated version of rrB in black. Probe sequences are presented as green lines and probes shown in transparent
colors were also used for mapping but detected no ends. The vertical broken lines denote the two ends detected in
this study. e: Circular RACE mapping of the RrB sequences co-precipitated with Hfq. The ten most abundant
RrB\textsuperscript{short} transcripts detected by circular RACE and deep sequencing are aligned to the genomic sequence. The size
of the characters correlates with their relative abundance, which is also stated as a percentage of total merged
reads (n=4722). Bold characters highlight the two terminators T1 and T2 and the direct repeat, which is also
found at the 3’ end of the mature 5S RNA is highlighted in red (see Figure 1).

Characterization of the RrB\textsuperscript{short}-Hfq interaction

To learn about the requirements for binding of RrB\textsuperscript{short} to purified Hfq\textsubscript{6}, we made a series of
RrB transcripts and used these transcripts in an electrophoresis mobility shift assay (EMSA).
The series of the RrB transcripts vary with respect to their 5’ and 3’ ends, and includes the 115
nt transcript we mapped by circular RACE, as well as the annotated RrB transcript (Rudd,
1999). Figure 5b shows the location of the ends of the transcripts tested; the 5’ ends are
marked by the numbers 1, 2, 3, and the 3’ ends are marked by A, B, C. The transcript 2B is the
transcript that mimics the Hfq\textsubscript{6} associated transcript, RrAB\textsuperscript{Fshort} except that it contains two
additional Gs at the 5’ end as a consequence of the sequence requirements of the T7-RNA
polymerase promoter used for the \textit{in vitro} transcription reaction. One of these Gs is present in
the genomic sequence already, the other is not.
Figure 5a summarizes the EMSA data of the seven versions of RrB tested for binding to Hfq6. The primary data are shown in Supplementary Figure S4a. The structures of the transcripts as suggested by the algorithm Mfold (Zuker, 2003) are shown in Figure S4b. A transcript corresponding to the OxyS sRNA, which has previously been shown to have a Kd of 5.3 nM for the lower shift and 53 nM for the super shift, was used as a positive control (Henderson et al., 2013b). Super shifts are a common feature in other Hfq6-sRNA EMSA experiments (Henderson et al., 2013b; Sun and Wartell, 2006), and was also seen in all our experiments except one.
G's at the 5' end, transcripts 2B and 2C contained one additional G (not belonging to the sequence) at the 5' end. The total sequence corresponds to nt 121 – 283 from Fig. 1 including the same annotation as in Fig. 1.

(Suppl. Fig. S4a). The super shifted complex between Hfq6 and OxyS was found to consist of 2 Hfq6 and 1 OxyS in the experiments of Henderson et al., therefore we find it very likely that a complex with the same ratio of 2 Hfq6 to 1 transcript is the origin of the super shifted complex in our experiments.

According to the results shown in Figure 5a, the transcripts bind to Hfq with different affinities, which rank as follows (highest to lowest affinity): 1C, 1A, 1B, 2B (alias RrB\textsuperscript{short}), 2C, 3B and 3A. From this order two rules can be deduced: First, a longer transcript does not simply result in stronger binding to Hfq\textsubscript{6}, since 1C binds better than 1A, and 3B binds better than 3A. Second, the location of the 5' end seems to be a more important determinant of the affinity than the 3' end, since the transcripts rank in affinity according to the location of their 5' end. The best binding transcript, 1C, binds Hfq\textsubscript{6} at a strength very similar to OxyS, and the 3A transcript, which corresponds to the annotated RrB (Rudd, 1999), is the transcript that binds with the lowest affinity. A comparison of the sequence of the five transcripts that bind Hfq\textsubscript{6} with the highest affinity reveal that they all are rich in Us at their very 5' end, perhaps indicating that the Us in the 5' end of the transcript could be important for Hfq binding. We note however that we observed no change in structure at the 5' end of RrB\textsuperscript{short} upon Hfq\textsubscript{6} binding in our structure probing assay (see below; Fig. 6a).

Another possibility is that the structures of the best binding transcripts are more favorable for Hfq\textsubscript{6} interaction, exposing a possible binding site better. Interestingly, the transcript 2B, mimicking the transcript RrABF\textsuperscript{short}, which we found to be associated with Hfq\textsubscript{6} in the co-precipitation experiments, displays intermediate binding in our assay. While this could be interpreted as an indication that a longer transcript binds Hfq\textsubscript{6} with high affinity and is 5' processed to yield the RrB\textsuperscript{short} in a Hfq\textsubscript{6}-associated manner, we remark that the biologically relevant transcript is not necessarily the one that binds Hfq with the highest affinity.
Figure 6: Hfq induces a specific structural change in the 2B transcript. a: An autoradiogram of an 8% acrylamide gel from the structure probing experiment with transcript 2B using Pb²⁺ and RNaseI as probing agents. The transcript was incubated with 0, 1.5 or 3 μM of Hfq, before adding the cleavage agent. C1 indicates the transcript in Pb²⁺ buffer without addition of Pb²⁺, C2 indicates the transcript in RNaseI buffer without addition of RNaseI. T1 indicates the heated transcript treated with RNase T1 and OH is the transcript heated in alkaline buffer. Gray bars on the left side of the gel mark areas of Pb²⁺ protection, Pb²⁺ enhancement and RNaseI enhancement (see text for details). The nt number with respect to the 5’ end is indicated at the right side of the gel. Intense bands at position 39 and 43 in all lanes can probably be attributed to background hydrolysis of the purified transcript before the analysis. b: Structure of 2B as predicted by Mfold (Zuker, 2003), folded with

\[ \Delta G = -39.5 \text{ kcal/mol} \]

\[ \Delta G = -44.5 \text{kcal/mol} \]

- Pb²⁺ protected after Hfq incubation
- Pb²⁺ enhanced cleavage after Hfq incubation
- RNaseI enhanced cleavage after Hfq6 incubation
Hfq6 induces a structural change in the RrBshort transcript

To further examine the interaction between RrBshort and Hfq6 we performed in vitro structure probing of the 2B transcript in the presence of varying amounts of the protein (Fig. 6). The probing agent we used for dsRNA was RNaseIII which digests long stems with decreasing efficiency as a function of number of bulges in the structure (Hjalt and Wagner, 1995). To identify single-stranded regions of RNA we used Pb2+ (Ciesiolka et al., 1998). The most notable Hfq6-induced change we found in the Pb2+-treated samples were the apparent single-stranded structure at nt 63-64, 66-68. Already at the low concentration of Hfq6, this sequence displayed less cleavage by Pb2+ (Fig. 6a). At the same low concentration of Hfq6, nt 53-56 show enhanced Pb2+ cleavage. Finally, nt 48-50 displays slightly enhanced RNaseIII cleavage after incubation with Hfq6 (Fig. 6a).

To interpret these data, we used the algorithm Mfold (Zuker, 2003) to predict the most thermodynamically favorable RrBshort structure with (Figure 6b) and without (Figure 6c) the constraint of a single stranded region from nt 63-68 (open red triangles). The differences between the two predicted structures in Figure 6b and 6c correspond very well to the experimentally observed differences between the absence and presence of Hfq6 in the structure probing assay (Figure 6a). Specifically, compared to the structure in Fig. 6b, the structure in Fig. 6c would be predicted to show increased protection of nt 63-68 for the Pb2+ ion (open triangles), more Pb2+-induced cleavage of the triple Gs at nt 53-56 that have moved from a stem to a loop (closed red triangles), and increased cleavage by RNaseIII of the stem involving nt 48-50 as it contains no bulges (red stars). For these reasons, it is meaningful to consider Figure 6b similar to the RrBshort structure before Hfq binding and Figure 6c similar to the structure after Hfq binding.

The structure shown in Fig. 6b is consistent with sequential folding of the growing ribonucleotide chain as it leaves the RNA polymerase, which would favor base pairing of the region around nt 50 as soon as it emerges from RNA polymerase with the region around nt 40 (Fig. 6b). Co-transcriptional folding is known to result in kinetical trapping of the full length transcripts in alternative conformations which are not the most thermodynamically favorable (Herschlag, 1995; Kramer and Mills, 1981; Meyer and Miklós, 2004). We suggest that after
incubation of the transcript with Hfq<sub>6</sub>, a structural change is induced, disrupting the long bulged stem and leading to formation of a new stem involving base paring between the region around nt 50 and the region around nt 60 (Fig. 6c), which is predicted by the algorithm Mfold to be the most favorable structure possible in terms of free energy. An alternative interpretation would be that Hfq<sub>6</sub> could bind and form a foot-print at the sequence motive UUUGAA at nt 63-68. Such an interaction would be predicted to result in protection of nt 63-68 from Pb<sup>2+</sup> cleavage. On the other hand, simple binding of Hfq would not explain the enhanced Pb<sup>2+</sup> cleavage at nt 53-56 after Hfq<sub>6</sub> addition, and therefore we favor the interpretation that Hfq<sub>6</sub> interaction allows refolding into the structure shown in Figure 6c and that we see no defined foot-print of Hfq<sub>6</sub> in our assay.

The structure and sequence of RrABF<sub>short</sub> is conserved among many members of the Enterobacteriaceae

To determine the degree of conservation of the sequence and genomic location of rrrB in other species, we performed a BLAST homology search (Coordinators, 2013) using the E. coli MG1655 sequence encoding rrfB, T1, rrB and T2 (Fig.1). Inclusion of the highly conserved rrfB sequence was necessary to avoid too much noise due to the repetitive nature of the T1-rrrB-T2 sequences in the NCBI database. The hits selected from the search contained homology from the start of rrfB to the end of the annotated rrB (Zhou and Rudd, 2013) and only the hit with the highest homology score from each species was selected. Both the sequence and the location of the rrB downstream of the gene encoding 5S was found to be conserved in 31 species from 13 genera, all belonging to the family of Enterobacteriaceae. The selected 31 hits can be sub-divided into 3 groups based on the 16S rRNA homology (Pruesse et al., 2007); the closely related: Salmonella, Shigella, Klebsiella, Escherichia, Pectobacterium, Brenneria; the more distant: Citrobacter, Enterobacter, Provedencia, Raoultella; and the most distant: Cedecea, Serratia and Hafnia. The location of the T2 terminator downstream of the rrB is conserved in 18 species from the genera; Salmonella, Shigella, Klebsiella, Enterobacter, Citrobacter, Provedencia and in one of the species from the genus of Serratia (Fig. S6a). Having found this homology, we compared sequence versus structure conservation. For that purpose we used the locARNA consensus structure prediction algorithm (Amit et al., 2012; Schmiedl et al., 2012; Will et al., 2012), and since 30 sequences is the maximum for locARNA, the most distant species, Serratia plymuthia, was left out of the analysis. Figure 7a shows the locARNA consensus structure for the 30 species.
Not surprisingly, there is extended homology among the 5S sequences and the 5S structure is well conserved (intense red base pairs). It is also evident that the T1 terminator structure is conserved (intense colors), while the sequence of T1 is less important (green-blue appearance). 16 nt of the 5′ end of the RrB\(^{\text{short}}\) are folded as part of the T1 terminator in this analysis and show therefore mainly structural conservation. However, the last 78 nt of the RrB\(^{\text{short}}\) sequence, containing stem 1, 2, 3, and 4 (Fig. 7a), show a high degree of structure conservation (intense colors), and an intermediate to high fraction of conserved base pairs (green-yellow-red appearance).

Interestingly, in this structure suggested by locARNA the direct repeat encoded just downstream of the annotated rrb\(^{\text{}}\) (indicated in red in Figure 1a) also displays a high amount of sequence conservation, and is predicted to base pair with the 5′ end of 5S (gray bar in Fig. 7a, see also Figure 1b). We speculate that this sequence may be necessary for proper excision of 5S from the primary transcript (see Discussion).

When we used the locARNA algorithm on the RrB\(^{\text{short}}\) sequence alone it lead to the structure shown in Figure 7b, retaining the stem 1, 2, and 4 and thus displaying almost the same degree of structural conservation as the longer sequence shown in Figure 7a. It is very compelling that the three major stem loops suggested by structure conservation are identical to the three major stem loops found in the structure probing analysis of RrB\(^{\text{short}}\) after Hfq binding, and that the three conserved A-U pairs in stem 1 (Fig. 7b) are formed after Hfq-assisted folding (Fig. 6bc).
Figure 7. LocARNA predicted consensus structure displays both structure and sequence conservation at RrB\textsuperscript{short}. The structure is based on 30 alignments found in a BLAST search (see Materials and Methods). The location of the 5S (gray) T1 (green) RrB\textsuperscript{short} (red) and T2 (blue) sequences are marked by colored areas on top of the locARNA structure. The base pairs of the locARNA structure are colored according to the level of both structure and sequence conservation; red means that the base pair of the predicted structure is found to consist of only one type of base pair (e.g., C-G), yellow means that the base pair of the predicted structure is found to consist of either of two types of compatible base pairs (e.g., C-G or U-G), etc. (color code inserted). The saturation level of the colors decreases according to the number of incompatible base pairs that must be formed to obtain the predicted structure. Full color saturation means that a compatible base pair is found at the position in all 30 structures. The compatible types of base pairs are: C-G, G-C, A-U, U-A, G-U and U-G. a) LocARNA consensus structure of the full-length sequences, from the 5' end of 5S to the 3' end of T2. a: LocARNA consensus structure of the \textit{rrfB-T1-rrB-T2} region. b: LocARNA consensus structure of RrB\textsuperscript{short}.
Discussion

RrABF is present in vivo and is stabilized by Hfq

The 22 homologous sequences in the *E. coli* K12 genome annotated as TLRs caught our attention because we found that they were conserved among species of the family Enterobacteriaceae. 19 of these TLRs overlap with what has recently been termed External or Internal Transcribed Spacers (ETS or ITS) of tRNA transcripts. Some of these have been suggested to play important roles in regulatory networks of gene expression, although the ETS’s that have been investigated experimentally until now do not have any overlap with TLR sequences (Lalaouna et al., 2015).

Here, we examine the remaining three TLRs, or 3’ETS’s, of ribosomal operons *rrnA/B/D*. Transcription in three of the seven *rrn* operons in *E. coli* is terminated by double termination signals T1 and T2 and contain each a TLR sequence in between the two terminators (Fig. 1). The three TLR sequences are practically identical (Fig. S7). We showed here that defined transcripts are expressed from these sequences (Fig. 2, 4, and S2). Specifically, we found a 173 nt long version, RrABF, probably processed out of the primary transcript by RNase E (discussed below), and a 115 nt further processed version RrABF\textsuperscript{short}. Both versions co-precipitated with Hfq, but RrABF\textsuperscript{short} was particularly enriched (Fig. 2). We found that both RNAs are relatively short lived after halt of transcription by addition of rifampicin or by inducing high ppGpp levels provoked by amino acid starvation (Fig. 2 and 3). We also found that the stability of RrABF\textsuperscript{short} was enhanced by the presence of functional Hfq in contrast to the longer RrABF whose half life seemed unaffected by the presence or absence of Hfq (Fig. 3). The stabilization of RrABF\textsuperscript{short} by binding to Hfq qualifies this part of the transcript to be the functional part since many sRNAs with known functions have been reported to exhibit increased half lives upon interaction with Hfq (Schu et al., 2015b; Zhang et al., 2013). This, together with the observation that versions of the RrABF\textsuperscript{short} transcript with extended 5’-ends bind Hfq with higher affinity in vitro (Fig. 5), could suggest that RrABF may be the precursor of RrABF\textsuperscript{short}, and that the processing could occur on Hfq. However, more experiments are needed to validate this hypothesis.

Processing of 5S and RrABF from the primary transcript

As described in the introduction, the 19 TLRs located in the operons of tRNA have 18-19 nt of sequence identity to the mature tRNA. A closer look at the sequences of the *rrfABF* and their
downstream TLR sequences revealed that also the TLRs of the *rrnA/B/D* operons share the same identity characteristics. Specifically, if the annotated *rrABF* sequences are extended five base pairs past the annotated 3’ end, they contain 18 bp of sequence identity to the 3’ region of *rrfA/B/F*. (See Fig.1 red sequences, and alignment in Fig. S7).

During transcription of the *rrn* operons, the precursors of 23S, 16S, and 5S are enzymatically relieved by RNasellII (Li et al., 1999a, b). In that process, the *rrfA/B/F* transcript is embedded in a 9S precursor and further maturation of 5S is performed by RNaseE (Roy et al., 1983). Roy et al. found evidence that RNaseE cleavage results in a still immature 3’ end of 5S, which is located at the bottom of the stem formed with the mature 5’ end of 5S (Fig. 1 and S5a). However, when there is read-through of the T1 terminator the exact same stem of base pairs could form between the mature 5’ end of 5S and the repeated sequence at the 3’ end of RrABF (Fig. 1 and S5b). This architecture of the 3’ end of the *rrnABD* operons strongly suggests that RNaseE cleaves at the 3’ end of RrABF, thereby separating the RrABF transcript from terminator T2, in the same way that terminator T1 is separated from 5S.

S1 mapping of the RrABF RNA (Fig. 4d) showed that the 5’ end of RrABF is identical to the RNase E processing site within the first direct repeat sequence reported by Roy and colleagues (Roy et al., 1983), while the 3’ end of RrABF was found to be cut five nucleotides further downstream with respect to the second direct repeat sequence (Fig. 4d). This is five nucleotides removed from the sequence with identity to the upstream RNase E processing site, but the processed RrABF also has a pair of adenosine bases at the 3’ end like the immature 5S transcript (Fig 4d).

**TLR Interaction with Hfq**

After the identification of the RrABF<sup>short</sup> transcript enriched in Hfq-bound RNA we wanted to determine the strength and specificity of this interaction. We used different *in vitro* transcribed variants of the mapped RrABF<sup>short</sup> in EMSA for this purpose (Fig. 5b). To our surprise, it was not the transcript 2B, which has the highest similarity to RrB<sup>short</sup> (two 5’ G’s too long) that showed the highest affinity to Hfq. Instead, longer transcripts with extended 5’ ends showed the highest affinity (Fig. 5a). The most unifying pattern of the five best binding transcripts is that they contained a stretch of Us in their 5’ end and that the longer transcripts contained an additional number of uridine bases (Fig. 5b). We propose that the Us at the 5’ end contributes to the affinity towards Hfq<sub>6</sub>, since it is well documented that Hfq<sub>6</sub> binds U-rich sequences (Brescia et al., 2003; Ishikawa et al., 2012; Mikulecky, 2004). However, we cannot exclude that the two extra G nucleotides at the 5’ end of the *in vitro* transcript 2B may
interfere with the binding to Hfq. If so, they could mask an affinity between Hfq and the \textit{in vivo} RrABF\textsuperscript{short} sequence that could exceed that between 2B and Hfq.

We used the sRNA OxyS as a positive control to confirm that our Hfq pull down assay was functional (Fig. S2). In our EMSA we also used OxyS RNA as a positive control. Here, we found that the 2B transcript had 50% bound Hfq at ~400 nM protein concentration and 50% of the 1A, 1C and the OxyS RNA are bound at 100-200 nM Hfq. These relatively identical affinities confirm that our transcripts bind Hfq with a binding strength in the range of already known Hfq-interacting sRNAs. Henderson and co-workers (Henderson et al., 2013a) reported a significantly higher affinity between OxyS and Hfq in their EMSA assay. The ionic strength and the pH for their and our experiments are similar. However, we included carrier RNA in our assay, and we assume that the apparent discrepancies in affinities are due to this fact. Since we do not yet know how and where the RrABF\textsuperscript{short} is processed to its final length, it is possible to suggest that it takes place at Hfq. If so, the lower binding strength of the final product could be biologically relevant for the RNA to be able to escape Hfq binding to accomplish its function.

**Hfq-assisted folding of RrABF\textsuperscript{short}**

In our study of the interaction between Hfq and the RrABF\textsuperscript{short} RNA we wanted to examine the structural changes in the RNA upon Hfq binding and also to see if we could detect protection from the reactive components by a “foot print” of Hfq on the interacting parts of the RNA molecule. The structure probing experiments revealed rather few changes in the probing pattern after Hfq binding (Fig. 6a). However, the results were all consistent with a model of Hfq-assisted refolding of the RNA, from a structure with a ΔG of -39.5 kCal/mol for folding (Fig. 6b) to the structure which is suggested by the algorithm Mfold to have the lowest possible free energy, namely ΔG = -44.5 kCal/mol (Fig. 6c). We propose that the first structure (Fig. 6b) is formed by sequential folding of the RNA during its transcription, although alternative stronger structures would be possible in the finished transcript if they were free to form. This situation resembles the sequential folding of transcriptional attenuators or riboswitches, where alternative structures also can be formed upon interaction with other binders like e.g. ribosomes in the Trp attenuator (Landick, 1996), and has been suggested to be a common problem in RNA folding that can be resolved by RNA chaperones such as Hfq (Herschlag, 1995; Moll et al., 2003).

This interpretation of our structure probing results left no obvious differences in the probing pattern for an Hfq footprint (Fig. 6a), which we would have expected from our EMSA where
we found an interaction, also after refolding. Unless Hfq binds RrABF\textsuperscript{short} in a very diverse or unspecific manner, we have no explanation for this lack of a clear footprint.

Our Hfq-assisted refolding model is strongly supported by the conservation of the RrABF structure and sequences among members of the Enterobacteriacea family (Fig. 7). First of all, the overall conservation suggests a specific role for these sequences and structures. Second, importantly, the base pairs that were formed in the structure only upon addition of Hfq, namely the three A-U pairs in the bottom of stem 1 (Fig. 7b and 6b,c) are strongly conserved both in terms of structure and sequence identity (Fig. 7b, intense red color), suggesting that Hfq-assisted refolding may be necessary to arrive at the functional, conserved, structure of the RrABF\textsuperscript{short}

The structure suggested by the locARNA algorithm shows that areas without sequence conservation can have a conserved secondary structure, as demonstrated by the T1 terminator alignment (Fig 7a). In the RrABF\textsuperscript{short} alignment, there is a mixture of sequence and structure conservation, indicating that not only the structure but also the identity of some of the base pairs are important.

Sequence and structural conservation is taken as an argument for an important function of the loci. Here, we note that the structure of RrABF, as well as its genomic location downstream of the gene encoding 5S is conserved in 31 species from 13 genera all belonging to the family of Enterobacteriaceae.

**Hfq and function of the TLR family**

The cellular level of RrABF\textsuperscript{short} is determined by the stability of the RrABF\textsuperscript{short}, the expression level of the rrnA/B/D operons, and the strength of the T1 termination signal (Orosz et al., 1991) upstream of rrABF. It is reasonable to assume the terminator T1 to function independently of growth conditions while the rrn promoters are known to respond with increasing activity as a function of growth rate and to be very sensitive to the level of the alarmone ppGpp (Ryals et al., 1982b). We therefore expect the presence of a constant level of RrABF during steady state growth conditions determined by the activity of the rrn promoters, unless RrABF is regulated post-transcriptionally or can be transcribed by an alternative promoter(s). Thomason and colleagues (Thomason et al., 2015), reported the existence of 5’ tri-phosphorylated nucleotides in transcripts at the very first nucleotide of the 5S RNA and again in a transcript beginning inside the T1 terminator sequence (Fig. 1) but experiments are needed to determine the activity of these promoters and whether they are regulated.
Hfq\textsubscript{6} has during the last 20 years been shown to be implicated in a variety of different functions in bacteria and still more are turning up. Recently, Lalaouna and co-workers showed that Hfq\textsubscript{6} is involved in the interaction between the small RNA 3’ETS\textsuperscript{LeuZ}, which is part of a polycistrionic tRNA transcript, and the sRNAs RyhB/RybB. The function of the 3’ETS\textsuperscript{LeuZ} is to sequester RyhB and RybB, absorbing their transcriptional noise under conditions where the activity of the two sRNAs is not required. The level of 3’ETS\textsuperscript{LeuZ} is linked to the expression of the tRNA and the nutritional state of the cell and so seems the expression of Hfq, even though there are conflicting reports (Ali Azam et al., 1999; Carmichael et al., 1975; Ishihama et al., 2008; Kajitani et al., 1994).

The 22 TLRs found in \textit{E. coli} K12 are all located in operons expressing tRNA or rRNA and the fact that the TLRs have 18-19 nt of sequence identity to the most distal 3’ end of the immediately upstream RNA gene could suggest a function linked to rRNA/tRNA; be it regulation, modification or processing. Since there are so relatively few transcripts of RrABF compared to the amount of rRNA, such function would require that one RrABF\textsuperscript{short} can act catalytically or that only a minority of rRNA molecules needs the regulation, modification, or processing affected by RrABF\textsuperscript{short}.

An alternative explanation for the sequence identity between 3’ ends of TLRs and their upstream RNA genes could be the re-use of a motif for 3’ end processing out of the primary transcript. This seems to be the case for the RrABF RNA, but does not exclude that the repeat sequence could serve another function(s) in the mature TLR (Fig. S5).

The Hfq co-precipitation and the Hfq-dependent stability of RrABF\textsuperscript{short} suggest an Hfq\textsubscript{6}-connected function. Further, the indication of a transcriptional start site of its own opens the possibility for a more classic sRNA/asRNA-like function of RrABF\textsuperscript{short} if it is induced under an as yet unknown growth condition.

In conclusion, we believe that we have four strong arguments that the transcript RrABF\textsuperscript{short} in \textit{E.coli} MG1655 serves as a functional RNA that interacts with Hfq\textsubscript{6}: First, RrABF\textsuperscript{short} is highly enriched by co-precipitation with Hfq. Second, Hfq induces specific structural changes in the RrABF\textsuperscript{short} transcript. Third, the stability of RrABF\textsuperscript{short} \textit{in vivo} is enhanced by Hfq, and finally, the genomic location and the structure of RrABF\textsuperscript{short} upon Hfq binding is conserved among many members of the Enterobacteriacea.

**Materials and methods**

Genotype and construction of cells and plasmids
Hfq co-precipitation experiment

Cells were grown exponentially in MOPS medium (Neidhardt et al., 1974b) supplemented with; 15 mg/l chloramphenicol, 0.2% glucose, 10 mg/l uracil, 50 μM biotin, in a volume of 300 ml at 37°C for at least 10 generations. At OD₄₃₆ 0.8 the culture was divided and isoleucine starvation was induced by addition of 400 mg/l valine, the other culture was transferred to ice. Cells were pelleted and washed in medium without biotin, and re-suspended in 2ml lysis buffer (50 mM Tris-HCl pH 7.5, 50 mM NaCl, 5% glycerol). The cells were lysed by sonication, cell debris spun down (“total RNA” sample was withdrawn) and the supernatant was
transferred to fresh tubes including 300 μl equilibrated SoftLink™ Avidin Resin (Promega), left ON at slow rotation at 4°C. Next day the resin was washed 4 times in lysis buffer and RNA was harvested by phenol extraction and ethanol precipitation. The RNA was separated on a 10% acryl amid/urea/PB gel. 0.01% of total RNA and 10% of the bound RNA were loaded at gel. After separation the RNA was transferred to a nylon membrane (GE Healthcare Life Science Hybond.) by electroblotting and fixed to the membrane by UV cross-linking.

Detection of RNA

$[^{32}\text{P}]\gamma\text{ATP}$ labeling of probes was done with T4 Polynucleotide Kinase (Thermo Scientific) in buffer:125 mM Tris-HCl, 25 mM MgCl$_2$, 0.25 mM EDTA, 12.5 mM dithiothreitol, 0.25 mM spermidine. For detection of RNA the membrane was pre-hybridized in; 5 x Denhardt (Thermo Fisher Scientific), 0.5% SDS, Herring Sperm (Sigma), 0.7 M NaCl, 4 mM EDTA, 40 Na$_2$PO$_4$, followed by overnight hybridization with the specific $^{32}\text{P}$-labeled DNA oligo. The membrane was washed in; 0.3 M NaCl 30 mM Na$_3$Citrate and 0.1% SDS and placed at a phosphor screen, subsequently the screen was read by phosphor imaging using a Typhoon scanner.

S1 nuclease analysis

One pmol of $[^{32}\text{P}]\gamma\text{ATP}$-labelled probe was hybridized to 30 μg of total RNA from the strain of interest. Hybridization was done in 50% formamid, 20 mM HEPES, 0.5 mM EDTA, 0.2 M NaCl, 0.05 % (w/v) SDS and performed overnight in a thermocycler. Starting at 68 °C for 10 min the temperature was lowered to 54 °C and then decreased 1 °C every 30 min until reaching 20 °C. Digestion was performed by adding 300 μl 0.28 M NaCl, 50 mM NaAc pH 4.6, 4.5 mM ZnSO$_4$ along with 300 U/ml S1 nuclease (Thermo Fisher Scientific) and incubating at RT for the indicated times. Samples were phenol/CHCl$_3$ extracted, ethanol precipitated, size separated by electrophoresis on 7 M urea, 10% poly acrylamide sequencing gels and detected by autoradiography.

RNA preparation for S1 analysis

Cell were grown in MOPS medium supplemented with 0.2% glucose and 10 mg/l uracil, for at least 10 generations, to an OD$_{436}$ of ~ 0.7. Total RNA was isolated after 1h of IPTG induction by the hot phenol extraction method.

Circular RACE mapping

The circular RACE mapping was done as described by McGrath (McGrath, 2011), however omitting the TAP treatment: 500 ng of RNA co-purified with Hfq was circularized in 1x buffer by adding T4 RNA ligase. Reverse transcription was carried out using Super Script III RT (Thermo Fischer) and primed by random hexamer oligos. The area of interest was amplified twice by PCR with two different sets of specific primers (cRACE rrB-2 1F + cRACE rrB-2 1R and cRACE rrB-2 2F + cRACE rrB-2 2R). The PCR-library was sequenced on a Illumina Mi-seq by 300
bp paired-end sequencing. The resulting sequences were merged and subsequently listed by abundance. All reads that could not be merged was left out in this analysis.

**Transcript stability**

The “spike in” cell MAS1074: Cells were grown in MOPS medium at 37° C. When they reached OD$_{436}$ 0.1 1 mM IPTG was added, and growth continued for 3 hours until the cells had reached an OD$_{436}$ of ~0.33. MG1655 and Δ Hfq: The cell cultures were exponentially grown at 37° C for at least 10 generations in MOPS medium supplemented with; 0.2% glucose, 10 mg/l uracil . At OD$_{436}$ = 0.7, 3x15 ml aliquots were collected (time -3). At time 0; 100 μg/ml rifampicin was added and 15 ml aliquots were withdrawn at; 2.5, 5, 10, and 20 minutes, all samples were kept on ice. After sampling, 1 ml of the “spike in” cells was added to each of the collected samples. The cell mix was pelleted and re-suspended in 1.8 ml TRI-reagent (Sigma) and treated according to the protocol. The RNA was quantified by NanoDrop 1000. 8-10 μg of RNA was loaded into each well of the 10% acrylamide/urea/PB gel. After separation, the RNA was transferred to a nylon membrane (GE Healthcare Life Science Hybond,) by electroblotting and crosslinked to the membrane. See section “Detection of RNA” for further treatment.

<table>
<thead>
<tr>
<th>Probe name</th>
<th>Probe sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>pextension rrB #1</td>
<td>gattgtctcctactcaggag</td>
</tr>
<tr>
<td>rrBrev#2</td>
<td>cccacctcgccggcc</td>
</tr>
<tr>
<td>pextension rrB#6</td>
<td>tggcagtttatgcccgggct</td>
</tr>
<tr>
<td>Ss</td>
<td>acactaccatcggcgtac</td>
</tr>
<tr>
<td>OxyS</td>
<td>gagatccgcaaaaaagttcagtt</td>
</tr>
<tr>
<td>tRNA$^{54}$</td>
<td>CACGACAACTGGAATCACAATCC</td>
</tr>
<tr>
<td>cRACE rrB-2 1F</td>
<td>CCGGAGCGGATTTGACGGT</td>
</tr>
<tr>
<td>cRACE rrB-2 1R</td>
<td>CGTCCCGGGGAGGATTTGT</td>
</tr>
<tr>
<td>cRACE rrB-2 2F</td>
<td>gtctcggctcggctagatgtgtaaagacagcGGTTGACGAGCAAGCAGGGC</td>
</tr>
<tr>
<td>cRACE rrB-2 2R</td>
<td>tcgtcggcagctcagatgtgtaaagacagcGGTTGACGAGCAAGCAGGGC</td>
</tr>
<tr>
<td>rrB68-8 upstrm (S1)</td>
<td>Complementary to the intergenic region between 5S and rrB and partly to 5S</td>
</tr>
<tr>
<td></td>
<td>CAGATACAAACGAAAAGGCCCAGTCTTTCGACTGACGCTTTTCGTTTTATATGCTGCTCGCA</td>
</tr>
<tr>
<td>rrB54-8 upstrm (S1)</td>
<td>Complementary to the intergenic region between 5S and rrB, used as size marker</td>
</tr>
<tr>
<td></td>
<td>CAGATACAAACGAAAAGGCCCAGTCTTTCGACTGACGCTTTTCGTTTTATATGCTGCTCGCA</td>
</tr>
<tr>
<td>rrB50-8 upstrm (S1)</td>
<td>Complementary to the intergenic region between 5S and rrB, used as size marker</td>
</tr>
<tr>
<td></td>
<td>AGATACAAACGAAAAGGCCCAGTCTTTCGACTGACGCTTTTCGTTTTATATGCTGCTCGCA</td>
</tr>
</tbody>
</table>

Table 3: The names and the sequences of the probes used in the Northern blots and for S1 analysis.
Transcription and 5’ end labeling of transcripts

<table>
<thead>
<tr>
<th>Primer</th>
<th>Probe sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>taatacgactctata(G)GGCCTTCGTTTATCTGTTTGGTC</td>
</tr>
<tr>
<td>2</td>
<td>taatacgactctataGGTTTATCTGTTTGGTGTAACG</td>
</tr>
<tr>
<td>3</td>
<td>taatacgactctataGGTAACGCTCTCCTGAGTAGG</td>
</tr>
<tr>
<td>A</td>
<td>TGGCAGTTTATGCGGGGCT</td>
</tr>
<tr>
<td>B</td>
<td>TTATGCGGG CGTCTCG</td>
</tr>
<tr>
<td>C</td>
<td>TCTGCGCCGCCACCCCT</td>
</tr>
</tbody>
</table>

Table 4: Primers used in the in vitro experiments. The lower case letters are the sequence encoding the T7 promoter, transcription start site is marked by the bolded G. An additional G was inserted in primer 1, marked by parenthesis, for creating product 1C and 1B since no transcription was accomplished without this G, (primer 1 used to make product 1A did not contain this G).

The template was made by PCR using plasmid pTSS2 as template, see table 4 for primers. In vitro transcription was carried out according to the MegaScript kit (Ambion) protocol. The transcript was separated on an 8% acrylamide, by UV shadowing the RNA containing slab was visualized, excised and transferred to 2 M NH₄Acetate for 1.5 hour at 17°C followed by addition of phenol, and ON shaking. Next day, the transcript was phenol/chloroform extracted and ethanol precipitated. The RNA was quantified on a NanoDrop 2000. The transcript was dephosphorylated by Shrimp Alkaline Phosphatase (NEB) and ³²P labeled by T4 Polynucleotide Kinase (NEB), in 1 x T4 Polynucleotide Kinase buffer (NEB). The labeled transcript was purified according to the protocol using NucleoSpin® miRNA (Macherey Nagel).

EMSA

In total volume of 10 ul, 40 fmol 5’ end labeled transcript, 2 pmol unlabeled random RNA oligo 75 nt long (TAG-Copenhagen), was incubated with or without Hfq in buffer; 100 mM KCl, 1 mM DTT, 20 mM HEPES pH 8. The samples were incubated 20 minutes at 37°C, transferred to ice for 10 minutes and subsequently separated at a 5% non-denaturing poly acrylamide gel at 4°C. The gel was dried (Bio Rad Gel Dryer) and visualized by phosphor imaging using a Typhoon scanner.

Structure probing

All samples had a final volume of 10 ul which included; 0.1 pmol of 5’ end ³²P labeled transcript and 50 nM unlabeled tRNA (E. coli). The transcript was incubated with; 3 μM Hfq₆₅.
1.5 μM Hfq<sub>6</sub> or no Hfq<sub>6</sub> for 100 minutes at 37°C in the buffer relevant for the cleavage agent, see below.

**Pb<sup>2+</sup> probing:** 1 x Structural Probing Buffer (Ambion AM2237), Pb<sup>2+</sup> was added to a final concentration of 10 mM. After 1 minute at 37°C the reaction was reduced as described below. Control (C1) had no Hfq<sub>6</sub> and no Pb<sup>2+</sup>.  

**Short Cut RNaseIII probing:** 1x Short Cut reaction buffer included 1x Short Cut MnCl<sub>2</sub> (NEB), 0.002 U Short Cut RNaseIII (NEB). After 20 minute at 37°C the reaction was reduced. Control (C2) had no Hfq<sub>6</sub> and no Short Cut RNaseIII.  

**Control T1:** 1 x Structural Probing Buffer (Ambion AM2237), sample was incubated at 95°C for 1 minute, transferred to 37°C for 1 minute. After addition of 0.05 U RNase T1 (Ambion AM2237) the sample was incubated at 37°C for 5 minutes.  

**OH ladder:** 1 x Alkaline Hydrolysis Buffer (Ambion AM2237), sample was incubated at 95°C for 5 minute.  

**After cleavage:** The cleavage reactions were reduced by adding; 200 μl of ice-cold H<sub>2</sub>O and transfer to ice.  

**After treatment of the samples:** The samples were phenol extracted, ethanol precipitated, and resuspended in 1 volume of H<sub>2</sub>O to 1x Loading Buffer II(Ambion AM2237). The RNA was separated on an 8% acrylamide/urea/PB gel at RT. The gel was dried (Bio Rad Gel Dryer) and visualized by phosphor imaging using a Typhoon scanner.

**Hfq**

The Hfq used in all *in-vitro* experiments, was a kind gift from Boysen, A. 2009 Odense: Southern Danish University, the faculty of Natural Science. Purification of Hfq was performed as described by Møller et al. (Møller et al., 2002). Hfq was stored at -80°C in Hepes buffer pH 8 until use.

For the data in Fig. 6a: No Hfq<sub>6</sub>-related cleavages is seen at nt 39 and 43 in all samples, these bands appear independently of the cleavage agents, buffer, and Hfq. Further, the control sample C2 (containing only RNaseIII buffer and labeled transcript) shows more cleavage than C1 (containing only structure probing buffer from Ambion and labeled transcript). The enhanced cleavage in C2 is presumably because the RNaseIII buffer includes the Mn<sup>2+</sup> ion, which is known to catalyze single stranded RNA cleavages (Wrzesinski et al., 1995).  

**BLAST search and locARNA folding**

The bait was 327 nt long, from first T at *rrfB* to last T in terminator T2, see Fig. 1. These search was done by Blastn at the database NCBI chromosomes, the setting was as default, except search excluded *Escherichia coli* (Tax id 562) and was performed by using dMegablast. The hits
chosen were those containing a full match from rrfB to the end of rrB annotated in the Ecogene database, selecting only the first hit from each species, and a FASTA file was manually created containing 31 species in total. This file was used in the ClustalX complete alignment shown in Fig. S6a. Since locARNA can only fold and align 30 sequences, Serratia plymuthia was removed, this file was uploaded at locARNA using setting as default (Amit et al., 2012; Schmiedl et al., 2012; Will et al., 2012).

Author contributions

MK, SLS, and MAS wrote the manuscript; MK, BK, TS, and MAS designed experiments; MK and TS (Mapping results, Fig. 4) performed experiments; MK, SLS, and MAS conceived and designed research.

Acknowledgments

The authors thank Marit Warrer and Pilar Menéndez Gil for excellent technical assistance. We also thank Anders Boysen for the kind supply of purified Hfq. This work was supported by the Lundbeck Foundation R108-A10583 (MAS), the Danish council for Independent Research | Natural Sciences (1323-00343B) (SLS) and the Danish National Research Foundation (DNRF120) (SLS; MAS).

References


Li, Z., Pandit, S., and Deutscher, M.P. (1999a). Maturation of 23S ribosomal RNA requires the exoribonuclease RNase T. Rna 5, 139-146.

Li, Z., Pandit, S., and Deutscher, M.P. (1999b). RNase G (CafA protein) and RNase E are both required for the 5’ maturation of 16S ribosomal RNA. The EMBO Journal 18, 2878-2885.


Supporting information

Test of epitope-tagged Hfq

In order to test the function of the epitope-tagged Hfq, Hfqbio, we used two plasmids; one expressing the Hfq-dependent small RNA qrr2 from Vibrio cholerae under control of a pBAD
promotor, the other expressing a *hapR-gfp* protein fusion. If qrr2 is expressed, it downregulates the *hapR-gfp* mRNA in a Hfq-dependent manner, resulting in a loss of GFP fluorescence (Svenningsen et al., 2008). If the tagged Hfq is functional, it is expected to mediate downregulation of fluorescence after addition of arabinose. As shown in Fig. S1A, the signal from GFP is reduced after addition of arabinose in MeK24 (unmodified Hfq). This reduction is not seen in MeK23 (∆hfq, Fig. S1b), but also seen in MeK22 (Epitope-tagged Hfq, Fig. S1c). We take this as evidence for a functional tagged Hfq. Figure S1d is a growth control of the MG1655, wt (MG1655+pBirA) and the MAS927 (MG1655 Epitope-tagged Hfq) and HfqBio (MG1655 Epitope-tagged Hfq + pBirA). The cells were grown in MOPS medium containing 0.2% glucose. Growth was not disturbed by the Epitope-tagged Hfq, but the growth rate was reduced in both strains containing the plasmid pBirA.

**Figure S1:** Test of the epitope-tagged Hfq. Panels a, b, and c contain optical density and fluorescence data plotted versus time from cells carrying the epitope-tagged Hfq against the cells containing the non-tagged Hfq: A; MeK24(MG1655 pqr2 and pHapR-GFP), B; MeK23 (SLS4119 Hfq::cat pqr2 and pHapR-GFP), C; MeK22 (MG1655 Hfq-tag pqr2 and pHapR-GFP). Figure A-C has two vertical logarithmic axes; the primary axis displays OD₆₅₀ and the secondary axis fluorescence in arbitrary units. In figure D, MG1655, wt (MG1655+pBirA), MAS927(MG1655 Epitope-tagged), and HfqBio (MG1655 Epitope-tagged Hfq + pBirA) are grown in MOPS including 0.2% glucose. The cells containing pBirA are used in the co-precipitation experiments seen at Fig. 2 and Materials and Methods.
**Estimating the 5’ and the 3’ ends of RrABF short by Northern**

Based on the Northern blots, the length of RrABF short was estimated to be between 104-117 nt, since the 120 nt long 5S RNA was migrating just above the RrABF short and the 109 nt long OxyS transcript was below or co-migrating with the band for RrABF short (Fig. S2). Using the probes pextension rrB #1, rrBrev #2 and pextension rrB #6 it was possible to estimate the location of the 5’ and 3’ ends of RrABF short.

pextension rrB #1 and rrBrev #2 detected the same band while probe pextension rrB #6 detected a longer band (see Fig. S2). Assuming that the Northern probe needs at least 15 nt of complementarity to bind the RNA on the membrane, the 3’ end must then be located between the arrows in Fig. 5. By considering the possible positions of the 3’ end, the 5’ end must be located in the area between the numbers 117 and 104 in Figure S2b. On the basis of this analysis we conclude that the Northern blot analysis supports the ends mapped by circular RACE of the RNA that co-precipitated with Hfq.

**Figure S2.** a: Estimates of the location of the 3’ and 5’ ends of RrABF short by Northern blot analysis. a.

Northern blot membrane probed with five different probes as indicated under each section. The lanes contain RNA prepared and loaded as in Fig. 2 of the main text. Arrows point to the bands specific to the probe. The 5S
probe could not be fully removed and a weak 5S signal was carried over to the OxyS and Pextension rrB#1 analysis. b: Complementarity of the probes used in the Northern analysis shown in a. The depicted sequence corresponds to nt 121 – 283 from Fig. 1 and contains the same annotation as in Fig.1. The sequences recognized by the probes are indicated by horizontal lines. The arrows flank the region where the 3’ end of the RrABFshort must be located as estimated by the Northern blot analysis, while the numbers 117 and 104 flank the region where the 5’ end must be located according to the Northern blot analysis.

Stabilization primary data

Figure S3. Stability of RrABF transcripts. Two Northern blots with total RNA from Δhfq and MG1655 strains as indicated. Each lane was loaded with 8-10 μg RNA. The probe used was pextension rrB #1. Time after rifampicin addition is indicated above the lanes. An arrow marks the location of the RrABFshort. Lane C is the lane containing total RNA from the tRNA^{tric}-expressing control cells used for normalization (see Materials and Methods). Bottom panels show the same Northern blots reprobed with the probe for tRNA^{tric}. 
Primary EMSA data and how it was processed

The data from EMSA was processed in IMAGE Quant 5.2. A broadened line through each lane at the gel was made and counts in the line were measured, background subtraction was set to the lowest point. Each lane was divided into three areas: 1) One below, containing fragments, 2) One in the middle, containing only the full length labeled transcript, 3) One containing the complex of Hfq and the labeled transcript. IMAGE Quant 5.2 was adjusted to give the result for each area in percentages due to the total counts in each lane. Percentage of area 2 at 0 nM was used as a normalization factor; this factor was divided with area 2 for all the other lanes, to see how much labeled transcript remained unbound, as the concentration of Hfq increased. In order to get the fraction bound, the fraction unbound was subtracted from 100 and this number was plotted against the concentration of Hfq6.

The three 3´ ends that we chose to test in our EMSA assay were; A the 3´ end annotated in EcoGene (Rudd, 1999; Zhou and Rudd, 2013) and C the 3´ end matching the 3´ end from the sequencing results of Raghavan et al. (Raghavan et al., 2011) and the B end found by RACE. The tested 5´ ends were: the best match (closest to a stretch of G) if the 3´ end of Raghaven et al. (Raghavan et al., 2011) was correct, the end found by RACE, and the annotated end (Rudd, 1999; Zhou and Rudd, 2013).

The 3B transcript was the only RNA that did not form a super shift when binding to Hfq6. (Fig. S4a). The only difference between 2B (Rrbshort) and 3B is the additional U-rich sequence present in the 5´ end of 2B. We suggest that this 5´ end could be a high affinity binding sequence for Hfq and that the binding to 3B is at the low affinity site that forms the super shift in the other experiments. However, the same 5´ end is also missing in the 3A transcript, but from the structure suggested in Fig. S4b a double ARN motif is exposed in the extended 3´ end of the 3A RNA.

Figure S4: 2B, mimicking RrABFshort, displays intermediate affinity for Hfq. A: Primary data of OxyS and the seven versions of the RrABF transcript tested for binding to Hfq, by EMSA. The data is plotted in Fig. 6 and a description of how the data treatment is performed is written in the supplementary text. At the upper right corner the name of the transcript is written. The concentration of Hfq is indicated above each lane. A small amount of the transcript 2B (RrABFshort) seems to remain unbound, we ascribe this to some form of a transcript not influencing our assay. B: The structures of the RrB transcripts tested by EMSA as predicted by the algorithm Mfold (Zuker, 2003), the values in the parentheses are the free energies calculated by the algorithm.
Predictions of the structure of 9s RNA and the same transcript in the case of terminator T1 read-through by Mfold

Figure S5. Structures of pre-5S and downstream sequences. a: The 9S RNA. The arrow indicates where RNaseE cuts (Roy et al., 1983). b: The 9S RNA plus the RrB and terminator T2 after terminator T1 read-through. The arrow indicates the 3’end of RrABF mapped by S1 mapping. a and b: The structures are predicted by the Mfold algorithm (Zuker, 2003). The location of individual structural components are shaded in gray and their names are indicated. The blue bars indicate the location of the 5’ region of mature 5S and the red bars indicate the part of the direct repeats that can base pair to the 5S 5’ end (see also Fig. 1).

ClustalX allignment of the BLAST hits
Figure S6: ClustalX alignment of the results obtained by the BLAST homology search. a: alignment of all 31 hits found. b: alignment of the 30 sequences used in the locARNa analysis. The bars above the sequences indicate the location of the sequences of 5S, T1, RrB<sup>prot</sup>, and T2.

**Alignment of rrfA/B/F and the downstream sequences**

![Alignment of rrfA/B/F and the downstream sequences](image)

Figure S7: ClustalX alignment of rrfA/B/F (dark gray bars) and the downstream sequences containing the **major terminator T1** (light gray). The locations of the gene coding RrAB<sup>prot</sup> (black bars) and the terminator T2 (light gray bars) are indicated. Stars below the letters mark sequence identity.

**References for Supplementary Material**


END OF PAPER 2
2.2 Discussion

2.2.1 The function of TLR

Finding the RrABF bound at Hfq actually raised more questions than answered. Most importantly, what is the function of the RrABF$^{short}$? Do the three TLRs downstream from the gene $rrfA$, $rrfB$, $rrfF$ make a sub group of its own, or do the rest of the TLRAs have similar functions? Thus, we created a mutant where all of the 22 TLRs were deleted. It was named the ‘Δ22 TLR’.

2.2.2 Phenotypic characterization

Our intention was to characterize the Δ22 TLR mutant phenotypically since the Δ22 TLR mutant displayed a really interesting phenotype: when outgrown in MOPS medium that was supplemented with 0.2 % glucose, shaking over night at 37°C, it grew to an OD ~10% higher than the isogenic MG1655, and the growth rate in balanced growth was 20-40% reduced for the Δ22 TLR as compared to the isogenic MG1655. Because the Δ22 TLR grew to a higher OD than MG1655 when outgrown in MOPS, we captured images of the two cell types by a polarized light microscope. One by one cell, the size of stationary phase cells was recorded, and it turned out that the Δ22 TLR mutant cells were ~10% bigger than the MG1655. The larger cell size of the Δ22 TLR, in stationary phase, was confirmed by flow cytometry, and no size difference could be measured during balanced growing cells. By flow cytometry, the number of origins per cell was counted during balanced growth and at the stationary phase. In balanced growth, the numbers of origins per mass was close to the same for the two cell types, but in stationary phase, the numbers of origins per cell mass were 8% smaller for the Δ22 TLR mutant than for the MG1655. So it was not DNA filling up the larger cells in the stationary phase. Unfortunately, we sequenced the Δ22TLR mutant and learned it contained a large deletion of 12.5 kbp. Therefore, these results were not included in the report. Due to a lack of time, the investigation was not resumed in a true Δ22TLR mutant.
2.2.3 Homology

Based on sequence homology, the 22 TLRs from *E.coli* have been subdivided into two subfamilies: TLRA and TLRB (See Fig. 5A and 5B.) The TLRs interacting with Hfq belong to the subfamily of TLRA and are located in operons of the *rrn*; the rest of the TLRA are located in operons of tRNA. After probing the membranes from the stability experiments shown in Figure S3, Paper 2, with a probe complementary to the TLR rtV1 located at the tyr operon, no signal is detected when the RNA is harvested from the MG1655. But a signal does appear when the RNA is harvested from the isogenic Δ*hfq*. This indicates that the RtV1 from the tyr operon is not stabilized by the presence Hfq. This behavior is in contrast to the stability of three TLRs from the operons of *rrn*, being stabilized by presence of Hfq. This reality can be taken together with the lower presence of structure homology as seen between the 17 TLRA (Fig. 5C) than between the TLRs from the 29 other species in Paper 2, Fig 7b. It perhaps suggests yet another subdivision of the TLRA maybe also with respect to function. Therefore the only role of the rtV1, located in the operon of tyr-tRNA, could be as rho-dependent transcription terminators, as discovered by Kupper, Madden and Landy (Kupper et al., 1978; Madden and Landy, 1989). Nonetheless, it should be stressed that these are only indications and must be further investigated to be verified.

2.2.4 Hfq

Hfq-mediated RNA regulation has been an intensive study for more than two decades, but many aspects in understanding the regulation are still unanswered. Some aspects are at a very basic level such as the following: What is the concentration of Hfq in the cell at different conditions? Is Hfq found in fibril *in vivo* or is the fibril configuration discovered by Taghbalout et al. (Taghbalout et al., 2014) an experimental artifact? At the moment, data is insufficient, but more data points towards the importance of the C-terminal. If it is true that Hfq is bound in fibril *in vivo*, the dynamic between the monomer, the hexamer, the fibril and the othermers *in vivo*, is rather interesting. Such knowledge would drastically increase the understanding of Hfq and would have an impact on model development when trying to comprehend the mechanism behind the fast Hfq$_{6}$-mediated stress response. Also, the RNA-binding properties of the Hfq fibril *in vivo* is of high interest if the affinity towards RNA is decreased in the fibril formation as shown *in vitro* (Panja and Woodson, 2012). As such, the
amount of Hfq available for RNA interaction must be reduced, and competition for Hfq must be elevated.

If Hfq is participating in compartmentation located together with RNase E, RhlB, RNaseIII, PAPI, RnaA, and RraB (Diestra et al., 2009; Taghbalout and Rothfield, 2008; Taghbalout et al., 2014), what is the role of Hfq at this location? Along these lines, learning if the RrABF\textsuperscript{short} preferably binds Hfq at one cellular compartment over the other could perhaps help clarify the function of the RrABF\textsuperscript{short}. But such information would also be relevant for other regRNAs dependent on Hfq for completing their task. At last, since the suggested fibril formation is connected to the C-terminal (Taghbalout et al., 2014), it might be worth noting that the last ~20 amino acids of the C-terminal are only conserved in 7 to 8 other species (Sun et al., 2002). Therefore, the fibril formation may only be relevant for these 7 species, and perhaps care should be taken when comparing data from different species. However, there has been a report of Hfq fibrils formation \textit{in vitro} at concentration of 18 µM Hfq. This particular Hfq was from \textit{Pseudomonas aeruginosa}, which has a C-terminal 20 amino acids shorter than Hfq from \textit{E.coli} (Murina et al., 2015).

2.2.5 Possible errors with the experiments

2.2.5.1 C-terminal Epitope-tagged Hfq

Like a large number of other laboratories (Pfeiffer et al., 2007; Sittka et al., 2008; Tree et al., 2014; Zhang et al., 1998), we used an epitope-tagged Hfq and checked the function of the epitope-tagged Hfq by measuring its ability to mediate down-regulation of over-expressed duplex RNA (See S1 in Paper 2.). The increasing evidence for the importance of the C-terminal at Hfq may question if this test was sufficient.

2.2.5.2 Co-precipitation and re-association of lysis components

When binding the epitope-tagged protein to the appropriate resin, re-association of non-cellular, relevant components could be a problem. The ionic condition, pH, and/or concentration of the molecule of interest is most likely is different during the incubation than
at the condition in the cell (Mili and Steitz, 2004). Therefore, co-precipitation experiments cannot stand alone.

2.3 Conclusion

Through co-precipitation, stability measurement, RACE mapping, structure conservation analysis, in vitro experiments, and careful analysis of the operons containing the RrABF<sup>short</sup>, we were able to substantiate that three TLRs are regRNAs with a function likely linked to Hfq. Since we do not yet know the function of RrABF<sup>short</sup>, these experiments were the beginning of an interesting journey towards understanding the function of RrABF<sup>short</sup>. In future perspectives, a list of experiments is suggested.

2.4 Future perspectives

*First of all, knowing the function of RrABF<sub>short</sub> seems most important and the first experiment suggested is an ongoing project in our laboratory.*

1. An aptamer-tagged RrABF<sup>short</sup> is expressed in vivo, where after the aptamer-tagged RrABF<sup>short</sup> and its interaction partners are co-purified by running the lysate over an affinity column, the co-purified proteins can thereafter be determined by mass spectroscopy and the co-purified RNA by sequencing (Said et al., 2009) (ongoing).

2. Resume characterizing the phenotype of true ΔTLR22; if the true ΔTLR22 mutant displays a similar phenotype to the ΔTLR22 with the large deletion, a phenotypical characteristic could lead to understanding the function of the TLRs; thus, the recording of macromolecular compositions during balanced growth, stationary phase, and perhaps using a kit-based screening of phenotype would provide useful information.

*Are the TLRA further subdivided? (Maybe not a demanding task, but easy)*
3. Probe the membranes from Hfq co-precipitation experiments (Paper 2, Fig. 2a), with probes complementary to other TLRs, to see if the TLR from the tRNA operons interacts with Hfq.

4. Probe the membranes from a stability experiment (Paper 2 Fig. S3), with probes complementary to other TLRs, to see if the TLR from the tRNA operons are stabilized or destabilized by the presence of Hfq.

If the TLR from the tRNA operons behave like RrABF\textsuperscript{short}, it could be indicative for the TLRs to belong to one big family, and future experiments should be performed in the ‘true’ Δ22 TLR.

5. If the TLR from the tRNA operons behave differently, subdivisions of the ΔTLR mutants only deleting a specific group, as for example, the three downstream rrfA, rrfB, and rrfF, thereafter characterizing the sub-mutants phenotypically, could be interesting. Perhaps instead of deleting the TLRs, we can silence the TLRs by using the method of Bak et al. (Bak et al., 2015) to overcome the difficulties when trying to delete a portion of DNA at areas of high-sequence homology.

\textit{Test if the full length 9s inclusive RrABF\textsuperscript{short} can be processed by RNase E}

6. Incubate internally labeled full-length-transcript 9s inclusive RrABF\textsuperscript{short} with RNase E perhaps by using the method of Roy et al. (Roy et al., 1983).

\textit{Investigate the interaction between Hfq\textsubscript{6} and RrABF\textsuperscript{short}.}

7. Label the 2B transcript at the 3’ end and repeat the structure probing experiment. Perhaps other cleavages agents as; RNase T1, nuclease S1 would reveal the Hfq interaction site at the transcript 2B.
8. Use 2B mutants and by EMSA test if specific mutations are able to inhibit the Hfq interaction. Especially the 5’ end or the site at 2B that where protected towards Pb²⁺ cleavages (nt 63-68), after incubation with Hfq.

9. Use Hfq mutants similar to Schu et al. (Schu et al., 2015b) to locate the main interaction site for RrABF<sub>short</sub> at Hfq<sub>6</sub> in vivo.

10. To locate the main interaction site for Hfq<sub>6</sub> at RrABF<sub>short</sub> in vivo, use RrABF<sub>short</sub> mutants, and test if the stability of the RrABF<sub>short</sub> mutant is changed as compared to the non-mutated RrABF<sub>short</sub>.

Can RrABF<sub>short</sub> homologs be found in other species?

11. Bioinformatically determine if there are any species where both Hfq and one or more of the TLRs are predicted to be present.

12. And if so, repeat the Hfq co-precipitation experiment on this cell type.

Last but not least, investigate the cellular concentration, conformation, and localization of Hfq, which is urgent in understanding the full role of Hfq.

13. Record the in vitro polymerization of Hfq in presence of RNA, at a concentration that is biological relevant. Therefore, conduct the experiments of Panja et al. only include RNA (Panja and Woodson, 2012)

14. Systematically, quantify the total amount of Hfq per cell, along with quantifying the amount of Hfq located in different fractions of the cell. This can be accomplished by using the protocol of Miczak et al. (Miczak et al., 1991) to fractionate the cell lysate. Perhaps it is possible to divide the fraction form cytosol and nucleoid by using the method of (Masayoshi Enami, 1979) or(Ishihama et al., 2008). The total cellular amount and fragments of Hfq, can be quantified by using the protocol of (Taghbalout et al., 2014). If it is possible to localize Hfq in different fragments, this type of experiment could serve as control experiment for a functional epitope-tagged Hfq. If the epitope-tagged Hfq fractionates as the wt, it will indicate a functional C-terminal. To this an Hfq specific antibody must be used instead of an epitope tagged Hfq.
15. Repeat experiment 14 under starvation (for amino acid or other nutrients), to see if Hfq fractionate differently during starvation. Also other growth medias could be used.

16. Repeat experiment 14 with Hfq having a truncated C-terminal.

17. Develop a protocol that combine fragmentation and co-precipitation of epitope tagged Hfq, to see if Hfq embedded in fractions from different cellular locations, change RNA-binding profile. Perhaps in combination with an UV crosslinking protocol, as for example PAR-CLIP (Li et al., 2014) to avoid artificial re association of the bound RNA. At first not include the steps of sequencing in the protocol, but only analyze the harvest RNA by using the method of Northern blot. As such, look for the fragments containing the RrABFshort. Perhaps the RrABFshort are over-represented in one particular fraction, compared to the numbers of Hfq.

References:


Escherichia coli small noncoding RNAs OxyS and RprA, which are involved in the riboregulation of rpoS. Rna 19, 1089-1104.


Ishikawa, H., Otaka, H., Maki, K., Morita, T., and Aiba, H. (2012). The functional Hfq-binding module of bacterial sRNAs consists of a double or single hairpin preceded by a U-rich sequence and followed by a 3′ poly(U) tail. Rna 18, 1062-1074.


Li, Z., Pandit, S., and Deutscher, M.P. (1999a). Maturation of 23S ribosomal RNA requires the exoribonuclease RNase T. Rna 5, 139-146.
Li, Z., Pandit, S., and Deutscher, M.P. (1999b). RNase G (CafA protein) and RNase E are both required for the 5' maturation of 16S ribosomal RNA. The EMBO Journal 18, 2878-2885.


