JUSTUS-LIEBIG-
UNIVERSITÄT
GIESSEN

Kumulative Inauguraldissertation zur Erlangung des
Doktorgrades der Naturwissenschaften
– Dr. rer. nat. –

The interplay of RNases and sRNAs contributes to a complex layer of post-transcriptional gene regulation in proteobacteria

Daniel-Timon Spanka

Justus-Liebig-Universität Gießen
Institut für Mikrobiologie und Molekularbiologie
Dezember 2021

The interplay of RNases and sRNAs contributes to a complex layer of post-transcriptional gene regulation in proteobacteria

Daniel-Timon Spanka

This project was started in July 2018 and ended in December 2021. It was supervised by Prof. Dr. Gabriele Klug, head of the Institute of Microbiology and Molecular Biology, Department 08, Justus Liebig University Gießen.

Dean:	Prof. Dr. Thomas Wilke Justus Liebig University Gießen Systematics & Biodiversity Lab Heinrich-Buff-Ring 26-32, 35392 Gießen
1. Reviewer:	Prof. Dr. Gabriele Klug Justus Liebig University Gießen Institute of Microbiology and Molecular Biology Heinrich-Buff-Ring 26-32, 35392 Gießen
2. Reviewer:	Prof. Dr. Annegret Wilde University of Freiburg Faculty of Biology, Institute of Biology III Schänzlestr. 1, 79104 Freiburg i. Breisgau

Contents

Summary

Bacteria face constantly changing environmental conditions, like changes in temperature or exposure to oxidative stress. Since bacterial cells have only limited options to physically escape, a rapid adaption to the surroundings is required. As key element, gene expression is altered on transcriptional and post-transcriptional levels. This study illustrates the crucial role of polynucleotide phosphorylase (PNPase) for adaption to low temperatures, growth during organic peroxide stress, and during microaerobic growth conditions in the α-proteobacterium *Rhodobacter sphaeroides*. Moreover an RNA-seq analysis reavealed that the abundance of at least 334 of in total 4104 transcripts (7 %) depends on PNPase function, among them rRNAs and tRNAs which are lower abundant in a *pnp* mutant strain compared to the parental wild type. Additionally, PNPase influences the half-lives of different regualtory small RNAs (sRNAs) by stabilization or destabilization, among them CcsR1-4 and UpsM. Based on RNA-seq data, RNase E-, RNase III-, and PNPase-dependent differential RNA 3′ ends were globally predicted using an advanced analysis pipeline developed for this study. The results highlight a stepwise RNA processing first by endo- then by exonucleases: 5.9 % of all RNase E-dependent RNA 3′ ends and 9.7 % of all RNase III-dependent 3′ ends are further degraded by PNPase.

Small RNAs are pivotal for post-transcriptional gene expression since they specificially regulate their target RNAs in various ways. During the last years, sRNAs were found to be not exclusively transcribed by an own promoter but to be derived from mRNA 5′ or 3′ untranslated regions (UTRs). So far only few examples of UTR-derived sRNAs are known from the model organisms *Escherichia coli*, *Salmonella enterica*, and *R. sphaeroides*. This study depicts the sRNA landscape in *R. sphaeroides*: According to global predictions, UTR-derived sRNAs are indeed numerous and account for 37 % of all sRNAs. In a genome wide screening five novel UTR-derived sRNAs were detected in *R. sphaeroides* and subsequently characterized. Some of these sRNAs, among them UdsC, were induced by oxidative stress or during different growth phases. Altered transcription rate is not always responsible for the observed changes of sRNA abundances indicating the involvement of other factors. The RNase E-dependent processing of several UTR-derived sRNAs varied for example during the early stationary growth phase and during growth under iron limiting conditions. *In vivo* RNase E, RNase III, and PNPase are the enzymes which are mainly involved in *R. sphaeroides* UTR-derived sRNA maturation, processing or degradation. These observations were strongly supported by a global predictive approach: based on multiple RNA-seq datasets obtained from different RNase mutant strains, transcription start sites and Rho-independent terminator predictions the generation mechanisms of all *R. sphaeroides* UTR-derived sRNA 5′ and 3′ ends were computed, emphasizing the importance of RNase E.

Taken together this work provides new insights in the complex networks of post-transcriptional gene regulating in bacteria. It sheds light on the interwoven relationship between sRNAs and RNases which can either act individually or work together to form a sophisticated layer of gene regulation.

Chapter 1

RNases and small RNAs as key players in bacterial post-transcriptional gene regulation

Review Article

The world is a dangerous place

When Heraclitus noted more than 2000 years ago "Panta rhei – Everything flows", he most likely had no bacterial cells in mind. However, the life of microorganisms could hardly be described more aptly, because stasis in terms of environmental and growth conditions is sought here in vain. The habitats of many bacteria are characterized by changing physical influences, be it through changes in ambient temperature, incidence of light or UV radiation. In addition, parameters such as salinity or the pH value, the availability of trace elements or energy sources change. Of course, not every bacterial cell lives on its own, but shares the habitat with various other species, whereby the entire population develops dynamically. Microorganisms must protect themselves from these sometimes very harmful environmental influences and adapt to them at the population and individual level. Since individual cells, in contrast to higher eukaryotes, can only isolate themselves from the environment to a limited extent, and in most cases the possibilities for escape are also extremely restricted, other strategies are used. In the case of a slow and gradual change, populations of bacteria can adapt through evolution. Those cells with beneficial mutations will assert themselves in the long run. But when microorganisms face sudden changes in environmental conditions, there are only two options: "Flight or fight!".

Flight – biofilms, persistence & sporulation

On a shorter time scale, many bacterial species can escape disadvantageous growth conditions by forming a biofilm (reviewed by Yin et al., 2019). To do this, individual cells first attach to a surface, where they form a matrix of self-synthesized polysaccharides, nucleic acids, and proteins. As the process continues, additional cells can attach, causing the biofilm to increase in size and grow three-dimensionally. This biofilm provides diverse protection, for example against antiobiotics (Anderl et al., 2000), UV-C radiation (Bernbom et al., 2011) or extreme pH values (de Paz et al., 2007). Nonetheless, conditions are never ideal for all cells since for example the oxygen or nutrient availability may be limited in the inner layers of the biofilm.

Persister cells are tightly connected to biofilms. Many bacterial species, among them various human pathogens, can make use of a special trick to circumvent dangerous habitats and form so-called persister cells. These cells represent a small subpopulation of an isogenic culture which exhibits different phenotypical characteristics than the other cells. Facing antibiotic treatment or other harmful surrounding conditions this subpopulation enters a dormant state whereas the rest of the population dies (reviewed by Fisher et al., 2017; Lewis, 2010). They grow very slowly or are even non-growing (Balaban et al., 2004). Since most antibiotics require an active metabolism to develop an effect (e.g. rifamycin: Calvori et al., 1965; ampicillin: Campoli-Richards & Brogden, 1987; aminoglycosides: Serio et al., 2018), dormant persister cells can hardly be targeted and their susceptibility to antibiotics is dramatically reduced. In contrast to resistant cells, persister cells can only tolerate antibiotics and exhibit the same minimal inhibitory concentration as the non-persistent cells (Brauner et al., 2016). Persistence is often regulated by endogenous toxin-antitoxin systems, for example encoded by the *hip* genes (Black

et al., 1991, 1994), *tisB/istR* (Dörr et al., 2010) or *hok/sok* (Gerdes et al., 1985) in *Escherichia coli*. The recovery process after dormancy is as specific as the mechanism is which initiated persistence. The toxin TisB is for example induced during DNA damage via the LexA regulator (Vogel et al., 2004) and accounts for a collapse of the proton motive force (Unoson & Wagner, 2008; Gurnev et al., 2012). TisB-dependent persister cells in *E. coli* seem to accumulate reactive oxygen species (ROS; Edelmann & Berghoff, 2019) thus components of the alkyl hydroperoxide reductase Ahp as well as the outer membran protein OmpF are crucial during the recovery phase (see Chapter 6). Another studied example is the TacT toxin from *Salmonella enterica* which inhibits translation by blocking an amine group of amino acids on charged tRNA molecules (Cheverton et al., 2016). *Salmonella* cells can detoxify using the peptidyl-tRNA hydrolase (Pth). Pth recycles damaged tRNAs and thereby counteracts the TacT effects which leads to resumption of growth (Cheverton et al., 2016).

Another way to escape unfavourable conditions is motility. Besides for example twitching and gliding, one widely spread mechanism of movement is based on rotating flagella which allow the cells to move. In parallel bacteria sense their surrounding area e.g. in terms of molecules (chemotaxis, aerotaxis), light (phototaxis), Earth's magnetic field (magnetotaxis) and direct the movement along a gradient towards the desired condition (e.g. reviewed by Nakamura & Minamino, 2019 and Wong-Ng et al., 2018). In a temporal dimension, bacteria often can escape by sporulation or other resistant dormant bodies. A multitude of bacterial genera – among them *Bacillus* and *Clostridium* – can form endospores, a type of cell which is metabolically inactive and has several protective properties. Sporulation is started in response to stress conditions and the spores can outlast harmful conditions such as dryness, radiation or toxic chemical compounds (reviewed by Cho & Chung, 2020). When the situation becomes more favourable, spores can be activated, germinate and start dividing. Of course the principle of sporulation as a way to survive harmful conditions is also widely spread among fungi (reviewed by Wyatt et al., 2013).

Fight – gene regulation

However, it is not always only external factors that affect bacterial cells and require rapid adaptation: Intracellular processes such as transcription/translation may be deregulated, leading to accumulation of non-functional intermediates or misfolded proteins. Regardless of whether external environmental conditions and growth conditions change or processes in the cell run incorrectly, bacteria must use effective mechanisms to adapt to the new conditions within the shortest possible time. To achieve this, genes are regulated at transcriptional and post-transcriptional levels. Due to the small volume of the bacterial cell and the lack of a nucleus, these processes are coupled both spatially and temporally. They are functionally interwoven, influence each other and thus form very finely balanced regulatory networks.

Transcriptional regulation

According to the central dogma of molecular biology, information cannot be passed from DNA to protein without RNA as an intermediate. Thus balanced RNA levels

help to ensure a proper proteome composition and are pivotal for cell viability. Essentially, two fundamental factors influence the abundance of an RNA species in the cell: the production of RNA by transcription and its degradation or decay.

The DNA-dependent RNA-polymerase is required for mRNA synthesis. The approximately 400 kDa complex is composed of the five subunits α, β, β´, ω and an interchangeable σ factor. This last-mentioned subunit binds to specific promoter sequence motifs upstream of genes and hence guides the RNA-polymerase complex to the gene to be transcribed. After binding the correct site and opening the promoter complex, the initiation process is finished and transcription starts. With this system, sets of genes –called regulons– coding for proteins with similar functions can easily be selected and expressed, for example in response to stress (reviewed by Feklístov et al., 2014; Hinkle & Chamberlin, 1972). Different σ factors compete for binding to the RNA-polymerase and the availability of these different proteins depends for example on their subcellular localization, sequestration, or the rate of synthesis (reviewed e.g. by Bervoets & Charlier, 2019). Furthermore, transcription rate is heavily influenced and modulated by various other proteins, called transcription factors. These are DNA binding proteins which can regulate transcription in a negative or positive manner. Common mechanisms of action include blocking the promoter region, deformation of DNA or forming a roadblock for the RNA-polymerase (negative). Activation of transcription can for example be enhanced by class I/II activators or by conformational changes of DNA in order to improve RNA-polymerase binding (positive) (reviewed e.g. by Bervoets & Charlier, 2019). Besides the described transcription initiation also elongation and termination are regulated in multiple ways to ensure optimal gene expression (reviewed by Washburn & Gottesman, 2015).

RNA thermometers & Co.

A very basic type of post-transcriptional gene regulation are RNA thermometers, which are short sequences usually located in 5′ UTRs of selected transcripts. These RNA sequences form secondary structures via base pairing upstream of the start codon and thus usually mask the ribosome binding site (RBS), repressing translation. As temperature increases, the previously structured region melts, making the RBS accessible to ribosomes, and translation initiation begins (reviewed by Narberhaus et al., 2006). In prokaryotes RNA thermometers control genes for the adaption to heat (*rpoH* mRNA in *E. coli*, Morita et al., 1999), probably also to cold stress (Yamanaka et al., 1999) and virulence genes (Waldminghaus et al., 2007). Mechanistically, RNA thermometers are the purists within post-transcriptional gene regulation because no RNases, sRNAs, or other molecules are required for proper functionality. The information for the complete switch is present in the sequence of the RNA thermometer, making the system inherently robust and less prone to failure. Conceptually related to RNA thermometers are riboswitches. These are 5′ UTR located RNA sequences which very specificially bind small target molecules such as thiamin pyrophosphate (Mironov et al., 2002) or S-adenosylmethionine (Wang & Breaker, 2008). After ligand binding, translation can be facilitated or repressed (re-

viewed by Tucker & Breaker, 2005). Besides the two mentioned mechanisms, proteolysis represents another important instance regarding post-translational gene regulation. Protein degradation, also, has to be controlled in multiple ways (i.e. selectivity and specificity), but this is outside of the scope of this review article (suggested further reading: Mahmoud & Chien, 2018).

RNases, key players of post-transcriptional gene regulation

In addition to transcriptional regulation, the control and modulation of RNA stability is a crucial mechanism in the course of differential gene regulation. RNA degradation is catalyzed in the bacterial cell by numerous ribonucleases (RNases). Many RNases are highly conserved (e.g. RNase III), but there are also differences between different classes and species of bacteria. Unlike for example *E. coli*, *R. sphaeroides* does not encode RNase II.

Messenger RNA degradation is often initiated in *E. coli* by an endonuclease (RNase E, G, III, or P) which breaks the transcript into shorter fragments in an initial reaction. In the case of RNase E, mRNAs with a 5′ monophosphate are preferentially processed and unstructured AU-rich regions are cut. In *E. coli* these fragments are subsequently hydrolytically or phosphorolytically cleaved by 3′-to-5′ exonucleases (PNPase, RNase II, and RNase R) (reviewed by Deutscher, 2006). However, in *E. coli* the aforementioned exonucleases sometimes cannot completely degrade the mRNA fragments, so that short RNA oligonucleotides result from this reaction. These can then be converted into mononucleotides by an oligoribonuclease (Ghosh & Deutscher, 1999).

Besides degradation of mRNAs, many RNases such as RNase E or RNase III are also involved in the processing and maturation of RNAs. A notable example is the maturation of 16S, 23S and 5S rRNAs, which are co-transcribed in many bacteria and then cut from the transcript by RNase E and RNase III (reviewed by Deutscher, 2009; Klein & Evguenieva-Hackenberg, 2002). Additional enzymes are involved, as the rRNA operon also contains sequences for up to three tRNAs. In *R. sphaeroides* the 23S rRNA is further cleaved by RNase III at stem-loop sequences (so-called intervening sequences) and then processed by RNase J, among others (Rische & Klug, 2012). Thus, the 23S rRNA is fragmented into a 5.8S-like rRNA, a 1.1 kb, and a 1.5 kb molecule (Evguenieva-Hackenberg & Klug, 2000; Zahn et al., 2000).

In contrast to an enzymatic degradation, RNA molecules are also subject to natural decay. In aqueous solution, the 2′-OH end in an RNA molecule can cause a nucleophilic attack on the phosphate group of the next nucleotide, resulting in strand breakage (Brown & Todd, 1955; Westheimer, 1968). This autohydrolysis occurs spontaneously and affects the stability of all RNAs equally; therefore, it is of no importance with respect to gene regulation.

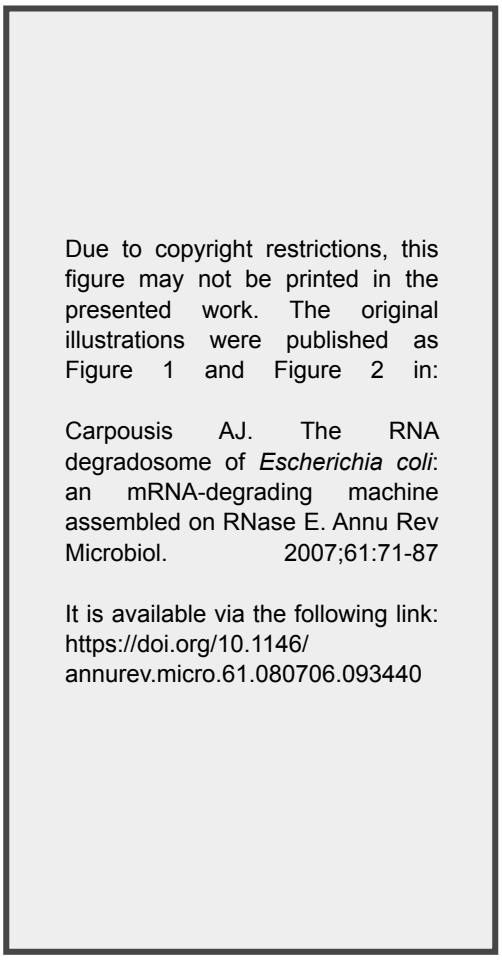


Figure 1.1: Schematic representation of RNA degradation and the degradosome in *E. coli*.
A Degradation of mRNAs (colored purple) is initiated by an endonucleolytic cut catalyzed mainly by RNase E or RNase III. The RNase E cleavage sites are often found near the RNA 5′ end. **B** Subsequently, the resulting mRNA fragments are digested to nucleotides by exoribonucleases. RNA helicase B and poly(A) polymerase serve as auxiliary enzymes as they facilitate the degradation of structured RNAs. Highlighted in red: Components of the degradosome. **C** The scaffold for the degradosome is the non-catalytic part of RNase E (C-terminal half; shown in blue). There are binding domains for RNA helicase B (purple), enolase (dimer, yellow), and one PNPase trimer (green). Four of these RNase E molecules with bound enzymes form the degradosome. Each of these subunits is structurally and functionally independent. Illustration modified after Carpousis, 2007.

In the following, the degradosome, and the enzymes RNase E, and PNPase are described in more detail, since they are of particular importance for gene regulation by maturation and degradation of sRNAs. Of course other enzymes such as RNase III or RNA pyrophosphohydrolase RppH can also contribute to these regulatory networks. However, as far as it is currently known, they are involved only to a minor extent (further reading: Court et al., 2013 and Bechhofer & Deutscher, 2019).

The bacterial degradosome

RNA molecules are subject to constant degradation. In *E. coli* this degradation is first initiated by an endonuclease – mostly RNase E – which cuts the transcript near the 5′ end or intercistronically if translation is halted (Figure 1.1A). The mRNA fragments generated in this way are then digested to nucleotides by various exoribonucleases (Figure 1.1B). These two steps are spatially closely coupled in *E. coli* and many other bacteria, as the degradation of RNA is often catalyzed by the degradosome. This is a very large complex which is composed of several enzymes and is often membrane-associated (Khemici et al., 2008). In *E. coli*, an ATP-dependent RNA helicase (RhlB), an enolase dimer, and at the RNase E C-terminus, a PNPase trimer bind to the non-catalytic part of RNase E (Figure 1.1C; Py et al., 1994, 1996; Carpousis et al., 1994). RNA is bound in this complex by RNase E and endonucleolytically cleaved. Thereafter, structured RNA molecules are linearized by RhlB and can be immediately degraded by PNPase in the 3′-to-5′ direction (reviewed by Carpousis, 2007). Enolase is actually an enzyme of glycolysis and its exact function in the context of the degradosome remains unknown

to date. However, at least two studies have shown that enolase is responsible for rapid degradation of transcripts encoding transporters of glucose and other carbon sources (Morita et al., 2004; Bernstein et al., 2004). However, the composition of the degradosome is dynamic; for example, in *E. coli* the helicase RhlB can be subsituted by CspA (cold shock protein A) under cold shock conditions (Prud'homme-Généreux et al., 2004). Also, the bound proteins vary among organisms: in *Rhodobacter capsulatus*, PNPase is probably not part of the degradosome (Jäger et al., 2001).

RNase E

The enzyme RNase E plays an important role in the maturation of ribosomal RNA (reviewed by Deutscher, 2009), mRNA and sRNA (Chao et al., 2017; Förstner et al., 2018; Eisenhardt et al., 2018) as well as in the degradation of various RNA species. Because of these important functions, RNase E is an essential enzyme in many Gram-negative model organisms (*E. coli*: Baba et al., 2006; Ow & Kushner, 2002; *S. enterica*: Hammarlof et al., 2011; *R. sphaeroides*: Förstner et al., 2018). However, RNase E homologs cannot be found in all bacteria; in the Gram-positive *Bacillus subtilis*, only the heterologous enzyme RNase Y (Shahbabian et al., 2009) exists. However, the functions of RNase E and RNase Y appear to be similar, such that a *rny* deletion in *B. subtilis* can be complemented by the gene *rne* from *E. coli* (Laalami et al., 2021).

RNase E preferentially cleaves at single-stranded AU-rich sequences (McDowall et al., 1994) and global analyses in *S. enterica* indicate a highly conserved uracil at position +2 in the consensus motif (Chao et al., 2017). An RNase E monomer subdivides into two functionally distinct domains, the catalytic (N-terminal half, NTH) and the macromolecular scaffold domain (C-terminal half, CTH; Figure 1.2A). The catalytic part consists of an S1 domain, a 5′ sensor and DNase I-like domain, and two RNase H-like domains (McDowall & Cohen, 1996). This NTH of the enzyme is necessary for RNA substrate binding and catalytic activity. The CTH exhibits binding sites for the enzymes RhlB, enolase, and PNPase, making the domain mainly important for the structure and interaction of the degradosome (reviewed by Carpousis et al., 2009 and Mackie, 2012). Two RNase E monomers interact via two Zn^{2+} binding cysteine residues between the large and small domains, forming a functional dimer. Again, two of these dimers bind via the small domains and form a tetramer (Figure 1.1C). In the active site of RNase E, a Mg^{2+} ion is chelated by the carboxyl groups of two aspartic acids (Callaghan et al., 2005). The single-stranded RNA reaches the active site through an S1 domain tunnel, is bound via interactions with three different amino acids, and is eventually hydrolyzed. This results in a monophosphorylated 5′ end and a 3′-OH end.

Basically, there are two pathways by which mRNAs can be degraded by RNase E. They are distinguished into the RNA 5′-dependent and the 5′-independent mechanism (direct-entry pathway, Baker & Mackie, 2002). Originally, it was assumed that only the phosphorylation state of the RNA 5′ end determines whether an mRNA is degraded by RNase E or not (Mackie, 1998). In the 5′-dependent mechanism degradation is initiated by the pyrophosphohydrolase RppH which first hy-

Figure 1.2: Schematic representation of ribonuclease E.
RNase E is a large enzyme and as a monomer consists of a total of 1061 amino acids. The structure can be subdivided into two major elements: The N-terminal half (NTH) consists of one DNase I-like domain, two RNase H-like domains, and one S1 and one 5′ sensor domain. The C-terminal half (CTH) has no catalytic activity, but is necessary for interaction with other enzymes. Therefore, binding sites for other enzymes of the degradosome RlhB, enolase and PNPase are found here. Illustration modified after Mackie, 2012.

drolyzes two phosphate groups of the RNA 5′ end (Deana et al., 2008). Alternatively, 5′ monophosphates can be generated by cuts from other endoribonucleases. RNase E preferentially catalyzes RNAs with monophosphorylated 5′ ends, whereas, in contrast, triphosphorylated RNA substrates are poorly cut (Mackie, 1998; Jiang & Belasco, 2004). This preference is generated by the 5′ sensor domain (also called phosphate-binding pocket), whose existence and function were first postulated and later verified by a crystal structure (Callaghan et al., 2005). An RNA end containing monophosphate is initially enclosed by the phosphate-binding pocket and the RNase H-like subdomain. This is followed by an allosteric conformational change of the S1 domain, pushing the RNA with the phosphate backbone close to the active site where it is hydrolyzed (mousetrap model, Callaghan et al., 2005). For 5′ triphosphates, the phosphate binding pocket is sterically too narrow, resulting in a reduced processing rate. However, the binding properties *per se* are not affected by the nature of the phosphorylation state. Further studies suggest that RNase E is catalytically active mainly through the 5′-independent degradation pathway (Clarke et al., 2014). In this process, RNase E can cut substrates without binding to the RNA 5′ end, thus bypassing the specificity described above. Particularly susceptible to this direct degradation pathway are RNAs that either have a stem-loop structure at the 5′ end or are particularly accessible to RNase E because they are not protected by translating ribosomes (Joyce & Dreyfus, 1998; Braun, 1998). Such a stem-loop structure located in the 5′ UTR is bound in the course of RNase E autoregulation, in which RNase E degrades its own *rne* transcript (Mudd & Higgins, 1993; Schuck et al., 2009). The feed-back mechanism is efficient: if the transcription rate of the *rne* gene in *E. coli* is increased by a factor of 21, it follows that RNase E levels will only slightly more than double (Jain & Belasco, 1995; Mudd & Higgins, 1993).

That RNase E is involved in multiple processing and degradation reactions of RNA is also evident from global TIER-seq (transiently inactivating an endoribonuclease followed by RNA-seq) analyses of

RNase E cleavage sites in *Salmonella* and *Rhodobacter*. In *S. enterica*, this revealed approximately 22,000 RNase E-dependent RNA 5′ ends, of which nearly 60 % were found in coding sequences. The remaining sites were distributed among rRNAs, tRNAs, 5′/3′ UTRs, sRNAs, and intergenic sequences (Chao et al., 2017). The cleavage sites in *R. sphaeroides* identified using the same methodology are comparable in terms of both number (23,000 sites) and distribution between the different genomic regions (Förstner et al., 2018). However, it is striking that no RNase E-dependent cleavage site could be found in about half of all genes (approx. 2100). Moreover, in *Salmonella* a high proportion of cleavage sites could be identified in the region of the stop codon (Chao et al., 2017) while in *R. sphaeroides* RNase E cuts more frequently in the immediate vicinity of the start codon (Förstner et al., 2018).

The importance of RNase E to the bacterial cell is further highlighted by the fact that RNase E plays an important structuring role in addition to its active catalytic function. The C-terminal half forms the basic scaffold for the binding of all other enzymes of the degradosome, so that this multiprotein complex can not assemble without RNase E. At the same time, part of the non-catalytic domain (segment A of RNase E) serves as an anchor in the inner cytoplasmic membrane (Khemici et al., 2008).

PNPase, the RNA shredder

In addition to RNase E, polynucleotide phosphorylase (PNPase) is involved in many RNA maturation and degradation processes and is often associated with the degradosome (reviewed by Briani et al., 2016). Bacterial PNPase is conserved and has strong structural and functional similarities to the exosome of archaea and eukaryotes (reviewed by Lin-Chao et al., 2007). Basically, PNPase catalyzes the 3′-to-5′ phosphorolysis of polyribonucleotides, releasing nucleoside diphosphates (NDPs), which requires Mg^{2+}. The equilibrium of the reaction can be shifted, for example, by increased concentrations of NDPs, so that the reverse 5′-to-3′ polymerization reaction is preferentially catalyzed. Degradation by PNPase of single-stranded RNA occurs until stable secondary structures are reached at the 3′-OH end. Subsequently, by extending the RNA in the 5′-to-3′ direction, the RNA can be further destabilized by recruiting exoribonucleases over the newly synthesized heteropolymeric RNA (Mohanty & Kushner, 2000). Thus, PNPase forms a complement to poly(A) polymerase (PAP), which is responsible for much of the 5′-to-3′ elongation of RNA in *E. coli* (Mohanty & Kushner, 1999).

In *E. coli*, PNPase is a very abundant enzyme. Depending on the dataset and study examined, it is found among the most abundant 5 % to 10 % of all proteins (Protein Abundance Database: PAXdb, 2021). In some organisms such as *E. coli* PNPase is not essential, while deletion of the *pnp* gene is not possible in *Pseudomonas aeruginosa* or *R. sphaeroides* (Chen et al., 2016; see Chapter 2). The functional homotrimer is composed of three Pnp monomers that form a ring-like structure with a centrally located channel (Figure 1.3). RNA molecules are bound mainly by the KH/S1 domain and subsequently degraded at the active site. Deletion of the KH/S1 domain leads to a widening of the channel in *E. coli* in addition to a strongly reduced binding

Figure 1.3: Crystal structure of PNPase from *E. coli*.
A Each PNPase monomer is composed of two RNase PH domains (blue and yellow) and an α-helical domain (green). The KH/S1 domain is dynamic and therefore could not crystallize. **B+C** The functional holoenzyme consists of three PNPase monomers that assemble into a ring-like structure. Of critical importance is the central channel in the middle of the trimer, at the ends of which amino acids relevant for processivity are localized. Illustration modified after Shi et al., 2008.

affinity to RNA, but without affecting the catalytic properties (Shi et al., 2008).

The functions of PNPase are diverse. They include degradation of mRNAs (Kinscherf & Apirion, 1975; Mohanty & Kushner, 2003) and sRNAs (Andrade & Arraiano, 2008; Andrade et al., 2012), processing of tRNAs (Li & Deutscher, 1994) and maturation of 16S rRNA in *E. coli* (Sulthana & Deutscher, 2013). In recent years, other functions of PNPase have been discovered, for example, the enzyme is involved in recognizing oxidized and thus defective RNA in *E. coli* (Hayakawa et al., 2001) and is part of homologous recombination and DNA repair systems (Carzaniga et al., 2017). Studies of the *E. coli* and *R. sphaeroides* transcriptome depict the versatile role of PNPase in regard to RNA degradation: levels of all classes of RNAs but mainly mRNAs are somehow influenced by deletion of *pnp* or of the RNA-binding domain KH/S1 (Dressaire et al., 2018; see Chapter 2). In *B. subtilis*, PNPase even degrades ssDNA (Cardenas et al., 2009) at high intracellular Mn^{2+} and low phosphate concentration. Moreover, PNPase was shown to be involved in phage defense in the human pathogenic organism *Listeria monocytogenes* by Sesto et al. (2014). In strain EGD-e, in which *cas* genes (CRISPR associated genes) are absent, PNPase undertakes the processing of CRISPRs (clustered regularly interspaced short palindromic repeats) and is also necessary for the correct activity of the CRISPR system.

An enzyme with such extensive functions requires strict and well-balanced regulation (reviewed by Condon, 2015 and Briani et al., 2016). In *E. coli* the *pnp* gene is located in the *rpsO-pnp* operon and is controlled by two different promoters (P_{rpsO} and P_{pnp}, Figure 1.4A). The 5′ UTR of the *pnp* messenger is structured and forms a stem-loop. The double-stranded hairpin loop can be cut by RNase III, but this does not yet affect *pnp* translation. The newly generated 37 nt-long RNA fragment is eventually degraded by PNPase in the 3′-to-5′ direction (Figure 1.4B,C). Through this degradation, the single-stranded monophosphorylated 5′-OH end of the *pnp* mRNA becomes accessible to RNase E, which then cleaves the

transcript within the open reading frame. In parallel, the accessible binding sites BS1 and BS2 recruit the protein CsrA as a dimer and binding to BS2 blocks the Shine-Dalgarno sequence (Park et al., 2015). This effectively prevents translational initiation (Figure 1.4D). Under cold shock conditions, the described autoregulation can be briefly suspended during the adaptation phase (Beran & Simons, 2001).

Figure 1.4: Autoregulation of the *E. coli* PNPase.
A Schematic representation of the *rpsO-pnp* operon from *E. coli*. P: promoter; ter: terminator; SL: stem-loop. **B-D** Stepwise processing of the *pnp* transcript, see main text for detailed description. BS1/BS2: CsrA binding sites. BS2 is part of the Shine-Dalgarno sequence. Illustration modified after Condon, 2015.

Another level of post-transcriptional regulation of *pnp* expression arises from the interaction of the PNPase mRNA with the sRNA SraG (Fontaine et al., 2016). The *sraG* gene is located between *rpsO* and *pnp* on the (–) strand and encodes the anti-sense sRNA (asRNA) SraG (Figure 1.4A). The approximately 215 nt-long primary transcript is processed by RNase E, RNase III, and PNPase to form the functional asRNA SraG (190 nt). Fontaine et al. (2016) demonstrated that SraG negatively affects the expression of PNPase by binding the *pnp* mRNA, preventing assembly of the translation initiation complex. *In vitro* experiments additionally revealed that upon binding to the *pnp* transcript, an RNase III cleavage site is formed in the double-stranded region, destabilizing the mRNA. Based on the mechanisms shown, the complexity of the regulation becomes apparent. Thus, PNPase levels are modulated at least by the transcription rate of *pnp*, activity of RNase E, RNase III, and PNPase itself, as well as by the asRNA SraG.

As a major 3′-to-5′ exonuclease, PNPase is not only responsible for the degradation of mRNAs but can also digest sRNAs. In *E. coli*, PNPase was found to be the major enzyme which degrades multiple sRNAs that are not bound to Hfq (Andrade & Arraiano, 2008). The effect was reported to be growth phase-dependent and is most prominent during the stationary phase (Andrade et al., 2012). Surprisingly, De Lay & Gottesman (2011) were able to show that the Hfq-bound sRNAs RyhB, SgrS, and CyaR are destabilized in a *pnp* mutant compared to wild type. Since the mechanism for this oberservation seemed counterintuitive, the group of De Lay studied the phenomenon again. A novel approach called ‘short RNA-seq’ revealed that PNPase degrades short mRNA fragments, which also bind to Hfq. They originate from mRNAs which have previously been targeted by sRNAs or can be generated by endonucleases. In the absence of a PNPase-mediated degradation, those fragments are stabilized and can interact with

Hfq and their respective sRNAs through base pairing. This in turn leads to an increased decay of the bound sRNAs through endonucleolytic cleavage (Cameron et al., 2019). In contrast to these observations, several analyzed sRNAs are highly stabilized in *R. sphaeroides* harbouring an inactive form of the PNPase (see Chapter 2 and Chapter 3). Remarkably of the three analyzed Hfq-dependent sRNAs only UpsM fits the model of De Lay and was stabilized in the presence of PNPase. CcsR1-4 and SorY showed the opposite effect despite their reported binding to Hfq. However, UdsA, as an Hfq-independent sRNA, was strongly stabilized in the *pnp* mutant strain. Those different results suggest the existence of other still unknown organism-specific factors that may influence the PNPase mediated decay of sRNAs.

Different RNases work together in RNA processing

Very recent studies adressed the cooperative degradation of RNA molecules. Data from high-throughput RNA next generation sequencing (NGS) is the basis for a novel type of analysis: the global determination of RNA 5′/3′ ends. To the best of our knowledge, the principle of an algorithm to globally map RNase E cleavage sites based on NGS RNA-Seq data was first developed by Clarke et al. (2014) and further enhanced by Konrad Förstner in cooperation with the Jörg Vogel lab (Chao et al., 2017). This algorithm allows to determine cleavage sites genome wide at single nucleotide resolution. In the following years, this concept was further developed and applied to investigate ribonuclease dependent cleavage sites in *E. coli*, *S. enterica*, *R. sphaeroides*, and *Streptococcus pyogenes* (Chao et al., 2017; Förstner et al., 2018; Lécrivain et al., 2018). Deeper insights were gained when the group of Emmanuelle Charpentier determined the cleavage sites of the endoribonuclease RNase Y (the RNase E orthologue in Gram-positive bacteria) and the three exoribonucleases PNPase, YhaM, and RNase R in *Streptococcus pyogenes* (Lécrivain et al., 2018). Next, the cleavage sites and so-called targetomes of these RNases were compared and analyzed in terms of overlapping RNA 5′/3′ ends. This global study highlights that endo- and exonucleases work together during RNA processing and degradation (Broglia et al., 2020). In more than half of all cases, RNase Y cleavage sites could be found in the proximity of 3′ RNA ends where exoribonucleolytic digestion stopped. Yet cleavage by RNase Y followed by an exoribonuclease dependent degradation starting at this very cleavage site was detected in a small fraction (6 % PNPase-dependent; 2 % YhaM-dependent). In a similar approach, differential RNA 3′ ends depending on either PNPase, RNase E or RNase III were globally mapped in the Gram-negative organism *R. sphaeroides* (see Chapter 2). Although the strains and the exact algorithm used were different, the results were very similar in *R. sphaeroides*. PNPase accounts for 6 % of all RNase E-generated and 9.7 % of all RNase III-dependent RNA 3′ ends. Both studies illustrate a cooperative RNA degradation by endo- and exonucleases. Moreover they both suggest that exoribonucleases other than PNPase or YhaM are likely to be involved in degradation of RNase E processed RNA.

Regulation of RNases

It is clear from the examples that RNases are major regulators for gene expression

since they either directly degrade mRNAs or act indirectly by modulating sRNA levels. Therefore, the question of the ancient Roman satirist Juvenal "Who guards the guardians themselves?" can also be rephrased in this context to "Who regulates the regulators?". So how is it ensured that both the number, activity, and stoichiometric ratio of RNases to each other are optimal at each time point in the cell so that no collateral damage occurs?

Several principles account for the ribonuclease regulation (e.g. reviewed by Deutscher, 2021). One common mechanism is the feed-back or autocatalytical regulation. In *E. coli*, at least RNase E (Mudd & Higgins, 1993), RNase III (Bardwell et al., 1989) and PNPase (reviewed by Condon, 2015 and Briani et al., 2016) control the amount and/or translation rate of their own messengers by binding and processing of the mRNAs. Another aspect is the nature of the potential substrate itself; RNase E, for example, preferes AU-rich sequences in many organisms (McDowall et al., 1994), whereas RNase III cleaves double-stranded RNA structures. Even known RNase E substrates are not all processed in the same manner. Under iron limiting conditions, processing of some RNase E-dependent UTR-derived sRNAs is decreased in *R. sphaeroides* whereas others are not affected or even cleaved at a higher rate (see Chapter 3).

In many organisms, enzymatic activity is altered by post-translational modifications such as phosphorylation or acetylation of specific amino acids (e.g. Magasanik, 1989; Soppa, 2010). The *E. coli* RNase R is a well studied example and an acetylation of the Lys544 strongly impacts the protein levels: During exponential growth phase, this acetylated lysine allows binding of the two *trans*-translation factors transfer-messenger RNA (tmRNA) and the small protein B (SmpB) resulting in a fast decay of RNase R (Liang & Deutscher, 2010). Since the acetyltransferase Pka responsible for this modification is absent during stationary phase, RNase R molecules cannot be acetylated (Liang & Deutscher, 2011; Liang et al., 2011). This in turn prevents binding of tmRNA/SmpB and thus strongly increases the RNase R half-life. Moreover activity can directly be influenced by external factors which bind to RNases. In *E. coli*, the proteins RraA and RraB (regulator of ribonuclease activity) can bind RNase E which then changes a) the degradosome composition and b) the set of RNAs which are degraded (Lee et al., 2003; Gao et al., 2006).

As already depicted in the previous section, expression of *pnp* is regulated in multiple ways and among them via the small RNA SraG (Fontaine et al., 2016). Although this anti-sense RNA is transcribed from its own promoter and does not require PNPase functionality for maturation, the mechanism highlights the interwoven relationship between sRNAs and RNases. On the one hand, ribonucleases maturate UTR-derived sRNAs resulting in functional regulators. On the other hand, RNase messengers can also be targeted by sRNAs.

Finally the target accessibility and localization of RNases play an important role, though it does not regulate RNase activity *per se*. One example is RNase E which is membrane associated in many bacteria. Moreover, ribonucleases cannot efficiently

cleave RNAs which are bound by proteins (e.g. ribosomal proteins) or are occupied by actively translating ribosomes. Both are protective factors which increase RNA stability.

Small regulatory RNAs

The first small RNAs (small in terms of molecule length) were discovered in the 1970s through improved biochemical methods. These include, for example, 4.5S RNA, 6S RNA, tmRNA, and RNase P RNA (Griffin, 1971; Jain et al., 1982). What they have in common is that they perform specific functions in the cell and do not necessarily fall into the definition of sRNAs with a regulatory function that is commonly used today. Only a few more sRNAs were detected in the following years, so that only a total of ten *E. coli* sRNAs were known by 1999 (reviewed by Wassarman et al., 1999). Major technical improvements have been made during the past decade: The developement of high-throughput sequencing platforms allowed the detection of novel sRNAs without prior time consuming cloning or using microarrays (Kröger et al., 2012). In addition to the technical achievements, the understanding of how small RNAs work has also improved and they have been detected in many bacterial species. Among them are *E. coli* (Wassarman et al., 1999), *S. enterica* (Vogel, 2009), *Streptococcus pyogenes* (Perez et al., 2009), *Staphylococcus aureus* (Bohn et al., 2010), *R. sphaeroides* (Berghoff et al., 2009), *B. subtilis* (Preis et al., 2009), and also Cyanobacteria (Axmann et al., 2005).

Small RNAs contribute to post-transcriptional gene regulation in various different pathways, e.g. starvation stress response (Amin et al., 2016), outer membrane composition (Vogel & Papenfort, 2006) or toxin-antitoxin systems (Sarpong & Murphy, 2021), and the effects on gene expression can be either negative or positive (e.g. reviewed by Jørgensen et al., 2020; Wagner & Romby, 2015; see Figure 1.5). First, regulation requires base pairing between sRNA and the target RNA. The number of base paired nucleotides varies. An initial seed region of 8 nt to 9 nt is commonly observed in *trans*-acting sRNAs. Pairing can then be extended depending on the two sequences. The highest number of paired nucleotides can be found in asRNAs. They are transcribed from the opposite strand of the mRNA target and have thus a perfect complementary sequence which allows a base pairing over the full length of the asRNA. *Trans*-encoded sRNAs (meaning being transcribed from a locus distant to the target RNA) often have a negative effect on the target gene expression. The *E. coli* sRNA OxyS for example is induced by oxidative stress and blocks translation of its target mRNA *fhlA* by covering the Shine-Dalgarno sequence (Altuvia et al., 1998; Figure 1.5A). Often degradation of sRNA and mRNA is initiated after duplex formation (e.g. MicC-*ompD* duplex; Pfeiffer et al., 2009; Figure 1.5B). This destabilization can be RNase E- (common; e.g. MicC) or RNase III-mediated (not so common; e.g. asPcrL from *R. sphaeroides*; Reuscher & Klug, 2021). Activation of gene expression can be achieved by increasing translation of the mRNA target (Figure 1.5C). The *rpoS* mRNA from *E. coli* has a long and highly structured 5′ UTR harbouring an inaccessible ribosome binding site (RBS). At least three sRNAs (DsrA, RprA and ArcZ) can bind in this

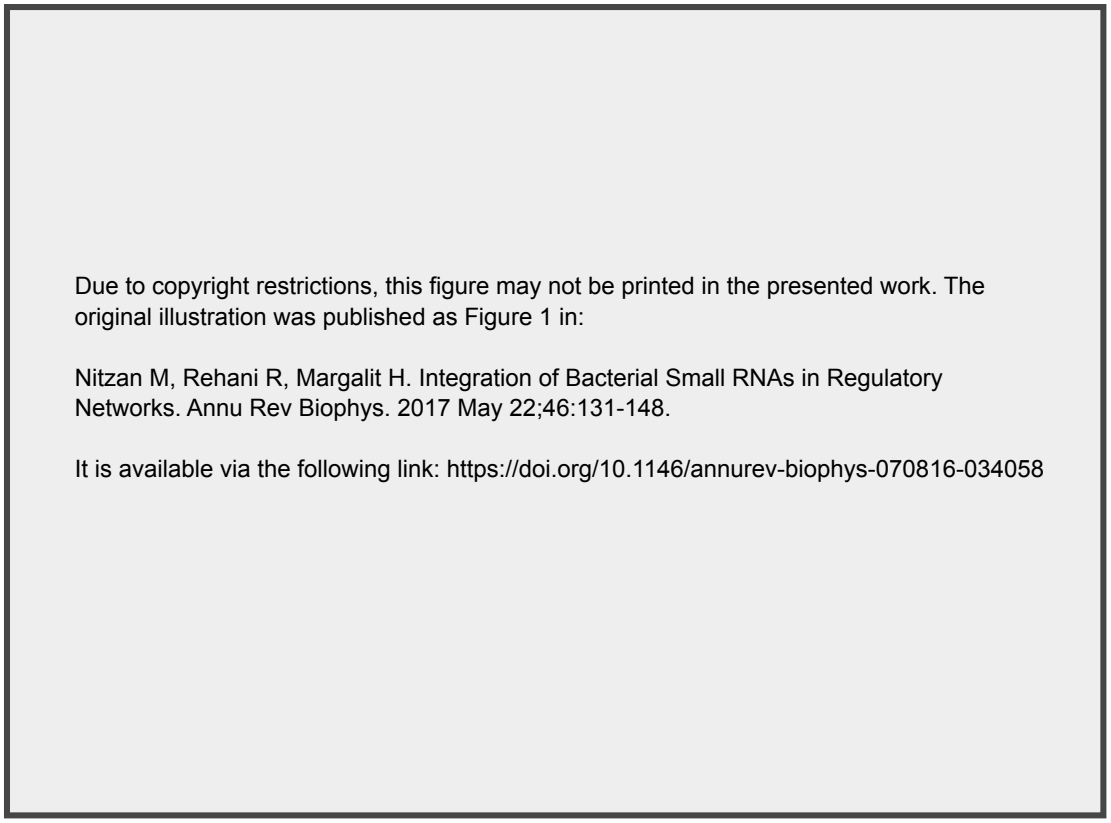


Figure 1.5: Small RNAs can regulate their target mRNAs in multiple ways.
Regulation of an mRNA target requires first base pairing with the respective sRNA in all described mechanisms. The gene can thus be negatively (**A**+**B**) or positively (**C**+**D**) regulated. Mechanisms for downregulation are translation inhibition by blocking the ribosome binding site (**A**) or recruting endonucleases such as RNase E which mediates the target RNA degradation (**B**). Target genes can also be upregulated: sRNA binding can make a previously blocked RBS accessible to the ribosome and thus elevate translation (**C**). Moreover target mRNAs can be stabilized due to structural changes which then prevent RNase cleavage (**D**). See main text for examples. Illustration modified after Nitzan et al., 2017.

UTR thus freeing the RBS by preventing secondary structure formation (Majdalani et al., 1998; reviewed by Papenfort & Vanderpool, 2015 and Mika & Hengge, 2014). Finally this allows translation of the alternative sigma factor RpoS. Another way to activate a target gene is achieved by stabilizing the mRNA transcript. This was for example observed for the *cfa* mRNA in *E. coli* which accumulates after being bound by the sRNA RydC (Fröhlich et al., 2013). Besides that, sRNAs can also act as sponges. In *B. subtilis* the sRNA RosA can bind the two target sRNAs RoxS and FsrA which in turn controls the target RNA levels and also the regulatory activity of the two sRNAs (Durand et al., 2021). A similar observation was made in *R. sphaeroides*, where an interaction of two sRNAs was described (Grützner et al., 2021). Both studies are special because of a) the sponge function and b) the uncommon sRNA-sRNA interaction (suggested further reading: Figueroa-Bossi & Bossi, 2018).

The above illustrated five mechanisms of sRNA action are of course simplified and do not cover all options of how sRNAs may work. Often sRNAs can bind multiple RNA targets and participate in complex regulatory networks where even single sRNAs may activate some targets and repress other ones (Balbontín et al., 2010; Jeon et al., 2021).

Rhodobacter sphaeroides – a metabolic Jack of all trades

Small RNAs have been investigated in several species and *R. sphaeroides* has become a relevant model organism in this field (Berghoff et al., 2009). The purple bacterium *R. sphaeroides* belongs to the class α-3 proteobacteria and is related to *Sinorhizobium meliloti*, *Agrobacterium tumefaciens*, and *Brucella melitensis*, among other species of purple bacteria (Woese et al., 1984; van Niel, 1944). Compared to *R. capsulatus*, the cells are relatively large, measuring 2.0 µm to 2.5 µm in length (Imhoff, 2006). *R. sphaeroides* thrives mainly in stagnant waters and has two flagella (reviewed by Camarena & Dreyfus, 2020 and de la Mora et al., 2015). The genome of *R. sphaeroides* is composed of two chromosomes CI (3.2 Mbp) and CII (0.9 Mbp) and five endogenous plasmids with lengths between 0.37 Mbp to 1.1 Mbp. Among them, chromosome CI has a DnaA-dependent origin of replication (*oriC*, origin of replication C), but CII has a *repABC* origin of replication that is completely independent of DnaA (Choudhary et al., 2007). Based on a recent taxonomic analysis of gene composition and diverse phenotypic characteristics, five *Rhodobacter* species were assigned to the genus *Cereibacter*, including *R. sphaeroides* (Hördt et al., 2020). To ensure better comparability with previously published data, the established name *Rhodobacter sphaeroides* was retained in this work.

R. sphaeroides can generate ATP under a wide range of environmental conditions and is characterized by its marked metabolic versatility. The purple bacterium produces ATP through aerobic respiration, anoxygenic photosynthesis, or anaerobic respiration depending on the oxygen partial pressure in the surrounding medium as well as on the light intensity. In order to ensure an optimal energy yield and at the same time to avoid the formation of toxic by-products, for example by reactive oxygen species, these pro-

cesses are regulated in a multilayered manner (Pandey et al., 2011; Jäger et al., 2007; Braatsch et al., 2002; Shimada et al., 1992). In this context, the post-transcriptional gene regulation mediated by sRNAs plays a major role. Several small RNAs have been described which are involved in these complex networks: PcrZ is for example expressed when the oxygen concentration drops. It has a negative effect on selected mRNAs from the photosynthetic gene cluster (e.g. *pucA* and *puc2A*, both encode subunits form the light-harvesting complex) and thus it balances and counteracts the redox-dependent induction of photosynthesis genes (Mank et al., 2012). Other sRNAs such as SorX and SorY are important during the defense of singlet oxygen or during photooxidative stress (Peng et al., 2016; Adnan et al., 2015). A prominent role could be attributed to StsR which is part of a complex regulatory network. The sRNA is transcribed in a RpoHI/RpoHII-dependent manner during the late stationary growth phase and modulates the *rpoE* messenger. Since RpoE promotes RpoHII expression, StsR is part of a negative feed-back loop. Moreover, several photosynthesis genes are regulated by StsR either directly or indirectly (see Chapter 5).

Auxiliary proteins increase efficacy of sRNA mediated regulation

Annealing of *trans*-derived sRNAs with their target RNAs is often facilitated by RNA chaperones to accelerate the process of regulation. The first helper enzyme to be discovered in *E. coli* was the host factor for bacteriophage Qβ RNA replication (Hfq) which is the major RNA chaperone in many organisms (e.g. reviewed in Updegrove et al., 2016 and Wagner, 2013). Hfq forms a hexameric ring-like structure and resembles the Sm-protein found in Eukarya and Archaea (Sauter et al., 2003). Small RNAs are classified according to their binding properties to Hfq: Class I sRNAs bind to the proximal face of the Hfq ring and also to the rim (Schu et al., 2015; Figure 1.6). The target mRNA associates with the distal face, and pairing of the two RNAs is mediated in sequence regions which are not bound to Hfq. In contrast, Class II sRNAs wrap around Hfq binding the proximal and distal face. This forces the mRNA to bind the rim, since Hfq harbours only these three binding positions. Small RNAs belonging to Class II are considered to be much more stable than those from Class I (Schu et al., 2015). What both classes have in common, is that the sRNA binding sequence is located in proximity of Rho-independent terminators, whereupon U-rich stretches are prefered to bind the proximal face of Hfq (Otaka et al., 2011). Aside from that, Hfq performs various other tasks. It interacts with rRNA and thus influences ribosome biogenesis and translation efficiency (Andrade et al., 2018), and most likely affects tRNA maturation (Lee & Feig, 2008). Also interactions with RNA degrading enzymes (Andrade et al., 2012) and the termination factor Rho were observed (Rabhi et al., 2011), highlighting a role for Hfq which is beyond a simple base-pairing promoting function (reviewed by dos Santos et al., 2019).

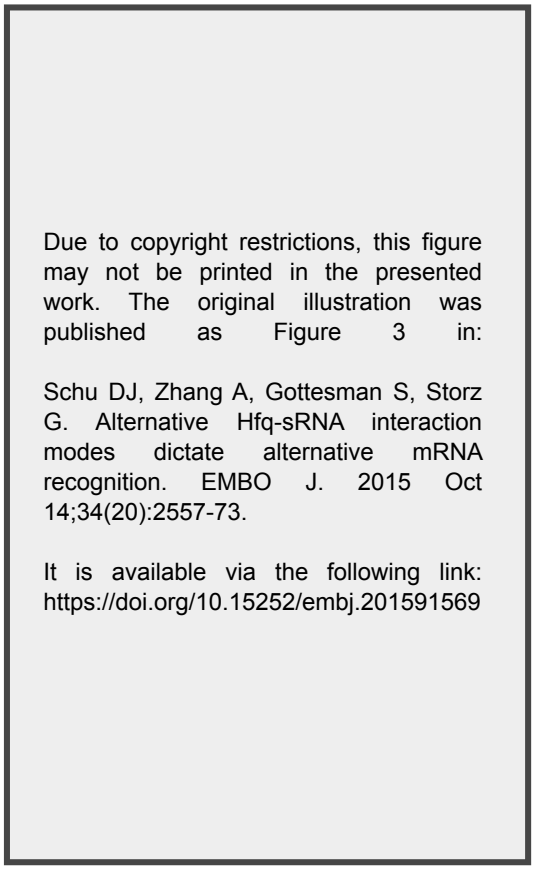


Figure 1.6: Small RNAs bind *E. coli* Hfq in different regions of the protein.
Small RNAs are classified according to their binding sites to Hfq. Class I bind the proximal face (colored red) of Hfq and the rim (colored purple), Class II sRNAs associate with the proximal and distal face (colored blue) by folding around Hfq forcing the target mRNA to bind the Hfq rim. Binding to the proximal face is facilitated by long U-stretches. Repeats of $(A\text{-}R\text{-}N)_n$ (adenine, purine nucleotide, any nucleotide) bind the distal face. UA-rich sequences are often found to attach to the rim. For more details see main text. Illustration modified after Schu et al., 2015.

Two other major RNA-binding proteins (RBPs) were identified and characterized: CsrA was first described as an regulator, which among other things controls different pathways of glycogen metabolism (Sabnis et al., 1995; Yang et al., 1996) and was later found to be involved in sRNA-mRNA binding in *E. coli* (Potts et al., 2017) and *B. subtilis* (Müller et al., 2019). However, its significance in sRNA binding seems to be low compared to Hfq. ProQ as a third sRNA binding protein gained attention in recent years (Smirnov et al., 2016; Olejniczak & Storz, 2017). In *S. enterica* it promotes stability of the sRNA ligands (Smirnov et al., 2016). Subsequent CLIP-seq experiments revealed that *E. coli* and *S. enterica* ProQ binds preferentially sRNAs and also mRNA 3′ UTRs, which is based on structural and not sequence motifs (Holmqvist et al., 2018). Most sRNAs could be crosslinked exclusively to Hfq, CsrA or ProQ and only a small fraction was linked to two or even all three proteins (Holmqvist et al., 2018). Nevertheless, not all species harbour genes coding these RNA chaperones. Hfq can be found in α-, β-, γ-, and δ-proteobacteria (Sun et al., 2002), and a structural homolog in cyanobacteria (Bøggild et al., 2009), whereas species from the group of ϵ-proteobacteria (e.g. *Campylobacter jejuni* and *Helicobacter pylori*) lack both *hfq* and *proQ* (Quendera et al., 2020). Other species of the α-proteobacteria (e.g. *R. sphaeroides*) seem to harbour exclusively Hfq.

UTR-derived sRNAs

During the first years after discovery, orphan sRNAs were regarded as the classical and only type of sRNA. They are located intergenically, transcription starts at an own promoter sequence and stops at a terminator structure. First published in 2015, mRNA 5′ (e.g. Weber et al., 2016) and 3′ UTRs (e.g. Miyakoshi et al., 2015 and Chao & Vogel, 2016) were recognized as reservoirs for sRNAs in different organisms (Figure 1.7). In contrast to orphan sRNAs, which only require transcription to be functional, UTR-derived sRNAs have

Figure 1.7: Genomic origins of sRNAs.
A small RNA can be located intergenic as an orphan gene, in mRNA 5′/3′ UTRs or also be on the opposite strand of the coding gene (anti-sense). A recent study points also to the existence of intragenic sRNAs in *E. coli*. Illustration modified after Hör et al., 2020.

to be matured by processing from a co-transcript consisting of the protein coding sequence and the small RNA. Besides UTRs as sRNA sources, a first study based on RNA sequencing also suggests the presence of intragenic sRNAs in *E. coli* (Dar & Sorek, 2018). One of the first UTR-derived sRNAs to be characterized was the CpxQ sRNA in *S. enterica* (Chao & Vogel, 2016). As a consequence of inner membrane damage, the gene of the membrane stress relevant chaperone CpxP is transcribed. In the next step, the sRNA is derived from the 3′ UTR of the *cpxP* mRNA by RNase E cleavage. After this maturation step, the translated protein and the functional sRNA can both counteract membrane stress either by refolding necessary proteins or by post-transcriptionally regulating the expression of required genes. In *R. sphaeroides*, several UTR-derived sRNAs were characterized in the last years (e.g. Weber et al., 2016; Peng et al., 2016; Grützner et al., 2021). The four homologous sRNAs CcsR1-4 are part of the oxidative stress defense and are co-transcribed with *ccaF1*, the gene which is located upstream. After transcription, RNase E is responsible for maturation (Billenkamp et al., 2015) and PNPase promotes degradation (see Chapter 2). Remarkably the DUF1127 protein CcaF1 binds its own co-transcript and is necessary for a correct maturation of the 3′ UTR-derived sRNAs CcsR1-4 (see Chapter 4).

Although examples from different organisms are characterized, the exact number of UTR-derived sRNAs is still unknown in many species. Based on a global analyis and *in vivo* validation, the relevance of UTR-derived sRNAs was depicted in *R. sphaeroides*. Approximately 38 % (30 of 79 in total) of all sRNAs are predicted to originate from 5′/3′ UTRs and are released from mRNA transcripts by ribonucleolytic cleavage. RNase E, RNase III and PNPase were confirmed to maturate or process these small RNAs *in vivo*, whereas endonucleolytic cleavages by RNase E are predicted to be most important for 3′ UTR-derived sRNAs (see Chapter 2 and Chapter 3). This study illustrates that sRNAs from UTRs are not a fancy and rare variant but seem to be rather common, at least in *R. sphaeroides*. Large scale experiments and RNA sequencing in other model organisms will reveal whether these observations hint to a general trend in proteobacteria.

Linking sRNAs and RNases creates new regulatory networks

Of course the very existence of UTR-derived sRNAs raises immediately the question as to why such a genomic organization could be beneficial. One simple answer could be that both components of mRNA-sRNA co-transcripts work together in similar pathways and that they form an operon like unit. This was reported in several cases so far, for example NarK/NarS (nitrite respiration in *S. enterica*; Wang et al., 2019), CpxP/CpxQ (inner membrane stress response in *S. enterica*; Chao & Vogel, 2016), OmpR1/SorX (oxidative stress response in *R. sphaeroides*; Peng et al., 2016; Zhao et al., 2018), and RhlS/RhlL (quorum sensing in *P. aeruginosa*; Thomason et al., 2019). Similar to a regular operon, this arrangement ensures a defined stoichiometry of all components. This principle is well illustrated by the example of photosynthesis gene regulation in *R. sphaeroides* and *R. capsulatus*. The genes *pufQBALMX* of the *puf* operon are co-transcribed (Belasco et al., 1985) and then directly processed by RNase E. Processing fragments possess stabilizing stem-loop structures that protect against further degradation by RNase E (stem-loop structure in 5′ UTR of *pufB*; Heck et al., 1996) or 3′-to-5′ exoribonucleases (stem-loop structure at RNA 3′ end; Klug et al., 1987). As a result, the fragments have different half-lives and are therefore present in fixed stoichiometric ratios. These specific differences in mRNA stability are necessary to ensure a correct ratio of the respective gene products (Klug et al., 1987).

But this is only one part of the story. Assuming a 1:1 ratio, such a model could also be achieved if the protein coding gene and the sRNA gene would each have its own promoter with the very same sequence. Instead, at least one more processing step is required for the UTR-derived small RNA to become functional. As illustrated in the example of the *puf* operon, the resulting ratios of the gene products are fixed. In this specific case this is useful, since the photosynthetic complexes encoded by the *puf* genes need to be composed in a fixed protein ratio in order to ensure a proper functionality. In other pathways, circuits or protein complexes, things might look very different and a dynamic change of ratios can be required. This is where UTR-derived sRNAs and RNases come into play: Adding the sRNA processing as an additional step of post-transcriptional gene regulation, the sRNA levels can be modulated if needed. This can range from a simple ON/OFF switch (sRNA is processed or not processed) to gradual changes in sRNA abundance which can be fine-tuned. So the extra layer of regulation allows dynamic modulation of the transcriptome composition, depending, for example, on RNase activity or RNase levels under certain (changing) growth conditions (Figure 1.8).

This mechanism is particularly useful if another layer of regulation needs to be integrated in a regulatory network, for example in the type I toxin-antitoxin system *tisB-istR1*. The protein TisB is expressed upon DNA damage and leads to severe effects for the cell. This is why a multi-layered and tight gene regulation is necessary (Vogel et al., 2004; Unoson & Wagner, 2008). The *tisB* primary transcript (inactive +1 *tisB* mRNA variant) cannot be translated since the ribosome binding site is covered in a highly structured 5′ UTR (Darfeuille et al., 2007). Only after

the mRNA is processed by an endonucleolytic cut (active +42 *tisB* mRNA variant) can translation start (reviewed by Wagner & Unoson, 2012; Berghoff et al., 2017). In this case ribonucleolytic cleavage ensures a transcriptional/translational uncoupling which acts as an additional barrier to protect the cell. At the same time, *tisB* translation is counteracted by binding of the σ^{70}-dependent antitoxin sRNA IstR1 which blocks the ribosome standby site and promotes the RNase III mediated co-degradation of sRNA and mRNA (Darfeuille et al., 2007).

Conclusion and perspective

This work provides novel and important aspects of UTR-derived sRNA maturation and thus contributes to a better understanding of post-transcriptional gene regulation in bacteria. Using *in vivo* data and predictive approaches the study systematically depicts the maturation of UTR-derived sRNAs and reveals RNase E, RNase III and PNPase as enzymes of major importance for this process. It highlightes furthermore, that on a genome wide scale sRNAs originating from mRNA 5′ or 3′ UTRs are not a fancy variant but are a common type of small regulatory RNA in the α-proteobacterium *R. sphaeroides*. Even considered individually, RNases and small RNAs are key players of post-transcriptional gene regulation because they control the general RNA turnover as well as target specific regulation. But when both elements are linked, the regulatory power even increases since the sRNA levels can be modulated in a fine balanced and condition-dependent manner in addition to the transcriptional regulation (Figure 1.8).

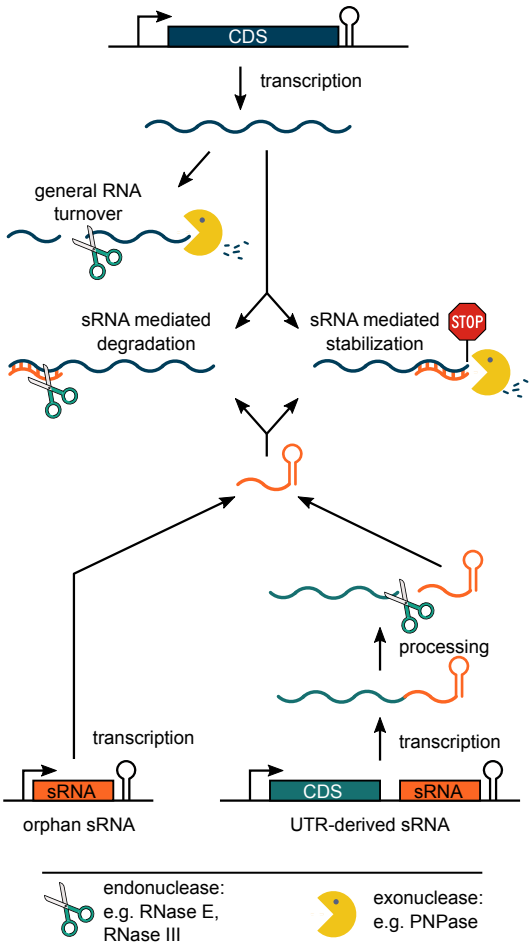


Figure 1.8: RNases and sRNAs work together in post-transcriptional gene regulation.
Endo- (scissors) and exonucleases (Pacman) are crucial for post-transcriptional gene regulation in bacteria. They are responsible for the regular RNA turnover and also for the sRNA mediated decay: sRNA-mRNA duplex formation often promotes mRNA cleavage by endonucleases such as RNase E or RNase III. Furthermore sRNAs can also stabilize the target (e.g. sRNA GadY, a roadblock for decay is assumed as mechanism (indicated by a stop sign), Fröhlich & Vogel, 2009). The involved small RNAs can be orphan or derived from an mRNA 5′ or 3′ UTR. In this case at least one additional processing step (commonly catalyzed by RNase E) is necessary after transcription of the mRNA-sRNA co-transcript. Lollipop structure indicates a Rho-independent terminator. mRNA transcripts are colored blue/green, sRNAs are colored orange. CDS: coding sequence.

Having the required tools and expertise at hand, prospecitve investigations could go one step further: Different types of RNA-

seq approaches, especially TIER-seq (transiently inactivating an endoribonuclease followed by RNA-seq), of as much mutant strains as possible lacking RNA processing enzymes could provide the information for a new type of data base. Enhanced bioinformatic tools could then reveal the processing and degradation patterns for every single transcript of an organism. Combining these data with other genome wide predictions such as transcription start sites, terminators or protein binding sites would allow monitoring of the fate of every RNA 'from the cradle to the grave'. However, since the overall goal is not just to accumulate complex data but to understand cellular processes, it remains important to unravel gene functions and make sense of the data from a biological point of view. This is why only a strong collaboration of predictive and microbiological approaches will justify such a meta-omics database which could help to shed light on the often still dark regulatory networks.

Abbreviations

The following abbreviations are used in this chapter:

asRNA	anti-sense small RNA
CDS	coding sequence
CLIP-seq	cross-linking immunoprecipitation followed by RNA-seq
CTH	C-terminal half of RNase E
GRAS	generally regarded as safe
NGS	next generation sequencing
NTH	N-terminal half of RNase E
PHA	poly-3-hydroxyalkanoate
PHB	poly-3-hydroxybutyrate
RBP	RNA-binding protein
RBS	ribosome binding site
RNA	ribonucleic acid
ROS	reactive oxygen species
SD	Shine-Dalgarno sequence
TEX	terminator 5′-phosphate dependent 5′-to-3′ exoribonuclease
TIER-seq	transiently inactivating an endoribonuclease followed by RNA-seq
tmRNA	transfer-messenger RNA
TSS	transcription start site
Uds	UTR-derived sRNA
UV	ultraviolett

Bibliography

F. Adnan, L. Weber, & G. Klug. The sRNA SorY confers resistance during photooxidative stress by affecting a metabolite transporter in *Rhodobacter sphaeroides*. *RNA Biology*, 12(5): 569–577, apr 2015. doi: 10.1080/15476286.2015.1031948.

S. Altuvia, A. Zhang, L. Argaman, A. Tiwari, & G. Storz. The *Escherichia coli* OxyS regulatory RNA represses *fhlA* translation by blocking ribosome binding. *EMBO J*, 17(20):6069–6075, oct 1998. doi: 10.1093/emboj/17.20.6069.

S. V. Amin, J. T. Roberts, D. G. Patterson, A. B. Coley, J. A. Allred, J. M. Denner, J. P. Johnson, G. E. Mullen, T. K. O'Neal, J. T. Smith, S. E. Cardin, H. T. Carr, S. L. Carr, H. E. Cowart, D. H. DaCosta, B. R. Herring, V. M. King, C. J. Polska, E. E. Ward, A. A. Wise, K. N. McAllister, D. Chevalier, M. P. Spector, & G. M. Borchert. Novel small RNA (sRNA) landscape of the starvation-stress response transcriptome of *Salmonella enterica* serovar typhimurium. *RNA Biology*, 13(3):331–342, mar 2016. doi: 10.1080/15476286.2016.1144010.

J. N. Anderl, M. J. Franklin, & P. S. Stewart. Role of antibiotic penetration limitation in *Klebsiella pneumoniae* biofilm resistance to Ampicillin and Ciprofloxacin. *Antimicrobial Agents and Chemotherapy*, 44(7):1818–1824, jul 2000. doi: 10.1128/aac.44.7.1818-1824.2000.

J. M. Andrade & C. M. Arraiano. PNPase is a key player in the regulation of small RNAs that control the expression of outer membrane proteins. *RNA*, 14(3):543–551, jan 2008. doi: 10.1261/rna.683308.

J. M. Andrade, V. Pobre, A. M. Matos, & C. M. Arraiano. The crucial role of PNPase in the degradation of small RNAs that are not associated with Hfq. *RNA*, 18(4):844–855, feb 2012. doi: 10.1261/rna.029413.111.

J. M. Andrade, R. F. Santos, I. Chelysheva, Z. Ignatova, & C. M. Arraiano. The RNAbinding protein Hfq is important for ribosome biogenesis and affects translation fidelity. *EMBO J*, 37 (11), apr 2018. doi: 10.15252/embj.201797631.

I. M. Axmann, P. Kensche, J. Vogel, S. Kohl, H. Herzel, & W. R. Hess. Identification of cyanobacterial non-coding RNAs by comparative genome analysis. *Genome Biol.*, 6(9):R73, 2005. doi: 10.1186/gb-2005-6-9-r73.

T. Baba, T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K. A. Datsenko, M. Tomita, B. L. Wanner, & H. Mori. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. 2(1), jan 2006. doi: 10.1038/msb4100050.

K. E. Baker & G. A. Mackie. Ectopic RNase E sites promote bypass of 5′-end-dependent mRNA decay in *Escherichia coli*. *Molecular Microbiology*, 47(1):75–88, dec 2002. doi: 10.1046/j.1365-2958.2003.03292.x.

N. Q. Balaban, J. Merrin, R. Chait, L. Kowalik, & S. Leibler. Bacterial persistence as a phenotypic switch. *Science*, 305(5690):1622–1625, sep 2004. doi: 10.1126/science.1099390.

R. Balbontín, F. Fiorini, N. Figueroa-Bossi, J. Casadesús, & L. Bossi. Recognition of heptameric seed sequence underlies multi-target regulation by RybB small RNA in *Salmonella enterica*. *Mol Microbiol*, 78(2):380–394, aug 2010. doi: 10.1111/j.1365-2958.2010.07342.x.

J. Bardwell, P. Régnier, S. Chen, Y. Nakamura, M. Grunberg-Manago, & D. Court. Autoregulation of RNase III operon by mRNA processing. *The EMBO Journal*, 8(11):3401–3407, nov 1989. doi: 10.1002/j.1460-2075.1989.tb08504.x.

D. H. Bechhofer & M. P. Deutscher. Bacterial ribonucleases and their roles in RNA metabolism. *Crit Rev Biochem Mol Biol*, 54(3):242–300, may 2019. doi: 10.1080/10409238.2019.1651816.

J. G. Belasco, J. Beatty, C. W. Adams, A. von Gabain, & S. N. Cohen. Differential expression of photosynthesis genes in *R. capsulata* results from segmental differences in stability within the polycistronic *rxcA* transcript. *Cell*, 40(1):171–181, jan 1985. doi: 10.1016/0092-8674(85)90320-4.

R. K. Beran & R. W. Simons. Cold-temperature induction of *Escherichia coli* polynucleotide phosphorylase occurs by reversal of its autoregulation. *Mol Microbiol*, 39(1):112–125, jan 2001. doi: 10.1046/j.1365-2958.2001.02216.x.

B. A. Berghoff, J. Glaeser, C. M. Sharma, J. Vogel, & G. Klug. Photooxidative stress-induced and abundant small RNAs in *Rhodobacter sphaeroides*. *Molecular Microbiology*, 74(6):1497–1512, dec 2009. doi: 10.1111/j.1365-2958.2009.06949.x.

B. A. Berghoff, M. Hoekzema, L. Aulbach, & E. G. H. Wagner. Two regulatory RNA elements affect TisB-dependent depolarization and persister formation. *Molecular Microbiology*, 103(6): 1020–1033, jan 2017. doi: 10.1111/mmi.13607.

N. Bernbom, B. Vogel, & L. Gram. *Listeria monocytogenes* survival of UV-C radiation is enhanced by presence of sodium chloride, organic food material and by bacterial biofilm formation. *International Journal of Food Microbiology*, 147 (1):69–73, may 2011. doi: 10.1016/j.ijfoodmicro.2011.03.009.

J. A. Bernstein, P.-H. Lin, S. N. Cohen, & S. Lin-Chao. Global analysis of *Escherichia coli* RNA degradosome function using DNA microarrays. 101(9):2758–2763, feb 2004. doi: 10.1073/pnas.0308747101.

I. Bervoets & D. Charlier. Diversity, versatility and complexity of bacterial gene regulation mechanisms: opportunities and drawbacks for applications in synthetic biology. *FEMS Microbiol Rev*, 43(3):304–339, feb 2019. doi: 10.1093/femsre/fuz001.

F. Billenkamp, T. Peng, B. A. Berghoff, & G. Klug. A cluster of four homologous small RNAs modulates C1 metabolism and the pyruvate dehydrogenase complex in *Rhodobacter sphaeroides* under various stress conditions. *Journal of Bacteriology*, 197(10):1839–1852, mar 2015. doi: 10.1128/jb.02475-14.

D. S. Black, A. J. Kelly, M. J. Mardis, & H. S. Moyed. Structure and organization of *hip*, an operon that affects lethality due to inhibition of peptidoglycan or DNA synthesis. *J Bacteriol*, 173(18):5732–5739, sep 1991. doi: 10.1128/jb.173.18.5732-5739.1991.

D. S. Black, B. Irwin, & H. S. Moyed. Autoregulation of *hip*, an operon that affects lethality due to inhibition of peptidoglycan or DNA synthesis. *J. Bacteriol.*, 176(13):4081–4091, jul 1994. doi: 10.1128/jb.176.13.4081-4091.1994.

A. Bøggild, M. Overgaard, P. Valentin-Hansen, & D. E. Brodersen. Cyanobacteria contain a structural homologue of the Hfq protein with altered RNA-binding properties. *FEBS Journal*, 276 (14):3904–3915, jul 2009. doi: 10.1111/j.1742-4658.2009.07104.x.

C. Bohn, C. Rigoulay, S. Chabelskaya, C. M. Sharma, A. Marchais, P. Skorski, E. Borezée-Durant, R. Barbet, E. Jacquet, A. Jacq, D. Gautheret, B. Felden, J. Vogel, & P. Bouloc. Experimental discovery of small RNAs in *Staphylococcus aureus* reveals a riboregulator of central metabolism. *Nucleic Acids Res*, 38(19):6620–6636, may 2010. doi: 10.1093/nar/gkq462.

S. Braatsch, M. Gomelsky, S. Kuphal, & G. Klug. A single flavoprotein, AppA, integrates both redox and light signals in *Rhodobacter sphaeroides*. *Molecular Microbiology*, 45(3):827–836, jul 2002. doi: 10.1046/j.1365-2958.2002.03058.x.

F. Braun. Ribosomes inhibit an RNase E cleavage which induces the decay of the *rpsO* mRNA of *Escherichia coli*. *The EMBO Journal*, 17(16): 4790–4797, aug 1998. doi: 10.1093/emboj/17.16.4790.

A. Brauner, O. Fridman, O. Gefen, & N. Q. Balaban. Distinguishing between resistance, tolerance and persistence to antibiotic treatment. *Nat Rev Microbiol*, 14(5):320–330, apr 2016. doi: 10.1038/nrmicro.2016.34.

F. Briani, T. Carzaniga, & G. Dehò. Regulation and functions of bacterial PNPase. *Wiley Interdiscip Rev RNA*, 7(2):241–258, jan 2016. doi: 10.1002/wrna.1328.

L. Broglia, A.-L. Lécrivain, T. T. Renault, K. Hahnke, R. Ahmed-Begrich, A. L. Rhun, & E. Charpentier. An RNA-seq based comparative approach reveals the transcriptome-wide interplay between 3′-to-5′ exoRNases and RNaseY. *Nature Communications*, 11(1), mar 2020. doi: 10.1038/s41467-020-15387-6.

D. M. Brown & A. R. Todd. Nucleic acids. *Annual Review of Biochemistry*, 24(1):311–338, jun 1955. doi: 10.1146/annurev.bi.24.070155.001523.

A. J. Callaghan, M. J. Marcaida, J. A. Stead, K. J. McDowall, W. G. Scott, & B. F. Luisi. Structure of *Escherichia coli* RNase E catalytic domain and implications for RNA turnover. *Nature*, 437(7062):1187–1191, oct 2005. doi: 10.1038/nature04084.

C. Calvori, L. Frontali, L. Leoni, & G. Tecce. Effect of Rifamycin on protein synthesis. *Nature*, 207 (4995):417–418, jul 1965. doi: 10.1038/207417a0.

L. Camarena & G. Dreyfus. Living in a foster home: The single subpolar flagellum Fla1 of *Rhodobacter sphaeroides*. *Biomolecules*, 10(5): 774, may 2020. doi: 10.3390/biom10050774.

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

D. M. Campoli-Richards & R. N. Brogden. Sulbactam/Ampicillin. A review of its antibacterial activity, pharmacokinetic properties, and therapeutic use. *Drugs*, 33(6):577–609, jun 1987. doi: 10.2165/00003495-198733060-00003.

P. P. Cardenas, B. Carrasco, H. Sanchez, G. Deikus, D. H. Bechhofer, & J. C. Alonso. *Bacillus subtilis* polynucleotide phosphorylase 3′-to-5′ DNase activity is involved in DNA repair. *Nucleic Acids Res*, 37(12):4157–4169, may 2009. doi: 10.1093/nar/gkp314.

A. J. Carpousis. The RNA degradosome of *Escherichia coli*: An mRNA-degrading machine assembled on RNase E. 61(1):71–87, oct 2007. doi: 10.1146/annurev.micro.61.080706.093440.

A. J. Carpousis, G. V. Houwe, C. Ehretsmann, & H. M. Krisch. Copurification of *E. coli* RNase E and PNPase: Evidence for a specific association between two enzymes important in RNA processing and degradation. 76(5):889–900, mar 1994. doi: 10.1016/0092-8674(94)90363-8.

A. J. Carpousis, B. F. Luisi, & K. J. McDowall. Chapter 3 endonucleolytic initiation of mRNA decay in *Escherichia coli*. In *Progress in Molecular Biology and Translational Science*, pages 91–135. Elsevier, 2009. doi: 10.1016/s0079-6603(08)00803-9.

T. Carzaniga, G. Sbarufatti, F. Briani, & G. Dehò. Polynucleotide phosphorylase is implicated in homologous recombination and DNA repair in *Escherichia coli*. *BMC Microbiol*, 17(1), apr 2017. doi: 10.1186/s12866-017-0980-z.

Y. Chao & J. Vogel. A 3′ UTR-derived small RNA provides the regulatory noncoding arm of the inner membrane stress response. *Molecular Cell*, 61(3):352–363, feb 2016. doi: 10.1016/j.molcel.2015.12.023.

Y. Chao, L. Li, D. Girodat, K. U. Förstner, N. Said, C. Corcoran, M. Śmiga, K. Papenfort, R. Reinhardt, H.-J. Wieden, B. F. Luisi, & J. Vogel. In vivo cleavage map illuminates the central role of RNase E in coding and non-coding RNA pathways. *Molecular Cell*, 65(1):39–51, jan 2017. doi: 10.1016/j.molcel.2016.11.002.

H. Chen, T. Dutta, & M. P. Deutscher. Growth phase-dependent variation of RNase BN/Z affects small RNAs. *Journal of Biological Chemistry*, 291(51):26435–26442, dec 2016. doi: 10.1074/jbc.m116.757450.

A. M. Cheverton, B. Gollan, M. Przydacz, C. T. Wong, A. Mylona, S. A. Hare, & S. Helaine. A *Salmonella* toxin promotes persister formation through acetylation of tRNA. *Molecular Cell*, 63 (1):86–96, jul 2016. doi: 10.1016/j.molcel.2016.05.002.

W.-I. Cho & M.-S. Chung. *Bacillus* spores: a review of their properties and inactivation processing technologies. *Food Sci Biotechnol.*, 29(11):1447–1461, oct 2020. doi: 10.1007/s10068-020-00809-4.

M. Choudhary, X. Zanhua, Y. X. Fu, & S. Kaplan. Genome analyses of three strains of *Rhodobacter sphaeroides*: Evidence of rapid evolution of chromosome II. *Journal of Bacteriology*, 189(5): 1914–1921, mar 2007. doi: 10.1128/jb.01498-06.

J. E. Clarke, L. Kime, D. R. A., & K. J. McDowall. Direct entry by RNase E is a major pathway for the degradation and processing of RNA in *Escherichia coli*. *Nucleic Acids Research*, 42(18):11733–11751, sep 2014. doi: 10.1093/nar/gku808.

C. Condon. Airpnp: Auto- and integrated regulation of polynucleotide phosphorylase. *Journal of Bacteriology*, 197(24):3748–3750, dec 2015. doi: 10.1128/jb.00794-15.

D. L. Court, J. Gan, Y.-H. Liang, G. X. Shaw, J. E. Tropea, N. Costantino, D. S. Waugh, & X. Ji. RNase III: Genetics and function, structure and mechanism. *Annu Rev Genet*, 47(1):405–431, nov 2013. doi: 10.1146/annurev-genet-110711-155618.

D. Dar & R. Sorek. Bacterial noncoding RNAs excised from within protein-coding transcripts. *mBio*, 9(5), nov 2018. doi: 10.1128/mbio.01730-18.

F. Darfeuille, C. Unoson, J. Vogel, & E. G. H. Wagner. An antisense RNA inhibits translation by competing with standby ribosomes. *Molecular Cell*, 26(3):381–392, may 2007. doi: 10.1016/j.molcel.2007.04.003.

J. de la Mora, K. Uchida, A. M. del Campo, L. Camarena, S.-I. Aizawa, & G. Dreyfus. Structural characterization of the Fla2 flagellum of *Rhodobacter sphaeroides*. *Journal of Bacteriology*, 197(17):2859–2866, sep 2015. doi: 10.1128/jb.00170-15.

N. De Lay & S. Gottesman. Role of polynucleotide phosphorylase in sRNA function in *Escherichia coli*. *RNA*, 17(6):1172–1189, apr 2011. doi: 10.1261/rna.2531211.

L. E. C. de Paz, G. Bergenholtz, G. Dahlén, & G. Svensäter. Response to alkaline stress by root canal bacteria in biofilms. *International Endodontic Journal*, 40(5):344–355, may 2007. doi: 10.1111/j.1365-2591.2006.01226.x.

A. Deana, H. Celesnik, & J. G. Belasco. The bacterial enzyme RppH triggers messenger RNA degradation by 5′ pyrophosphate removal. *Nature*, 451(7176):355–358, jan 2008. doi: 10.1038/nature06475.

M. P. Deutscher. Degradation of RNA in bacteria: comparison of mRNA and stable RNA. *Nucleic Acids Research*, 34(2):659–666, jan 2006. doi: 10.1093/nar/gkj472.

M. P. Deutscher. Chapter 9 maturation and degradation of ribosomal RNA in bacteria. In *Progress in Molecular Biology and Translational Science*, pages 369–391. Elsevier, 2009. doi: 10.1016/s0079-6603(08)00809-x.

M. P. Deutscher. Regulation of bacterial ribonucleases. *Annual Review of Microbiology*, 75(1), jun 2021. doi: 10.1146/annurev-micro-020121-011201.

R. F. dos Santos, C. M. Arraiano, & J. M. Andrade. New molecular interactions broaden the functions of the RNA chaperone Hfq. *Curr Genet*, 65 (6):1313–1319, may 2019. doi: 10.1007/s00294-019-00990-y.

C. Dressaire, V. Pobre, S. Laguerre, L. Girbal, C. M. Arraiano, & M. Cocaign-Bousquet. PNPase is involved in the coordination of mRNA degradation and expression in stationary phase cells of *Escherichia coli*. *BMC Genomics*, 19(1), nov 2018. doi: 10.1186/s12864-018-5259-8.

S. Durand, A. Callan-Sidat, J. McKeown, S. Li, G. Kostova, J. R. Hernandez-Fernaud, M. T. Alam, A. Millard, D. Allouche, C. Constantinidou, C. Condon, & E. L. Denham. Identification of an RNA sponge that controls the RoxS riboregulator of central metabolism in *Bacillus subtilis*. *Nucleic Acids Res*, 49(11):6399–6419, jun 2021. doi: 10.1093/nar/gkab444.

T. Dörr, M. Vulić, & K. Lewis. Ciprofloxacin causes persister formation by inducing the TisB toxin in *Escherichia coli*. *PLoS Biol.*, 8(2):e1000317, feb 2010. doi: 10.1371/journal.pbio.1000317.

D. Edelmann & B. A. Berghoff. Type I toxin-dependent generation of superoxide affects the persister life cycle of *Escherichia coli*. *Scientific Reports*, 9(1), oct 2019. doi: 10.1038/s41598-019-50668-1.

K. M. Eisenhardt, C. M. Reuscher, & G. Klug. PcrX, an sRNA derived from the 3′- UTR of the *Rhodobacter sphaeroides puf* operon modulates expression of *puf* genes encoding proteins of the bacterial photosynthetic apparatus. *Molecular Microbiology*, 110(3):325–334, oct 2018. doi: 10.1111/mmi.14076.

E. Evguenieva-Hackenberg & G. Klug. RNase III processing of intervening sequences found in helix 9 of 23S rRNA in the alpha subclass of proteobacteria. *Journal of Bacteriology*, 182 (17):4719–4729, sep 2000. doi: 10.1128/jb.182.17.4719-4729.2000.

A. Feklístov, B. D. Sharon, S. A. Darst, & C. A. Gross. Bacterial sigma factors: A historical, structural, and genomic perspective. *Annu Rev Microbiol*, 68(1):357–376, sep 2014. doi: 10.1146/annurev-micro-092412-155737.

N. Figueroa-Bossi & L. Bossi. Sponges and predators in the small RNA world. *Microbiology Spectrum*, 6(4), jul 2018. doi: 10.1128/microbiolspec.rwr-0021-2018.

R. A. Fisher, B. Gollan, & S. Helaine. Persistent bacterial infections and persister cells. *Nat Rev Microbiol*, 15(8):453–464, may 2017. doi: 10.1038/nrmicro.2017.42.

F. Fontaine, E. Gasiorowski, C. Gracia, M. Ballouche, J. Caillet, A. Marchais, & E. Hajnsdorf. The small RNA SraG participates in PNPase homeostasis. *RNA*, 22(10):1560–1573, aug 2016. doi: 10.1261/rna.055236.115.

K. S. Fröhlich & J. Vogel. Activation of gene expression by small RNA. *Current Opinion in Microbiology*, 12(6):674–682, dec 2009. doi: 10.1016/j.mib.2009.09.009.

K. S. Fröhlich, K. Papenfort, A. Fekete, & J. Vogel. A small RNA activates CFA synthase by isoform-specific mRNA stabilization. *EMBO J*, 32(22):2963–2979, oct 2013. doi: 10.1038/emboj.2013.222.

K. U. Förstner, C. M. Reuscher, K. Haberzettl, L. Weber, & G. Klug. RNase E cleavage shapes the transcriptome of *Rhodobacter sphaeroides* and strongly impacts phototrophic growth. *Life Science Alliance*, 1(4):e201800080, aug 2018. doi: 10.26508/lsa.201800080.

J. Gao, K. Lee, M. Zhao, J. Qiu, X. Zhan, A. Saxena, C. J. Moore, S. N. Cohen, & G. Georgiou. Differential modulation of *E. coli* mRNA abundance by inhibitory proteins that alter the composition of the degradosome. *Molecular Microbiology*, 61(2):394–406, jul 2006. doi: 10.1111/j.1365-2958.2006.05246.x.

K. Gerdes, J. E. Larsen, & S. Molin. Stable inheritance of plasmid R1 requires two different loci. *J Bacteriol.*, 161(1):292–298, jan 1985. doi: 10.1128/jb.161.1.292-298.1985.

S. Ghosh & M. P. Deutscher. Oligoribonuclease is an essential component of the mRNA decay pathway. *Proceedings of the National Academy of Sciences*, 96(8):4372–4377, apr 1999. doi: 10.1073/pnas.96.8.4372.

B. E. Griffin. Separation of 32 P-labelled ribonucleic acid components. the use of polyethylenimine-cellulose (TLC) as a second dimension in separating oligoribonucleotides of 4.5 S and 5 S from *E. coli*. *FEBS Letters*, 15(3):165–168, jun 1971. doi: 10.1016/0014-5793(71)80304-6.

J. Grützner, B. Remes, K. M. H. Eisenhardt, D. Scheller, J. Kretz, R. Madhugiri, M. McIntosh, & G. Klug. sRNA-mediated RNA processing regulates bacterial cell division. *Nucleic Acids Research*, jun 2021. doi: 10.1093/nar/gkab491.

P. A. Gurnev, R. Ortenberg, T. Dörr, K. Lewis, & S. M. Bezrukov. Persister-promoting bacterial toxin TisB produces anion-selective pores in planar lipid bilayers. *FEBS Lett.*, 586(16):2529–2534, jun 2012. doi: 10.1016/j.febslet.2012.06.021.

D. L. Hammarlof, L. Liljas, & D. Hughes. Temperature-sensitive mutants of RNase E in *Salmonella enterica*. 193(23):6639–6650, sep 2011. doi: 10.1128/jb.05868-11.

H. Hayakawa, M. Kuwano, & M. Sekiguchi. Specific binding of 8-oxoguanine-containing RNA to polynucleotide phosphorylase protein. *Biochemistry*, 40(33):9977–9982, jul 2001. doi: 10.1021/bi010595q.

C. Heck, R. Rothfuchs, A. Jäger, R. Rauhut, & G. Klug. Effect of the pufQ-pufB intercistronic region on *puf* mRNA stability in *Rhodobacter capsulatus*. *Molecular Microbiology*, 20(6):1165–1178, jun 1996. doi: 10.1111/j.1365-2958.1996.tb02637.x.

D. C. Hinkle & M. J. Chamberlin. Studies of the binding of *Escherichia coli* RNA polymerase to DNA. *J Mol Biol*, 70(2):157–185, sep 1972. doi: 10.1016/0022-2836(72)90531-1.

E. Holmqvist, L. Li, T. Bischler, L. Barquist, & J. Vogel. Global maps of ProQ binding in vivo reveal target recognition via RNA structure and stability control at mRNA 3′ ends. *Mol Cell*, 70(5):971–982.e6, jun 2018. doi: 10.1016/j.molcel.2018.04.017.

A. Hördt, M. G. López, J. P. Meier-Kolthoff, M. Schleuning, L.-M. Weinhold, B. J. Tindall, S. Gronow, N. C. Kyrpides, T. Woyke, & M. Göker. Analysis of 1,000+ type-strain genomes substantially improves taxonomic classification of alphaproteobacteria. *Frontiers in Microbiology*, 11, apr 2020. doi: 10.3389/fmicb.2020.00468.

J. Hör, G. Matera, J. Vogel, S. Gottesman, & G. Storz. Trans-acting small RNAs and their effects on gene expression in *Escherichia coli* and *Salmonella enterica*. *EcoSal Plus*, 9(1), mar 2020. doi: 10.1128/ecosalplus.esp-0030-2019.

J. F. Imhoff. The phototrophic alphaproteobacteria. In *The Prokaryotes*, pages 41–64. Springer New York, 2006. doi: 10.1007/0-387-30745-1_2.

A. Jäger, S. Braatsch, K. Haberzettl, S. Metz, L. Osterloh, Y. Han, & G. Klug. The AppA and PpsR proteins from *Rhodobacter sphaeroides* can establish a redox-dependent signal chain but fail to transmit blue-light signals in other bacteria. *Journal of Bacteriology*, 189(6):2274–2282, mar 2007. doi: 10.1128/jb.01699-06.

C. Jain & J. G. Belasco. RNase E autoregulates its synthesis by controlling the degradation rate of its own mRNA in *Escherichia coli*: unusual sensitivity of the *rne* transcript to RNase e activity. 9(1):84–96, jan 1995. doi: 10.1101/gad.9.1.84.

S. K. Jain, M. Gurevitz, & D. Apirion. A small RNA that complements mutants in the RNA processing enzyme ribonuclease P. *Journal of Molecular Biology*, 162(3):515–533, dec 1982. doi: 10.1016/0022-2836(82)90386-2.

H. J. Jeon, Y. Lee, P. A. N. Monford, X. Wang, D. K. Chattoraj, & H. M. Lim. sRNA-mediated regulation of *gal* mRNA in *E. coli*: Involvement of transcript cleavage by RNase E together with Rho-dependent transcription termination. *PLoS Genet*, 17(10):e1009878, oct 2021. doi: 10.1371/journal.pgen.1009878.

X. Jiang & J. G. Belasco. Catalytic activation of multimeric RNase E and RNase G by 5'-monophosphorylated RNA. *Proceedings of the National Academy of Sciences*, 101(25):9211–9216, jun 2004. doi: 10.1073/pnas.0401382101.

M. G. Jørgensen, J. S. Pettersen, & B. H. Kallipolitis. sRNA-mediated control in bacteria: An increasing diversity of regulatory mechanisms. *Biochim Biophys Acta Gene Regul Mech*, 1863 (5):194504, may 2020. doi: 10.1016/j.bbagrm.2020.194504.

S. A. Joyce & M. Dreyfus. In the absence of translation, RNase E can bypass 5′ mRNA stabilizers in *Escherichia coli*. *Journal of Molecular Biology*, 282(2):241–254, sep 1998. doi: 10.1006/jmbi.1998.2027.

S. Jäger, O. Fuhrmann, C. Heck, M. Hebermehl, E. Schiltz, R. Rauhut, & G. Klug. An mRNA degrading complex in *Rhodobacter capsulatus*. *Nucleic Acids Res*, 29(22):4581–4588, nov 2001. doi: 10.1093/nar/29.22.4581.

V. Khemici, L. Poljak, B. F. Luisi, & A. J. Carpousis. The RNase E of *Escherichia coli* is a membrane-binding protein. *Mol Microbiol.*, 70(4):799–813, nov 2008. doi: 10.1111/j.1365-2958.2008.06454.x.

T. G. Kinscherf & D. Apirion. Polynucleotide phosphorylase can participate in decay of mRNA in *Escherichia coli* in the absence of ribonuclease II. 139(4):357–362, dec 1975. doi: 10.1007/bf00267975.

F. Klein & E. Evguenieva-Hackenberg. RNase E is involved in 5′-end 23S rRNA processing in α-proteobacteria. *Biochemical and Biophysical Research Communications*, 299(5):780–786, dec 2002. doi: 10.1016/s0006-291x(02)02738-9.

G. Klug, C. W. Adams, J. Belasco, B. Doerge, & S. N. Cohen. Biological consequences of segmental alterations in mRNA stability: effects of deletion of the intercistronic hairpin loop region of the *Rhodobacter capsulatus puf* operon. *The EMBO Journal*, 6(11):3515–3520, nov 1987. doi: 10.1002/j.1460-2075.1987.tb02677.x.

C. Kröger, S. C. Dillon, A. D. S. Cameron, K. Papenfort, S. K. Sivasankaran, K. Hokamp, Y. Chao, A. Sittka, M. Hebrard, K. Handler, A. Colgan, P. Leekitcharoenphon, G. C. Langridge, A. J. Lohan, B. Loftus, S. Lucchini, D. W. Ussery, C. J. Dorman, N. R. Thomson, J. Vogel, & J. C. D. Hinton. The transcriptional landscape and small RNAs of *Salmonella enterica* serovar typhimurium. *Proc Natl Acad Sci USA*, 109(20):E1277–E1286, apr 2012. doi: 10.1073/pnas.1201061109.

S. Laalami, M. Cavaiuolo, S. Roque, C. Chagneau, & H. Putzer. *Escherichia coli* RNase E can efficiently replace RNase Y in *Bacillus subtilis*. 49(8):4643–4654, mar 2021. doi: 10.1093/nar/gkab216.

A.-L. Lécrivain, A. L. Rhun, T. T. Renault, R. Ahmed-Begrich, K. Hahnke, & E. Charpentier. In vivo 3′-to-5′ exoribonuclease targetomes of *Streptococcus pyogenes*. *Proc Natl Acad Sci*, 115(46):11814–11819, oct 2018. doi: 10.1073/pnas.1809663115.

K. Lee, X. Zhan, J. Gao, J. Qiu, Y. Feng, R. Meganathan, S. N. Cohen, & G. Georgiou. RraA. A protein inhibitor of RNase E activity that globally modulates RNA abundance in *E. coli*. *Cell*, 114(5):623–634, sep 2003. doi: 10.1016/j.cell.2003.08.003.

T. Lee & A. L. Feig. The RNA binding protein hfq interacts specifically with tRNAs. *RNA*, 14 (3):514–523, jan 2008. doi: 10.1261/rna.531408.

K. Lewis. Persister cells. *Annu Rev Microbiol*, 64(1):357–372, oct 2010. doi: 10.1146/annurev.micro.112408.134306.

Z. Li & M. Deutscher. The role of individual exoribonucleases in processing at the 3' end of *Escherichia coli* tRNA precursors. *J Biol Chem*, 269(8):6064–71, Feb. 1994.

W. Liang & M. P. Deutscher. A novel mechanism for ribonuclease regulation: transfer-messenger RNA (tmRNA) and its associated protein SmpB regulate the stability of RNase R. *Journal of Biological Chemistry*, 285(38):29054–29058, sep 2010. doi: 10.1074/jbc.c110.168641.

W. Liang & M. P. Deutscher. Post-translational modification of RNase R is regulated by stress-dependent reduction in the acetylating enzyme Pka (YfiQ). *RNA*, 18(1):37–41, nov 2011. doi: 10.1261/rna.030213.111.

W. Liang, A. Malhotra, & M. P. Deutscher. Acetylation regulates the stability of a bacterial protein: Growth stage-dependent modification of RNase R. *Molecular Cell*, 44(1):160–166, oct 2011. doi: 10.1016/j.molcel.2011.06.037.

S. Lin-Chao, N.-T. Chiou, & G. Schuster. The PNPase, exosome and RNA helicases as the building components of evolutionarily-conserved RNA degradation machines. *Biomed Sci*, 14(4): 523–532, may 2007. doi: 10.1007/s11373-007-9178-y.

G. A. Mackie. Ribonuclease E is a 5′-end-dependent endonuclease. *Nature*, 395(6703): 720–724, oct 1998. doi: 10.1038/27246.

G. A. Mackie. RNase E: at the interface of bacterial RNA processing and decay. *Nature Reviews Microbiology*, 11(1):45–57, dec 2012. doi: 10.1038/nrmicro2930.

B. Magasanik. Regulation of transcription of the *glnALG* operon of *Escherichia coli* by protein phosphorylation. *Biochimie*, 71(9-10):1005–1012, sep 1989. doi: 10.1016/0300-9084(89)90104-1.

S. A. Mahmoud & P. Chien. Regulated proteolysis in bacteria. *Annual Review of Biochemistry*, 87 (1):677–696, jun 2018. doi: 10.1146/annurev-biochem-062917-012848.

N. Majdalani, C. Cunning, D. Sledjeski, T. Elliott, & S. Gottesman. DsrA RNA regulates translation of RpoS message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription. *Proc Natl Acad Sci U S A*, 95(21):12462–12467, oct 1998. doi: 10.1073/pnas.95.21.12462.

N. N. Mank, B. A. Berghoff, Y. N. Hermanns, & G. Klug. Regulation of bacterial photosynthesis genes by the small noncoding RNA PcrZ. *Proceedings of the National Academy of Sciences*, 109(40):16306–16311, sep 2012. doi: 10.1073/pnas.1207067109.

K. McDowall, S. Lin-Chao, & S. Cohen. A+U content rather than a particular nucleotide order determines the specificity of RNase E cleavage. *Journal of Biological Chemistry*, 269 (14):10790–10796, 1994. ISSN 0021-9258. doi: https://doi.org/10.1016/S0021-9258(17)34129-7. URL `https://www.sciencedirect.com/science/article/pii/S0021925817341297`.

K. J. McDowall & S. N. Cohen. The N-terminal domain of the *rne* gene product has RNase E activity and is non-overlapping with the arginine-rich RNA-binding site. *Journal of Molecular Biology*, 255(3):349–355, jan 1996. doi: 10.1006/jmbi.1996.0027.

F. Mika & R. Hengge. Small RNAs in the control of RpoS, CsgD, and biofilm architecture of *Escherichia* coli. *RNA Biol*, 11(5):494–507, apr 2014. doi: 10.4161/rna.28867.

A. S. Mironov, I. Gusarov, R. Rafikov, L. E. Lopez, K. Shatalin, R. A. Kreneva, D. A. Perumov, & E. Nudler. Sensing small molecules by nascent RNA. *Cell*, 111(5):747–756, nov 2002. doi: 10.1016/s0092-8674(02)01134-0.

M. Miyakoshi, Y. Chao, & J. Vogel. Regulatory small RNAs from the 3′ regions of bacterial mRNAs. *Current Opinion in Microbiology*, 24: 132–139, apr 2015. doi: 10.1016/j.mib.2015.01.013.

B. K. Mohanty & S. R. Kushner. Analysis of the function of *Escherichia coli* poly(a) polymerase I in RNA metabolism. *Mol Microbiol*, 34(5):1094–1108, dec 1999. doi: 10.1046/j.1365-2958.1999.01673.x.

B. K. Mohanty & S. R. Kushner. Polynucleotide phosphorylase functions both as a 3′ to 5′ exonuclease and a poly(A) polymerase in *Escherichia coli*. *Proc Natl Acad Sci USA*, 97(22):11966–11971, oct 2000. doi: 10.1073/pnas.220295997.

B. K. Mohanty & S. R. Kushner. Genomic analysis in *Escherichia coli* demonstrates differential roles for polynucleotide phosphorylase and RNase II in mRNA abundance and decay. *Mol Microbiol*, 50(2):645–658, sep 2003. doi: 10.1046/j.1365-2958.2003.03724.x.

M. Morita, M. Kanemori, H. Yanagi, & T. Yura. Heat-induced synthesis of σ 32 in *Escherichia coli*: Structural and functional dissection of *rpoH* mRNA secondary structure. *J Bacteriol*, 181(2):401–410, jan 1999. doi: 10.1128/jb.181.2.401-410.1999.

T. Morita, H. Kawamoto, T. Mizota, T. Inada, & H. Aiba. Enolase in the RNA degradosome plays a crucial role in the rapid decay of glucose transporter mRNA in the response to phosphosugar stress in *Escherichia coli*. *Mol Microbiol*, 54(4):1063–1075, oct 2004. doi: 10.1111/j.1365-2958.2004.04329.x.

E. A. Mudd & C. F. Higgins. *Escherichia coli* endoribonuclease RNase E: autoregulation of expression and site-specific cleavage of mRNA. 9(3):557–568, aug 1993. doi: 10.1111/j.1365-2958.1993.tb01716.x.

P. Müller, M. Gimpel, T. Wildenhain, & S. Brantl. A new role for CsrA: promotion of complex formation between an sRNA and its mRNA target in *Bacillus subtilis*. *RNA Biol*, 16(7):972–987, may 2019. doi: 10.1080/15476286.2019.1605811.

S. Nakamura & T. Minamino. Flagella-driven motility of bacteria. *Biomolecules*, 9(7):279, jul 2019. doi: 10.3390/biom9070279.

F. Narberhaus, T. Waldminghaus, & S. Chowdhury. RNA thermometers. *FEMS Microbiol Rev*, 30(1):3–16, jan 2006. doi: 10.1111/j.1574-6976.2005.004.x.

M. Nitzan, R. Rehani, & H. Margalit. Integration of bacterial small RNAs in regulatory networks. *Annu Rev Biophys*, 46(1):131–148, may 2017. doi: 10.1146/annurev-biophys-070816-034058.

M. Olejniczak & G. Storz. ProQ/FinO-domain proteins: another ubiquitous family of RNA matchmakers? *Mol Microbiol*, 104(6):905–915, may 2017. doi: 10.1111/mmi.13679.

H. Otaka, H. Ishikawa, T. Morita, & H. Aiba. Poly-U tail of Rho-independent terminator of bacterial small RNAs is essential for Hfq action. *Proc Natl Acad Sci U S A*, 108(32):13059–13064, jul 2011. doi: 10.1073/pnas.1107050108.

M. C. Ow & S. R. Kushner. Initiation of tRNA maturation by RNase E is essential for cell viability in *E. coli*. 16(9):1102–1115, may 2002. doi: 10.1101/gad.983502.

R. Pandey, D. Flockerzi, M. J. Hauser, & R. Straube. Modeling the light- and redox-dependent interaction of PpsR/AppA in *Rhodobacter sphaeroides*. *Biophysical Journal*, 100(10):2347–2355, may 2011. doi: 10.1016/j.bpj.2011.04.017.

K. Papenfort & C. K. Vanderpool. Target activation by regulatory RNAs in bacteria. *FEMS Microbiol Rev*, 39(3):362–378, may 2015. doi: 10.1093/femsre/fuv016.

H. Park, H. Yakhnin, M. Connolly, T. Romeo, & P. Babitzke. CsrA participates in a PNPase autoregulatory mechanism by selectively repressing translation of *pnp* transcripts that have been previously processed by RNase III and PNPase. *J Bacteriol*, 197(24):3751–3759, dec 2015. doi: 10.1128/jb.00721-15.

PAXdb. Protein abundance database - abundance of PNPase in *E. coli*, Oct. 2021. URL `https://pax-db.org/protein/6883773`.

T. Peng, B. A. Berghoff, J.-I. Oh, L. Weber, J. Schirmer, J. Schwarz, J. Glaeser, & G. Klug. Regulation of a polyamine transporter by the conserved 3′ UTR-derived sRNA SorX confers resistance to singlet oxygen and organic hydroperoxides in *Rhodobacter sphaeroides*. *RNA Biology*, 13(10):988–999, aug 2016. doi: 10.1080/15476286.2016.1212152.

N. Perez, J. Treviño, Z. Liu, S. C. M. Ho, P. Babitzke, & P. Sumby. A genome-wide analysis of small regulatory RNAs in the human pathogen group A *Streptococcus*. *PLoS One*, 4(11):e7668, nov 2009. doi: 10.1371/journal.pone.0007668.

V. Pfeiffer, K. Papenfort, S. Lucchini, J. C. D. Hinton, & J. Vogel. Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nat Struct Mol Biol*, 16(8):840–846, jul 2009. doi: 10.1038/nsmb.1631.

A. H. Potts, C. A. Vakulskas, A. Pannuri, H. Yakhnin, P. Babitzke, & T. Romeo. Global role of the bacterial post-transcriptional regulator CsrA revealed by integrated transcriptomics. *Nat Commun*, 8(1), nov 2017. doi: 10.1038/s41467-017-01613-1.

H. Preis, R. A. Eckart, R. K. Gudipati, N. Heidrich, & S. Brantl. CodY activates transcription of a small RNA in *Bacillus subtilis*. *J Bacteriol.*, 191(17):5446–5457, sep 2009. doi: 10.1128/jb.00602-09.

A. Prud'homme-Généreux, R. K. Beran, I. Iost, C. S. Ramey, G. A. Mackie, & R. W. Simons. Physical and functional interactions among RNase E, polynucleotide phosphorylase and the cold-shock protein, CsdA: evidence for a 'cold shock degradosome'. *Mol Microbiol*, 54 (5):1409–1421, oct 2004. doi: 10.1111/j.1365-2958.2004.04360.x.

B. Py, H. Causton, E. A. Mudd, & C. F. Higgins. A protein complex mediating mRNA degradation in *Escherichia coli*. 14(4):717–729, nov 1994. doi: 10.1111/j.1365-2958.1994.tb01309.x.

B. Py, C. F. Higgins, H. M. Krisch, & A. J. Carpousis. A DEAD-box RNA helicase in the *Escherichia coli* RNA degradosome. 381(6578): 169–172, may 1996. doi: 10.1038/381169a0.

A. P. Quendera, A. F. Seixas, R. F. dos Santos, I. Santos, J. P. N. Silva, C. M. Arraiano, & J. M. Andrade. RNA-binding proteins driving the regulatory activity of small non-coding RNAs in bacteria. *Frontiers in Molecular Biosciences*, 7, may 2020. doi: 10.3389/fmolb.2020.00078.

M. Rabhi, O. Espéli, A. Schwartz, B. Cayrol, A. R. Rahmouni, V. Arluison, & M. Boudvillain. The Sm-like RNA chaperone Hfq mediates transcription antitermination at Rho-dependent terminators. *EMBO J*, 30(14):2805–2816, jun 2011. doi: 10.1038/emboj.2011.192.

C. M. Reuscher & G. Klug. Antisense RNA asPcrL regulates expression of photosynthesis genes in *Rhodobacter sphaeroides* by promoting RNase III-dependent turn-over of *puf* mRNA. *RNA Biology*, pages 1–13, jan 2021. doi: 10.1080/15476286.2020.1857520.

T. Rische & G. Klug. The ordered processing of intervening sequences in 23S rRNA of *Rhodobacter sphaeroides* requires RNase J. *RNA Biology*, 9(3):343–350, mar 2012. doi: 10.4161/rna.19433.

N. A. Sabnis, H. Yang, & T. Romeo. Pleiotropic regulation of central carbohydrate metabolism in *Escherichia coli* via the gene *csrA*. *J Biol Chem*, 270(49):29096–29104, dec 1995. doi: 10.1074/jbc.270.49.29096.

D. D. Sarpong & E. R. Murphy. RNA regulated toxin-antitoxin systems in pathogenic bacteria. *Frontiers in Cellular and Infection Microbiology*, 11, may 2021. doi: 10.3389/fcimb.2021.661026.

C. Sauter, J. Basquin, & D. Suck. Sm-like proteins in Eubacteria: the crystal structure of the Hfq protein from *Escherichia coli*. *Nucleic Acids Res*, 31(14):4091–4098, jul 2003. doi: 10.1093/nar/gkg480.

D. J. Schu, A. Zhang, S. Gottesman, & G. Storz. Alternative Hfq-sRNA interaction modes dictate alternative mRNA recognition. *EMBO J*, 34(20):2557–2573, sep 2015. doi: 10.15252/embj.201591569.

A. Schuck, A. Diwa, & J. G. Belasco. RNase E autoregulates its synthesis in *Escherichia coli* by binding directly to a stem-loop in the *rne* 5′ untranslated region. *Mol Microbiol*, 72 (2):470–478, apr 2009. doi: 10.1111/j.1365-2958.2009.06662.x.

A. W. Serio, T. Keepers, L. Andrews, & K. M. Krause. Aminoglycoside revival: Review of a historically important class of antimicrobials undergoing rejuvenation. *EcoSal Plus*, 8(1), feb 2018. doi: 10.1128/ecosalplus.esp-0002-2018.

N. Sesto, M. Touchon, J. M. Andrade, J. Kondo, E. P. C. Rocha, C. M. Arraiano, C. Archambaud, É. Westhof, P. Romby, & P. Cossart. A PNPase dependent CRISPR system in *Listeria*. *PLoS Genet*, 10(1):e1004065, jan 2014. doi: 10.1371/journal.pgen.1004065.

K. Shahbabian, A. Jamalli, L. Zig, & H. Putzer. RNase Y, a novel endoribonuclease, initiates riboswitch turnover in *Bacillus subtilis*. 28(22): 3523–3533, sep 2009. doi: 10.1038/emboj.2009.283.

Z. Shi, W.-Z. Yang, S. Lin-Chao, K.-F. Chak, & H. S. Yuan. Crystal structure of *Escherichia coli* PNPase: Central channel residues are involved in processive RNA degradation. *RNA Biology*, 14(11):2361–2371, sep 2008. doi: 10.1261/rna.1244308.

H. Shimada, K. Iba, & K. ichiro Takamiya. Blue-light irradiation reduces the expression of *puf* and *puc* of *Rhodobacter sphaeroides* under semi-aerobic conditions. *Plant and Cell Physiology*, jun 1992. doi: 10.1093/oxfordjournals.pcp.a078276.

A. Smirnov, K. U. Förstner, E. Holmqvist, A. Otto, R. Günster, D. Becher, R. Reinhardt, & J. Vogel. Grad-seq guides the discovery of ProQ as a major small RNA-binding protein. *Proc Natl Acad Sci U S A*, 113(41):11591–11596, sep 2016. doi: 10.1073/pnas.1609981113.

J. Soppa. Protein acetylation in archaea, bacteria, and eukaryotes. *Archaea*, 2010:1–9, 2010. doi: 10.1155/2010/820681.

S. Sulthana & M. P. Deutscher. Multiple exoribonucleases catalyze maturation of the 3′ terminus of 16S ribosomal RNA (rRNA). *J Biol Chem*, 288(18):12574–12579, may 2013. doi: 10.1074/jbc.c113.459172.

X. Sun, I. Zhulin, & R. M. Wartell. Predicted structure and phyletic distribution of the RNA-binding protein Hfq. *Nucleic Acids Research*, 30(17):3662–3671, sep 2002. doi: 10.1093/nar/gkf508.

M. K. Thomason, M. Voichek, D. Dar, V. Addis, D. Fitzgerald, S. Gottesman, R. Sorek, & E. P. Greenberg. A *rhlI* 5′ UTR-derived sRNA regulates RhlR-dependent quorum sensing in *Pseudomonas aeruginosa*. *mBio*, 10(5), oct 2019. doi: 10.1128/mbio.02253-19.

B. J. Tucker & R. R. Breaker. Riboswitches as versatile gene control elements. *Curr Opin Struct Biol*, 15(3):342–348, jun 2005. doi: 10.1016/j.sbi.2005.05.003.

C. Unoson & E. G. H. Wagner. A small SOS-induced toxin is targeted against the inner membrane in *Escherichia coli*. 70(1):258–270, aug 2008. doi: 10.1111/j.1365-2958.2008.06416.x.

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

C. B. van Niel. The culture, general physiology, morphology, and classification of the non-sulfur purple and brown bacteria. *Microbiology and Molecular Biology Reviews*, 8(1):1–118, 1944. ISSN 0005-3678. URL `https://mmbr.asm.org/content/8/1/1`.

J. Vogel. A rough guide to the non-coding RNA world of *Salmonella*. *Mol Microbiol.*, 71(1):1–11, jan 2009. doi: 10.1111/j.1365-2958.2008.06505.x.

J. Vogel & K. Papenfort. Small non-coding RNAs and the bacterial outer membrane. *Current Opinion in Microbiology*, 9(6):605–611, dec 2006. doi: 10.1016/j.mib.2006.10.006.

J. Vogel, L. Argaman, E. H. Wagner, & S. Altuvia. The small RNA IstR inhibits synthesis of an SOS-induced toxic peptide. *Curr Biol*, 14(24):2271–2276, dec 2004. doi: 10.1016/j.cub.2004.12.003.

E. G. H. Wagner. Cycling of RNAs on Hfq. *RNA Biol*, 10(4):619–626, apr 2013. doi: 10.4161/rna.24044.

E. G. H. Wagner & P. Romby. Small RNAs in bacteria and archaea. In *Advances in Genetics*, pages 133–208. Elsevier, 2015. doi: 10.1016/bs.adgen.2015.05.001.

E. G. H. Wagner & C. Unoson. The toxin-antitoxin system tisB-istR1. *RNA Biology*, 9(12):1513–1519, dec 2012. doi: 10.4161/rna.22578.

T. Waldminghaus, N. Heidrich, S. Brantl, & F. Narberhaus. FourU: a novel type of RNA thermometer in *Salmonella*. *Molecular Microbiology*, 65(2):413–424, jul 2007. doi: 10.1111/j.1365-2958.2007.05794.x.

C. Wang, Y. Chao, G. Matera, Q. Gao, & J. Vogel. The conserved 3′ UTR-derived small RNA NarS mediates mRNA crossregulation during nitrate respiration. *Nucleic Acids Research*, 48(4):2126–2143, dec 2019. doi: 10.1093/nar/gkz1168.

J. X. Wang & R. R. Breaker. Riboswitches that sense S-adenosylmethionine and S-adenosylhomocysteine. *Biochem Cell Biol*, 86(2):157–168, apr 2008. doi: 10.1139/o08-008.

R. Washburn & M. Gottesman. Regulation of transcription elongation and termination. *Biomolecules*, 5(2):1063–1078, may 2015. doi: 10.3390/biom5021063.

K. M. Wassarman, A. Zhang, & G. Storz. Small RNAs in *Escherichia coli*. *Trends in Microbiology*, 7(1):37–45, jan 1999. doi: 10.1016/s0966-842x(98)01379-1.

L. Weber, C. Thoelken, M. Volk, B. Remes, M. Lechner, & G. Klug. The conserved *dcw* gene cluster of *R. sphaeroides* is preceded by an uncommonly extended 5′ leader featuring the sRNA UpsM. *PLOS ONE*, 11(11):e0165694, nov 2016. doi: 10.1371/journal.pone.0165694.

F. H. Westheimer. Pseudo-rotation in the hydrolysis of phosphate esters. *Accounts of Chemical Research*, 1(3):70–78, mar 1968. doi: 10.1021/ar50003a002.

C. Woese, E. Stackebrandt, W. Weisburg, B. Paster, M. Madigan, V. Fowler, C. Hahn, P. Blanz, R. Gupta, K. Nealson, & G. Fox. The phylogeny of purple bacteria: The alpha subdivision. *Systematic and Applied Microbiology*, 5(3):315–326, oct 1984. doi: 10.1016/s0723-2020(84)80034-x.

J. Wong-Ng, A. Celani, & M. Vergassola. Exploring the function of bacterial chemotaxis. *Current Opinion in Microbiology*, 45:16–21, oct 2018. doi: 10.1016/j.mib.2018.01.010.

T. T. Wyatt, H. A. Wösten, & J. Dijksterhuis. *Fungal Spores for Dispersion in Space and Time*. Elsevier, 2013. doi: 10.1016/b978-0-12-407672-3.00002-2.

K. Yamanaka, M. Mitta, & M. Inouye. Mutation analysis of the 5′ untranslated region of the cold shock *cspA* mRNA of *Escherichia coli*. *J Bacteriol*, 181(20):6284–6291, oct 1999. doi: 10.1128/jb.181.20.6284-6291.1999.

H. Yang, M. Y. Liu, & T. Romeo. Coordinate genetic regulation of glycogen catabolism and biosynthesis in *Escherichia coli* via the CsrA gene product. *J Bacteriol*, 178(4):1012–1017, feb 1996. doi: 10.1128/jb.178.4.1012-1017.1996.

W. Yin, Y. Wang, L. Liu, & J. He. Biofilms: The microbial “protective clothing” in extreme environments. *International Journal of Molecular Sciences*, 20(14):3423, jul 2019. doi: 10.3390/ijms20143423.

K. Zahn, M. Inui, & H. Yukawa. Divergent mechanisms of 5' 23S rRNA IVS processing in the alpha-proteobacteria. *Nucleic Acids Research*, 28(23):4623–4633, dec 2000. doi: 10.1093/nar/28.23.4623.

Z. Zhao, T. Peng, J.-I. Oh, J. Glaeser, L. Weber, Q. Li, & G. Klug. A response regulator of the OmpR family is part of the regulatory network controlling the oxidative stress response of *Rhodobacter sphaeroides*. *Environmental Microbiology Reports*, 11(2):118–128, dec 2018. doi: 10.1111/1758-2229.12718.

Chapter 2

Impact of PNPase on the transcriptome of *Rhodobacter sphaeroides* and its cooperation with RNase III and RNase E

Spanka *et al. BMC Genomics* (2021) 22:106
https://doi.org/10.1186/s12864-021-07409-4

BMC Genomics

RESEARCH ARTICLE Open Access

Impact of PNPase on the transcriptome of *Rhodobacter sphaeroides* and its cooperation with RNase III and RNase E

Daniel-Timon Spanka, Carina Maria Reuscher and Gabriele Klug*

Abstract

Background: The polynucleotide phosphorylase (PNPase) is conserved among both Gram-positive and Gram-negative bacteria. As a core part of the *Escherichia coli* degradosome, PNPase is involved in maintaining proper RNA levels within the bacterial cell. It plays a major role in RNA homeostasis and decay by acting as a 3′-to-5′ exoribonuclease. Furthermore, PNPase can catalyze the reverse reaction by elongating RNA molecules in 5′-to-3′ end direction which has a destabilizing effect on the prolonged RNA molecule. RNA degradation is often initiated by an endonucleolytic cleavage, followed by exoribonucleolytic decay from the new 3′ end.

Results: The PNPase mutant from the facultative phototrophic *Rhodobacter sphaeroides* exhibits several phenotypical characteristics, including diminished adaption to low temperature, reduced resistance to organic peroxide induced stress and altered growth behavior. The transcriptome composition differs in the *pnp* mutant strain, resulting in a decreased abundance of most tRNAs and rRNAs. In addition, PNPase has a major influence on the half-lives of several regulatory sRNAs and can have both a stabilizing or a destabilizing effect. Moreover, we globally identified and compared differential RNA 3′ ends in RNA NGS sequencing data obtained from PNPase, RNase E and RNase III mutants for the first time in a Gram-negative organism. The genome wide RNA 3′ end analysis revealed that 885 3′ ends are degraded by PNPase. A fair percentage of these RNA 3′ ends was also identified at the same genomic position in RNase E or RNase III mutant strains.

Conclusion: The PNPase has a major influence on RNA processing and maturation and thus modulates the transcriptome of *R. sphaeroides*. This includes sRNAs, emphasizing the role of PNPase in cellular homeostasis and its importance in regulatory networks. The global 3′ end analysis indicates a sequential RNA processing: 5.9% of all RNase E-dependent and 9.7% of all RNase III-dependent RNA 3′ ends are subsequently degraded by PNPase. Moreover, we provide a modular pipeline which greatly facilitates the identification of RNA 5′/3′ ends. It is publicly available on GitHub and is distributed under ICS license.

Keywords: Alphaproteobacteria, *Rhodobacter sphaeroides*, PNPase, Exoribonuclease, Transcriptomics, RNA 3′ end identification, RNase E, XPEAP

* Correspondence: Gabriele.Klug@mikro.bio.uni-giessen.de
Institute of Microbiology and Molecular Biology, Justus Liebig University Giessen, IFZ, Giessen, Germany

Background

Prokaryotes populate nearly every imaginable habitat. In contrast to higher multicellular eukaryotes, they are directly exposed to all types of environmental stress. Since escaping is not an option, prokaryotes need mechanisms to quickly adapt to their changing surrounding. This can be achieved by modifying the transcriptome and/or the proteome. One essential mechanism in bacterial adaptation is to exchange the sigma factor, a subunit of the RNA polymerase. Alternative sigma factors target different DNA sequences and thus activate the expression of a specific set of genes. This activates transcription of genes needed for the cell to deal with the present growth condition [1, 2].

Besides and in addition to the transcriptional initiation, posttranscriptional regulation plays a major role in bacterial adaptation [3]. During the past decades, more and more bacterial non-coding RNAs were discovered and found to be involved in various posttranscriptional regulatory networks (reviewed in [4]).

Current studies documented, that the prokaryotic transcriptome is heavily influenced by processing and maturation reactions mediated by the endoribonuclease E [5–7]. Endonucleolytic RNA cleavage by RNase E is mostly followed by further degradation. 3′-to-5′ exonucleases can attack the new 3′ end and RNase E can bind to the monophosphorylated new 5′ end and promote further endonucleolytic degradation in 5′-to-3′ direction. Secondary structures can protect against 3′-to-5′ degradation [8] and can also impede RNase E mediated 5′-to-3′ processing [9]. A key player during RNA turnover is a multicomponent degradation complex called the degradosome. In *Escherichia coli*, this complex is composed of RNase E, which serves as catalytic and scaffold protein, a DEAD box RNA helicase (RhlB), an exoribonuclease (polynucleotide phosphorylase, PNPase) and an enolase (reviewed in [10]). In contrast to that, studies of the α-proteobacterium *Rhodobacter capsulatus* suggest, that PNPase is most likely not part of its degradosome [11] that in addition to RNase E includes 2 deadbox helicases and the transcriptional termination factor Rho. Moreover, the composition of the *R. capsulatus* degradosome changes in response to altering environmental conditions [12].

The PNPase is a trimer comprising three Pnp monomers that form a ring-like structure. In *E. coli*, each monomer consists of two RNase PH-like domains and a KH and S1 domain [13, 14]. A deletion of *pnp* is possible in *E. coli*, whereas a double knockout of PNPase and RNase II is not viable [15]. The *R. sphaeroides* genome does not harbour an RNase II gene and it is not possible to delete the *pnp* gene. The same effect was also observed in at least one other organism, *Pseudomonas aeruginosa* [16]. Removal of the RNA-binding KH/S1 domains of PNPase leads to an eightfold reduced binding affinity to RNA in *E. coli* [14]. Further, the trimer formation is less stable, which leads to a wider central channel [14]. PNPase not only serves as an important 3′-to-5′ exoribonuclease involved in mRNA degradation but also in tRNA processing and degradation [15, 17]. Besides that, PNPase can also prolong RNA molecules in 5′-to-3′ direction using nucleotide diphosphates present in the cytoplasm. This tail allows recruitment of single-strand dependent exoribonucleases thus reducing the RNA half-life [18]. Since PNPase is an enzyme with such a widespread influence on the cellular RNAs, the *pnp* expression has to be tightly regulated. Similar to *rne* mRNA levels, *pnp* mRNA levels are balanced in an autoregulatory manner. The endoribonuclease RNase III first cleaves a stem-loop located in the *pnp* leader sequence. The newly generated 3′ end in this RNA duplex is then targeted and degraded by PNPase. Ultimately this leads to reduced *pnp* mRNA stability [19, 20]. Besides PNPase, several other exoribonucleases are likely involved in RNA processing, maturation and degradation in the α-proteobacterium *R. sphaeroides*. These are the RNase R, RNase D and RNase PH which catalyze mainly tRNA and rRNA processing reactions and all act in 3′-to-5′ direction [21–23]. In addition, RNase J1 is responsible for the maturation of the 23S rRNA and very few other transcripts [24, 25]. In contrast to the other RNases, it processes RNA molecules in 5′-to-3′ end direction [26]. The endoribonucleases RNase E, III and G (homolog of RNase E) are mainly responsible for RNA maturation and turnover [7, 27, 28].

In order to understand bacterial adaptation, it is important to elucidate the complex interplay between different RNases and how they sequentially process RNA molecules. A common way for degradation of mRNA and maturation of RNA precursors requires two steps: First, the endoribonucleases III, E or P catalyse the endonucleolytic cleavage of the RNA molecule. Second, the enzymes PNPase, RNase R, RNase PH or RNase II can further degrade the RNA fragments from 3′-to-5′-direction (reviewed in [29, 30]). In both steps, new RNA 3′ ends are generated (Fig. 1a+b). Recent studies in the Gram-positive human pathogen *Streptococcus pyogenes* illustrate how initial processing by endoribonuclease Y is followed by further maturation reactions catalyzed by the exoribonucleases PNPase and RNase R [31]. The other principal mechanisms for RNA 3′ end generation are transcription termination by RNA polymerase and the 3′-terminal elongation mediated by PNPase (Fig. 1c + d).

In this study, we report that in the *Rhodobacter sphaeroides pnp* mutant strain several physiological characteristics are affected by the deletion of the KH and S1 domains, including growth behavior and

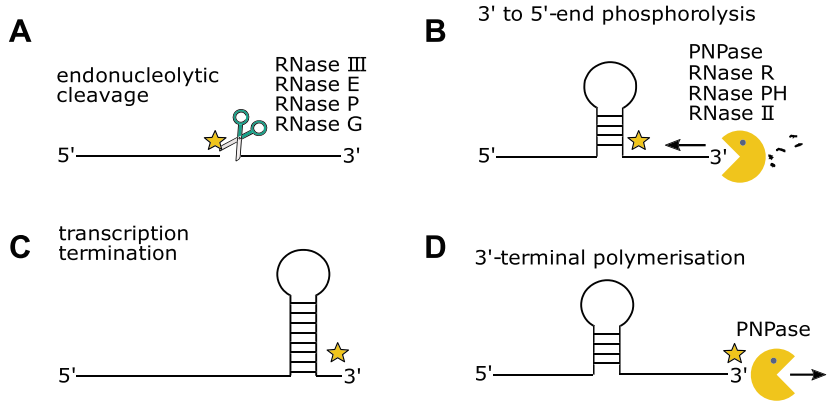


Fig. 1 Generation of RNA 3′ ends in bacteria and action of PNPase. **a** RNA 3′-OH ends (highlighted by yellow stars) can be generated via endonucleolytic cleavage by RNase III/E/P/G, **b** by 3′-to-5′ degradation mediated by PNPase and RNase R/PH/II, or **c** by transcriptional termination. **d** PNPase can also produce new 3′-OH ends by a 3′-terminal oligonucleotide polymerase reaction

pigmentation. In a global approach, we further used RNA-Seq data and identified all RNA 3′ ends that are PNPase-, RNase III- or RNase E-dependent. Intersection analysis sheds light on important processing events by the analyzed RNases that shape the transcriptome in a cooperative manner. Finally, we could demonstrate, that homeostasis of the regulatory sRNAs CcsR1–4 rely on initial RNase E cleavage followed by PNPase degradation.

Methods

Bacterial strains and growth conditions

The strains used in this study are listed in Table S1 [32]. Microaerobic *Rhodobacter sphaeroides* cultures (dissolved oxygen concentration of 25–30 μM) were cultivated in 50 ml Erlenmeyer flasks filled with 40 ml malate minimal medium at 32 °C under continuous shaking at 140 rpm in the dark [33]. To perform phototrophic cultivation, Metplat bottles were completely filled and sealed. Afterwards the cultures were constantly exposed to white light with an intensity of 40 W/m^2 at 32 °C.

Construction of *pnp* KH and S1 deletion strain

The deletion of the *pnp* C-terminal KH and S1 domains in *Rhodobacter sphaeroides* 2.4.1 was carried out by homologous recombination. Since *pnp* is essential in *R. sphaeroides*, only the RNA binding domains KH/S1 were replaced by a gentamicin resistance gene on the chromosome. The up and down fragments were generated using the primer pairs pnpFragAfw/pnpFragArev (5′-gaaTTCAAGAAGCTGGAAAGCTCGAT, 5′-ggatcctcAGGTTTCCACGATCTCGCGG, 870 bp) and pnpFragBfw/pnpFragBrev (5′-ggaTCCGTCTCGGCATGAAGATG, 5′-aagcTTCTCGTCCGAAGACGTGCTG, 631 bp), introducing an in-frame TGA stop codon within the reverse primer of the up fragment (see underlined bases in primer pnpFragArev). The stop codon is located at position 1755 in the *pnp* gene and leads to translation termination directly upstream of the deleted KH/S1 region (Fig. 2a). Both fragments were cloned in pPHU281 using EcoRI/BamHI and BamHI/HindIII cleavage sites. The gentamicin resistance gene was taken from pPHU45Ω and inserted between the up and down fragment on the plasmid with BamHI. The final construct was transformed to *E. coli* strain S17–1 and subsequently transferred to *Rhodobacter sphaeroides* 2.4.1 by diparental conjugation. The conjugants were selected on malate minimal agar containing 10 μg/ml gentamicin.

Measurement of bacteriochlorophyll and carotenoids

The determination of bacteriochlorophyll and carotenoid concentrations was performed as described in [34]. The calculations rely on the extinction coefficients (76 $mM^{-1} \cdot cm^{-1}$ for bacteriochlorophyll *a*, 128 $mM^{-1} \cdot cm^{-1}$ for carotenoids) published in [35].

Spot assay

A volume of 5 μl taken from a liquid culture during the exponential growth phase was placed on malate minimal agar plates. The plates were first incubated at 4 °C or 42 °C for 1 day and then shifted to 32 °C and cultivated for three more days. To test resistance to organic peroxides, *tert*-butyl hydroperoxide (tBOOH) was added to the agar (300 μM final concentration). That plate as well as the control without any tBOOH were subsequently incubated at 32 °C for 3 days.

Determination of RNA half-life

Rhodobacter cultures were cultivated under microaerobic conditions. During the exponential phase, the sample

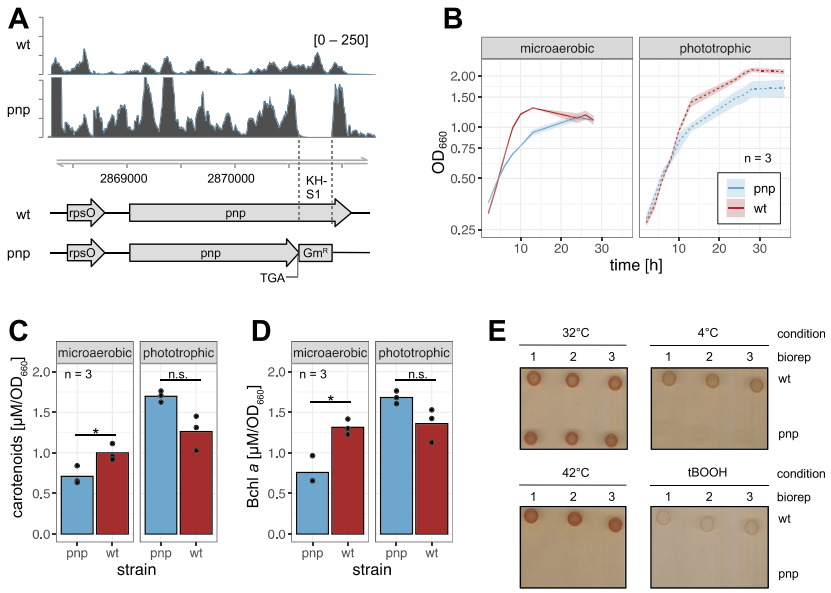


Fig. 2 The *pnp* mutant and the wild type strain differ in growth behaviour, pigmentation as well as in growth under different temperatures and under organic peroxide stress. **a** Schematic overview of the *pnp* operon. In the *pnp* mutant, the KH-S1 domains were deleted and substituted with a gentamicin resistance gene. A stop codon was inserted at the end of the remaining *pnp* coding region. Upper panels show the RNA read coverage in the wild type and *pnp* mutant strain. **b** The *pnp* mutant grows slower than the wild type under microaerobic cultivation and does not reach the wild type optical density during stationary phase when cultivated under phototrophic conditions. Red: wild type; blue: *pnp* mutant; $n = 3$. **c, d** Exponentially growing *pnp* mutant cultures exhibit reduced carotenoid and bacteriochlorophyll *a* (Bchl *a*) concentrations under microaerobic conditions in comparison to the wild type strain. Phototrophically cultivated, the pigment concentrations are increased in the mutant. The *p*-values were calculated using two-sided Student's *t*-test (*: $p < 0.05$; n.s.: not significant). **e** On solid malate minimal agar, the growth of the *pnp* mutant strain is strongly impaired when the plates are incubated at 4 °C or 42 °C. The organic peroxide tBOOH (300 µM final concentration) diminishes growth of the wild type but prevents growth of the *pnp* mutant strain. Biological triplicates are shown for each growth condition

t_0 was harvested. Immediately after that the transcription inhibitor rifampicin was added to a final concentration of 0.2 mg/ml. The following samples were taken at the indicated time points. All cells were harvested on ice and total RNA was isolated and blotted as described below.

Northern blot analysis

The hot phenol method was used to isolate total RNA [36]. The procedure was followed by a DNase treatment (Invitrogen #AM1907) according to the manufacturer's protocol to digest remaining DNA fragments. The electrophoretic separation in a gel and subsequent Northern blot analysis was performed as described earlier [37]. The oligonucleotide end-labelling was performed using T4 polynucleotide Kinase (T4-PNK, Thermo Scientific) according to the manufacturer's instructions. Radioactive [γ^{32}P]-ATP was obtained from Hartmann Analytic (SRP-301), the oligonucleotides used for labeling are listed in Table S2 in Additional file 1. After overnight incubation with labeled oligonucleotides, the membrane was washed in 5x SSC buffer and exposed to a screen for 1 day. The QuantityOne 1-D Analysis Software (BioRad, version 4.6.6) was used to quantify the signals. All signal intensities were normalized to the corresponding 5S rRNA signal.

Library preparation

Three single colonies were used to inoculate three independent pre-cultures. Every culture was then used to inoculate three main cultures (nine in total). During the exponential growth phase, all three replicates belonging to one biological pre-culture were harvested and pooled. Total RNA was extracted followed by DNase treatment.

RNA quality was checked using a 2100 Bioanalyzer with the RNA 6000 Nano kit (Agilent Technologies). Five hundred nanograms of high quality total RNA were used for the preparation of a cDNA library with the NEBNext Multiplex Small RNA Library Prep kit for Illumina (NEB) in accordance with the manufacturers' instructions with modifications: RNA samples were fragmented with Mg^{2+} at 94 °C for 3 min 15 s using the NEBNext Magnesium RNA Fragmentation Module (NEB) followed by RNA purification with the Zymo

Oligo Clean & Concentrator kit. Fragmented RNA was dephosphorylated at the 3′ end, phosphorylated at the 5′ end and decapped using 10 U T4-PNK +/− 40 nmol ATP and 5 U RppH, respectively (NEB). After each enzymatic treatment RNA was purified with the Zymo Oligo Clean & Concentrator kit. The RNA fragments were ligated for cDNA synthesis to 3′ single-read (SR) adapter and 5′ SR adapter diluted 1:2 with nuclease-free water before use. PCR amplification to add Illumina adaptors and indices to the cDNA was performed for 14 cycles. Barcoded DNA Libraries were purified using magnetic MagSi-NGS^{PREP} Plus beads (AMSBIO) at a 1.5 ratio of beads to sample volume. Libraries were quantified with the Qubit 3.0 Fluorometer (ThermoFisher) and the library quality and size distribution was checked using a 2100 Bioanalyzer with the DNA-1000 kit (Agilent). Sequencing of pooled libraries, spiked with 10% PhiX control library, was performed in single-end mode on the NextSeq 500 platform (Illumina) with the High Output Kit v2.5 (75 Cycles). Demultiplexed FASTQ files were generated with bcl2fastq2 (Illumina). The sequencing data are available at NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo) under the accession number GSE156818.

Bioinformatical analysis

The 3′ end analysis was performed with XPEAP, a pipeline programmed for this study. First, the adapter sequences were removed and all raw reads trimmed for quality with Trim Galore (version 0.6.3). All filtered reads were mapped to the *Rhodobacter sphaeroides* 2.4.1 genome (assembly GCF_000012905.2) using READemption (version 0.4.3 [38];) with the mapper segemehl (version 0.2.0 [39];). The DESeq2 package (version 1.26.0 [40];) was used for the normalization of read counts and the full transcriptome analysis. The results were validated with the R package baySeq (version 2.20.0 [41];) with the gene quantification table obtained from READemption. Coverage generation for both full coverage and 3′ end coverage was done with READemption. The 3′ end coverage files were converted to BED file format with Bedops (version 2.4.37) and filtered. All bases without a minimal read coverage of 10 were rejected. Further, all positions with a signal ratio lower than 5% comparing the 3′ end and the full read coverage were excluded. The nucleotide-wise fold changes were calculated with DESeq2 and all nucleotide positions kept which passed the $\log_2$-fold change cutoff ≤ -1 or $\geq +1$ and exhibited an adjusted *p*-value (Benjamini-Hochberg algorithm) lower than 0.05. All positions within a maximal distance of three nucleotides were merged to one 3′ end with BEDtools' subcommand *merge* (version 2.25.0 [42];), the mean $\log_2$-fold change was computed for every differential 3′ end. BEDtools *intersect* was used to identify genes with overlapping differential 3′ ends and all ends without any overlapping feature were assigned to untranslated regions.

The intersection of the differential 3′ ends between different RNase mutant strains was analyzed with BEDtools *window* using a window size of 1 nt while only matches on the same strand were considered for further analysis. Fisher's exact test was calculated for all intersection files using BEDtools' subcommand *fisher*.

XPEAP is published under ISC license and can be accessed via Zenodo/GitHub (DOI: https://doi.org/10.5281/zenodo.8475, https://github.com/datisp/XPEAP). The raw reads and analyzed data from all experiments are deposited on NCBI Gene Expression Omnibus: PNPase and RNase III mutant strains (NCBI GEO accession number: GSE156818) and thermosensitive RNase E mutant strain (NCBI GEO accession number: GSE71844, published in [7]).

For the 3′ elongation analysis, reads that could not be mapped in end-to-end mode with segemehl were mapped with bowtie2 (version 2.2.6) in local mode with option *–very-sensitive-local* and flags *-f -p 24 –no-hd*. Reads with less than 10 nt matching at the 5′ end were rejected. The sequences following the matching regions were extracted with awk (version 4.1.3).

Results and discussion

Physiological consequences of altered PNPase activity

To analyze the functionality of PNPase in vivo, we designed and cloned a *pnp* mutant strain of *Rhodobacter sphaeroides* 2.4.1. The KH-S1 RNA binding domains were removed and a stop codon was introduced at the end of the remaining coding sequence of *pnp* resulting in a truncated enzyme lacking those domains. The knockout was confirmed via selection on agar containing gentamicin and subsequent RNA sequencing analysis (Fig. 2a). Growth behavior of this strain differed from that of the wild type (Fig. 2b). When cultivated under microaerobic conditions, the growth rate was reduced, but both wild type and mutant finally reached the identical OD_{660}. Under phototrophic conditions the mutant leaves exponential phase earlier than the wild type reaching a lower final OD_{660}. A previous study revealed that reduced RNase E activity strongly impeded phototrophic growth of *R. sphaeroides*, while it had no effect on chemotrophic growth [7]. Moreover, the *pnp* mutant and the parental wild type strain vary in pigment composition (Fig. 2c+d). These differences are strongly dependent on the cultivation conditions: A significantly lower concentration of carotenoids and bacteriochlorophyll *a* was observed in the *pnp* mutant under microaerobic conditions (*p*-values < 0.05), while the *pnp* mutant exhibited repeatedly higher pigment

concentrations under phototrophic conditions. However, this difference was statistically not significant.

In *E. coli, Yersinia enterocolitica* and *Photorhabdus sp.* PNPase plays an important role in the cold shock response due to selective degradation of mRNAs for cold shock proteins at the end of the acclimation phase to low temperature [43–46]. Based on this observation, we decided to test the *R. sphaeroides* strains for their ability to adapt to low and high temperatures. Wild type and *pnp* mutant cells were incubated at 4 °C or 42 °C on agar plates for 1 day and then shifted to an optimal temperature of 32 °C. In both cases growth of the *pnp* mutant was strongly impeded, while the wild type was able to grow at 42 °C and 4 °C (Fig. 2e). Also, in contrast to the wild type, the *pnp* mutant was not able to grow on malate minimal agar containing 300 µM tBOOH, while the wild type showed weak growth. Tertiary butylalcohol is representing organic peroxides that are produced e. g. during photo-oxidative stress.

Our results show that PNPase of *R. sphaeroides* is involved in cold adaptation as other bacterial PNPases and also is strongly impeded in its adaptation to heat. Whether the same molecular mechanisms are responsible for the phenotype as in other bacteria remains to be elucidated. This study for the first time analyses the function of PNPase in a phototrophic bacterium. The effect of PNPase on the bacteriochlorophyll levels and on carotenoid levels depends on growth conditions. Many genes are involved in the formation of photosynthetic complexes and it is not possible to correlate these phenotypic changes to specific changes of the transcriptome. We observed before that a temperature-sensitive variant of RNase E had little effect on growth under microaerobic conditions but strongly impeded phototrophic growth [7]. For the PNPase mutant we observed slower growth under both conditions, phototrophic growth was less affected, in contrast to the *rne* mutant.

PNPase modulates the transcriptome of *R. sphaeroides*

PNPase is an enzyme involved in many RNA processing reactions, and a global influence on the transcriptome can be expected as also shown for the Gram-positive *S. pyogenes* [31]. For the transcriptome analysis, three pre-cultures of the wild type and the *pnp* mutant strain of *R. sphaeroides* were inoculated with cells from three different single colonies. With each of these pre-cultures, three main cultures were inoculated (nine in total), grown under microaerobic conditions and later harvested during the exponential growth phase. All cultures initially derived from one colony in the first step were pooled. Total RNA was isolated and the DNA-free RNA was sequenced on an Illumina NextSeq 500 platform. The overall reproducibility within the replicates was fair, only one replicate obtained from the wild type strain showed some deviation to the other samples of the group (Supplementary Fig. S1, Additional file 1). In total 98% of the entire variance can be explained by the first two principal components.

Figure 3 shows the result of the DESeq2 analysis (version 1.26.0 [40];) and illustrates the $\log_2$-fold changes of the normalized read numbers in the *pnp* mutant versus the wild type strain (see Supplementary Table S3, Additional file 2). All transcripts with a $\log_2$-fold change ≤ −1 or ≥ +1 and an adjusted *p*-value ≤0.05 (Benjamini

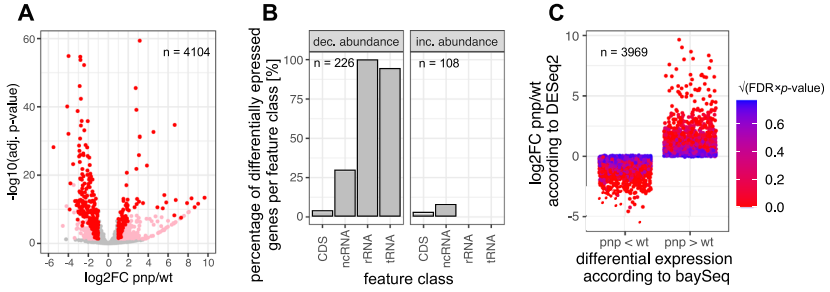


Fig. 3 The *Rhodobacter sphaeroides* transcriptome composition is strongly influenced in the PNPase mutant. **a** Volcano plot of the observed $\log_2$-fold changes based on RNA-Seq data analyzed with DESeq2. Genes with significant change in abundance are colored red (adjusted *p*-value ≤0.05, $\log_2$fold change ≤ −1 or ≥ +1, basemean ≥50) and pink (adjusted *p*-value ≤0.05, $\log_2$-fold change < −1 or > +1, basemean < 50). Grey dots: adjusted *p*-value > 0.05 or −1 ≤ $\log_2$-fold change ≥ +1. Altogether the transcripts of 334 genes were observed to differ in a statistically significant manner and exhibited a basemean above the threshold. **b** Feature-wise distribution of these significant genes, classified in decreased and increased abundance (*pnp* mutant/wild type). Most tRNAs and all rRNAs showed a reduced abundance in the mutant strain. x-axis: feature class; y-axis: percentage of differentially expressed genes per feature class [%] **c**) Comparison of data computed with DESeq2 and baySeq, which show a very good match. Almost all transcripts that are lower abundant in the *pnp* mutant according to DESeq2 ($\log_2$-fold change (*pnp*/wt) < 0) are also classified to be lower abundant by baySeq (*pnp* < wt) and vice versa. Since baySeq does not provide *p*-values, the color coding represents the square root of the product of the false discovery rate (FDR, obtained from baySeq) and the adjusted *p*-value (obtained from DESeq2). Every dot represents one gene

Hochberg algorithm) were considered to have a significant differential abundance within the two strains (coloured dots). We then decided to only keep those differentially expressed genes which have a basemean ≥50 (red dots) in order to further decrease the number of false positive hits. In total 334 transcripts met these strict criteria, 226 of them showed lower abundance in the *pnp* mutant strain and 108 showed higher abundance in the *pnp* mutant strain compared to the wild type. The most prominent differences were observed in the feature classes tRNA and rRNA: 94% of all tRNAs (51 out of 54) and 100% of all rRNAs (9 out of 9) showed a lower abundance in the *pnp* mutant strain (Fig. 3b). Altogether 37% of all non-coding RNAs, here merged of sRNAs and ncRNAs (including 6S, SRP RNA and tmRNA), were observed to have a differential abundance. Within the groups of RNAs with increased or decreased abundance, no distinct orthologous group of encoded proteins (COG) could be found to be prominent (Supplementary Fig. S2A + B, Additional file 1).

The transcriptome is directly affected by the action of RNases. Moreover, the RNA entity is modulated through secondary effects by the PNPase-mediated processing of sRNAs and mRNAs that code for regulatory elements, for example transcription factors. Thus, our transcriptome analysis reflects both direct and indirect PNPase dependent regulations and does not allow a distinction. In either case, our data emphasize the effect which PNPase has especially on stable RNAs (rRNA, tRNA). A similar effect was also observed in *E. coli*, although both rRNAs and tRNAs were more abundant in the *pnp* mutant despite a conducted rRNA depletion prior to RNA sequencing [47]. Further, Płociński et al. [48] demonstrated, that PNPase is involved in processing of ribosomal RNA and tmRNA in *Mycobacterium smegmatis* and *M. tuberculosis*.

We further validated these predictions using a different algorithm. An empirical Bayes approach integrated in the baySeq package (version 2.20.0 [41];) was used to identify differential expression (Supplementary Table S4, Additional file 2). The results of the two methods perfectly agree, since virtually all genes could be properly assigned. Every transcript (blue dot) with a $\log_2$-fold change ≤0 (*pnp* mutant/wild type) according to the DESeq2 analysis was also observed to be lower abundant in the *pnp* mutant according to the baySeq algorithm and vice versa (Fig. 3c). This includes every differently expressed gene which fulfills the strict criteria as mentioned above.

The RNA sequencing data was further used to investigate the cellular RNA 3′ elongation. All reads that could not be mapped end-to-end were instead mapped in very sensitive local mode with bowtie2 (version 2.2.6). To increase the quality of the analysis, all reads without a minimal matching sequence of 10 nt at the 5′ end were excluded. Only soft clipped sequences at the 3′ ends of the remaining reads were extracted with awk (version 4.1.3) (Supplementary Fig. S3A, Additional file 1). The overall results are similar for both strains: The lengths of elongated sequences are comparable, the majority of them (95%) is shorter than 39 nt in length. Further, the base frequency for each nucleotide position of the 3′ tail reveals an enrichment of guanine within the first 20 bases (Supplementary Fig. S3B + C + D, Additional file 1). A sequence motif which is related to a PNPase-dependent elongation could not be identified. Since both the lengths and base frequencies of the 3′ tails do not differ in between the analyzed strains, we conclude that the deletion of the KH-S1 domains does not have a major impact on the overall RNA 3′ elongation events in *R. sphaeroides*.

Levels of regulatory sRNAs are influenced by PNPase

An important effect of the PNPase on levels of small RNAs was reported: the enzyme does not only influence mRNA but also sRNA stability [49–51]. We were especially interested in those sRNAs that are derived from 5′ or 3′ UTRs and wanted to investigate the role of PNPase during the maturation process. For further analysis, we selected five sRNAs which showed a different pattern in the read coverage comparing *pnp* mutant and wild type. Two of them, CcsR1 and SorY, are known to have a regulatory function during the oxidative stress response in *Rhodobacter sphaeroides* [52, 53]. UpsM is processed from the *mraZ* 5′ UTR in a stress-dependent manner by RNase E [54]. The other two sRNAs have not been described so far and their function is still unknown. One is located in the intergenic region between RSP_1711 and *rpsL* and is derived from the *rpsL* 5′ UTR. The second one is derived from the 5′ UTR of RSP_6083. During the exponential growth phase, three of these sRNAs differed in abundance comparing the total RNA from the *pnp* mutant and the wild type strain (Fig. 4a+b). Moreover, processing products of the sRNAs IGR_1711_*rpsL* and 5′ UTR_6083 were prominently enriched in the *pnp* mutant. Interestingly, the abundance of the mature transcript of SorY and 5′ UTR RSP_6083 does not vary between the strains. To further evaluate the sRNA stability, we added rifampicin during the exponential phase and determined the RNA half-lives (Fig. 4c+d). CcsR1, SorY, IGR_1711_*rpsL* and 5′UTR_6083 are strongly stabilized in the mutant lacking PNPase, resulting in prolonged half-lives. In contrast to that, the half-life of UpsM drops form 12.2 min in the wild type to 4.0 min in the *pnp* mutant. The changed stabilities are in agreement with the observed sRNA levels during exponential phase (Fig. 4a). These observations highlight the role of PNPase during the maturation of sRNAs and in the homeostasis of their

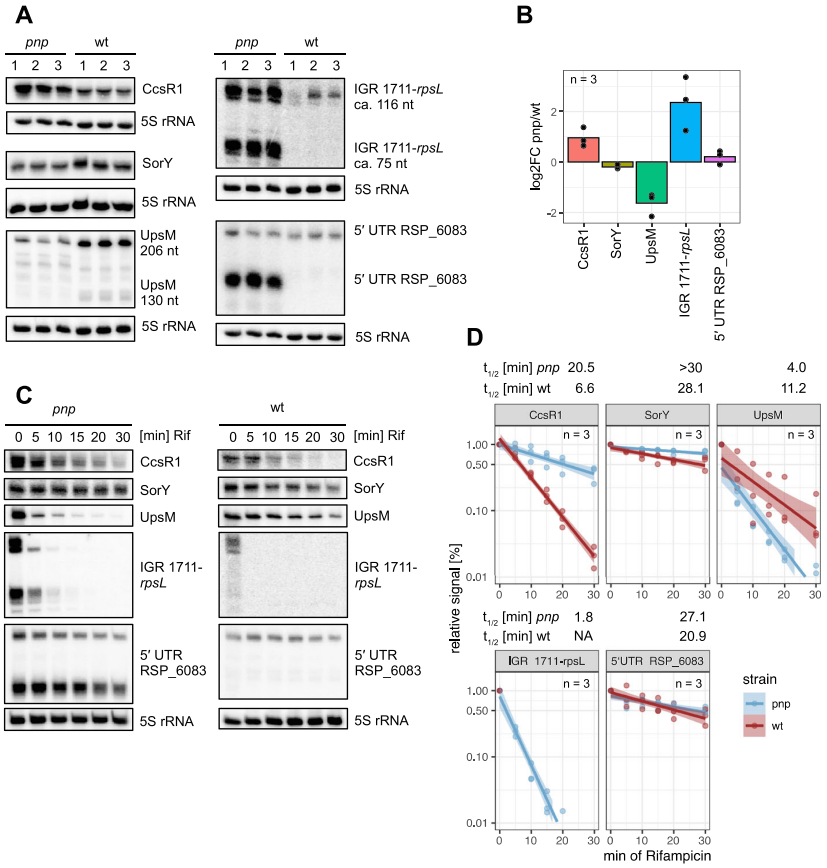


Fig. 4 a + b Northern blot analysis of the sRNAs CcsR1, SorY, UpsM, IGR_1711_*rpsL* and 5′ UTR RSP_6083. Total RNA was isolated during exponential growth phase (OD_{660} = 0.45) from microaerobic cultures of wild type or *pnp* mutant. Loading control: 5S rRNA. The barchart illustrates the log_2–fold changes within the indicated sRNAs. *n* = 3. **c** Wild type and *pnp* mutant cells were cultivated under microaerobic conditions until exponential phase. Samples were harvested before (t_0) and after (t_5-t_{30}) addition of 0.2 mg/ml rifampicin. Total RNA was isolated and used for Northern blot analysis. Loading control: 5S rRNA. **d** The signal intensities for the indicated probes were normalized to the 5S rRNA signal, timepoint t_0 was set to 1. x-axis: minutes after addition of 0.2 mg/ml rifampicin; y-axis: relative signal intensity. Solid lines represent mean value for biological triplicates, light colors indicate the standard deviation (*n* = 3). Blue: *pnp* mutant; red: wild type. The full-length Northern blots are presented in Supplementary Fig. S4, S5, S6 and S7, Additional file 1

levels which has been described in *E. coli*. Cameron and De Lay [50] reported a stabilizing function by PNPase on Hfq-dependent sRNAs during the exponential growth phase in *E. coli*. They speculate, that PNPase may for example protect Hfq-bound sRNAs by degrading binding sites for other ribonucleases. Our data show a different trend in *Rhodobacter sphaeroides*, which suggests a mainly destabilizing effect on the selected sRNAs in this study. Even though CcsR1–4, UpsM and SorY are Hfq-dependent [52, 54, 55], CcsR1–4 and SorY are destabilized by PNPase and only UpsM fits the model proposed for *E. coli*. The two UTR-derived sRNAs are also destabilized by PNPase. The reason for these differences remains unidentified. Given this major influence of PNPase on the abundance of regulatory sRNAs, the pleiotropic effect of a *pnp* mutant becomes even more perspicuous.

Deletion of the KH-S1 domains of PNPase leads mainly to enriched RNA 3′ ends

As a 3′-to-5′ exoribunuclease, PNPase plays an important role in RNA turnover and decay from the 3′ end. Therefore, we analyzed how the RNA 3′ ends differ in abundance comparing the *pnp* mutant to the wild type. For this study we developed XPEAP, an analysis pipeline. It allows the detection of RNA 5′ or 3′ ends in prokaryotic NGS data and covers all relevant steps form data preprocessing to the final statistical analysis. As input the raw read files, three replicates of each strain, were used. After trimming and read alignment to the

reference genome, READemption's subcommand *coverage* was used to generate coverage files which contain the nucleotide positions of the 3′ end bases of each aligned read. The coverage files from the plus and minus strand were integrated in one data set. Subsequently all nucleotide positions that did not exhibit a coverage of ≥10 in at least one of the analyzed libraries were excluded. We further calculated the ratio of the 3′ end coverage and the full read coverage for every replicate and every position. To improve the signal noise ratio, only positions with a ratio higher than 0.05 were kept. A DESeq2 analysis was conducted based on this nucleotide-wise coverage files to detect differences within the two strains. All positions which showed a $\log_2$-fold change ≤ − 1 or ≥ + 1 and an adjusted *p*-value ≤0.05 (Benjamini Hochberg algorithm) were kept for further analysis, the other positions were rejected. Due to the fact that the 3′-to-5′ end processing is a dynamic process, it is supposed that in some cases several 3′ ends per RNA molecule will be detected. This is why we decided to merge all nucleotide positions within a range of 3 nt with BEDtools' subcommand *merge*. The range of mapped positions which belong to one 3′ end is defined as distribution size. All resulting positions are regarded as true differential 3′ ends.

In total 1072 differential 3′ ends could be detected, the majority of them (around 68%) were mapped to a single position (Fig. 5a). By far most of all 3′ ends (82%) were strongly enriched in the *pnp* mutant strain showing a $\log_2$-fold change between + 5 and + 9 (Fig. 5b). These ends represent either termination sites or arise from cleavage by endoribonucleases and are further processed by PNPase in the wild type. The ends were strand-specifically assigned to the different feature classes according to their genomic position with BEDtools *intersect*. All 3′ ends that did not overlap with any feature were classified in the group UTR. This group may also contain additional sRNAs that have not yet been identified. The most prominent changes could be observed within the coding sequences and the untranslated regions: In both feature classes, around 10 times more RNA 3′ ends were enriched in the *pnp* mutant compared to the wild type (Fig. 5c). These 885 ends are interpreted as PNPase-degraded. To search for putative motifs, the 15 nt upstream sequence of every *pnp* enriched differential 3′ end was extracted with BEDtools *getfasta*. Only non-overlapping 3′ ends were kept for the following analysis to reduce bias during the motif analysis. No binding motif or consensus sequence could be found with MEME Suite (version 5.3.0).

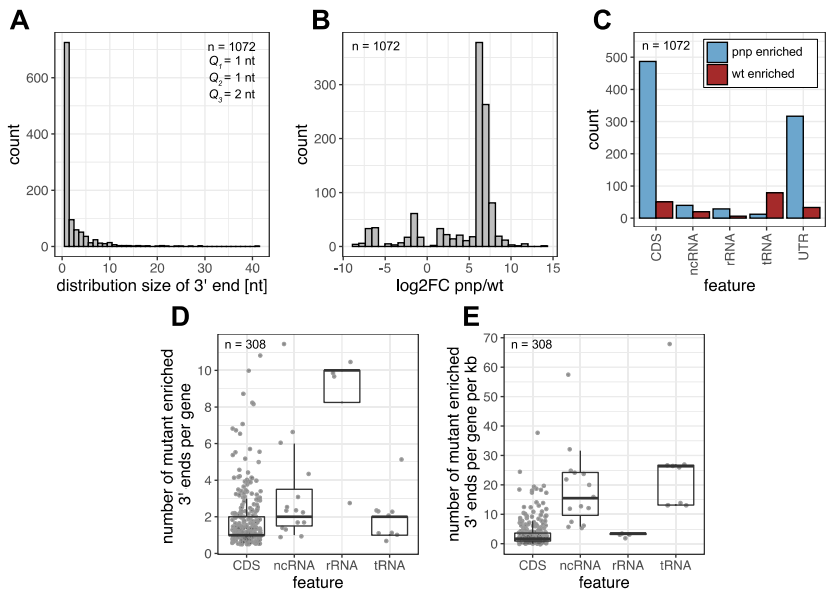


Fig. 5 The PNPase mutation leads to a change in RNA 3′ end composition. **a** Histogram illustrating the different distribution sizes of all 1072 differential 3′ ends. Q_1 = 1 nt; Q_2 = 1 nt; Q_3 = 3 nt; arithmetic mean = 3.86 nt. **b** Distribution of the $\log_2$-fold changes of all differential 3′ ends, comparing the *pnp* mutant strain and the wild type. 82% of the 3′ ends are enriched in the mutant. **c** All 3′ ends were assigned to genes or UTRs according to their genomic position and then grouped in feature classes. Most 3′ ends in CDS (coding sequences) and UTRs were enriched in the *pnp* mutant, but more tRNA 3′ ends were detected in the wild type. **d** + **e** The boxplots show the absolute number of 3′ ends enriched in the mutant per gene **d** and the relative number (normalized to the gene length) (**e**). Overall 306 genes were observed to harbour at least one 3′ end enriched in the mutant. Every dot represents one gene

We further detected 189 3′ ends which are enriched in the wild type. One possible explanation for this observation is, that these RNAs are degraded by PNPase from the 3′ end, but at the detected positions stable secondary structures prevent further RNA 3′-to-5′ end degradation in the wild type. For a test of this hypothesis all sequences within windows of 20 nt, 30 nt, 40 nt and 50 nt upstream of the detected wild type enriched 3′ ends were extracted as described above. As a control we selected sequences of the same length but located downstream of the 3′ ends (Supplementary Fig. S8A, Additional file 1). These RNA sequences are supposed to have no effect on the stop of PNPase decay since they are properly degraded. RNAfold (version 2.4.17) was used to compute the minimal folding energy (MFE) of every sequence. Independently of the window size, the distributions of both groups are highly similar and no shift to low MFE values was observable in the upstream sequences (Supplementary Fig. S8B, Additional file 1). But remarkably, the number of unstructured sequences (MFE = 0.0 kcal/mol) was three times higher downstream of the 3′ ends (n = 29, 13.8%) compared to the regions upstream of the 3′ ends (n = 9, 4.8%) (Supplementary Fig. S8C, Additional file 1). We therefore conclude that unstructured sequences may at least enhance degradation by PNPase while highly structured RNA sequences may facilitate a stop of decay. Moreover, other still unknown factors, e.g. binding proteins, are likely to influence degradation by PNPase. Using MEME, no recurring motif that could hint to conserved binding sites was detectable in the upstream sequences.

The group of tRNAs exhibited less differential 3′ ends in the *pnp* mutant strain, all with rather minor $\log_2$-fold changes (median – 1.3). This affects 32 of in total 54 tRNAs. We observed characteristic 3′ ends which show a clear edge in the wild type RNA coverage profile of the $tRNA^{Gly}$ and $tRNA^{Leu}$ (see Supplementary Fig. S9A, Additional file 1). The other tRNAs coverage profiles of the wild type harbour only minor edges but are stronger sloped than in the *pnp* mutant (see Supplementary Fig. S9B, Additional file 1). The reason for these differences still remains concealed. Nevertheless, the detection of those 3′ ends depicts the high sensitivity of XPEAP. The influence of the *E. coli* PNPase on tRNA maturation and degradation has been investigated intensively. PNPase is involved in the repair process of several tRNAs [56], although this enzyme does not affect the tRNA poly(A) tail length [57]. Furthermore, both PNPase and RNase II remove the Rho-independent terminator structure of the leuX tRNA [58]. In chloroplasts of *A. thaliana*, the PNPase activity is directly linked to the decay of tRNAs [59].

Moreover, our data suggest that PNPase-dependent degradation is not limited to only one site per gene. In many cases, more than one RNA 3′ end per gene was enriched in the mutant (Fig. 5d+e). This affects in particular the 23S rRNAs, but also several mRNAs and ncRNAs. This is not surprising, since often several RNase E cleavage sites per gene could be detected in *Rhodobacter sphaeroides*, but also in *Salmonella enterica* [6, 7]. After RNase E cleavage, the RNA fragments become potential new substrates for PNPase.

It should be mentioned that some of the 1072 differential 3′ ends may arise from a stop of the sequencing reaction after about 75 nucleotides in the single-end sequencing. 3′ ends at such a distance from the 5′ end account for only 4.8% of all differential 3′ ends. Moreover, this number also contains real 3′ ends as depicted in Fig. 4: as predicted, a 75 nt RNA derived from IGR 1711-*rpsL* occurs only in the mutant but not in the wild type.

Intersection analysis

Multiple ribonucleases are involved in RNA processing and turnover. In many organisms, the ribosomal RNA maturation requires an initial endonucleolytic cleavage of the long nascent precursor transcript by RNase III. This is followed by further enzymatic reactions, performed for example by RNase E, J, G and various other enzymes. PNPase can process those RNA species which are newly generated by endonucleases from the 3′ end. Furthermore, both PNPase and RNase E are part of the *E. coli* degradosome and work together in RNA degradation. In the *R. capsulatus* degradosom fraction PNPase activity could be detected, but only a small amount [11]. To get more insight into the interplay of RNase E, RNase III and PNPase we analyzed the correlation of 3′ ends from the wild type, the *pnp* mutant, an RNase III mutant and a strain with reduced RNase E activity. In the RNase III mutant strain, the *rnc* gene was removed by homologous recombination [24]. Since RNase E is an essential enzyme in *Rhodobacter sphaeroides*, the *rne* mutant strain was achieved by replacing the native *rne* by the gene of the thermosensitive RNase E from *E. coli* [54]. At 32 °C the enzymatic activity is already reduced and is even more reduced at 42 °C [7].

First, the differential 3′ ends were detected as described in the previous section (see Additional file 3 for the full list of detected RNA 3′ ends). For the comparison of a mutant to the corresponding wild type strain, we only analyzed data that was obtained from the very same sequencing chip. Next, the overlap between the detected 3′ ends was investigated with BEDtools' subcommand *window*. Using this function, a window of 1 nt

upstream and 1 nt downstream of every differential *pnp* mutant 3′ end was set prior to the intersection analysis to compensate a potential inaccuracy during 3′ end determination. Two hundred eighty-nine differential 3′ ends could be detected in the *rnc* mutant, more than half of them (166; 57%) were uniquely overlapping with PNPase dependent 3′ ends (Fig. 6a). Within this group, no 3′ end showed a reduced abundance in the *pnp* mutant and an increased abundance in the *rnc* mutant strain (Fig. 6b). By far the highest number of overlapping ends could be found in tRNAs (Fig. 6c). Taken together, 9.7% of all RNase III generated RNA 3′ ends are further processed by PNPase (Fig. 6a + b: quadrant *rnc* depleted, *pnp* enriched). Although in *E. coli* RNase III is involved in *pnp* mRNA processing [20, 60], we could not observe any effect on the *pnp* mRNA levels in the RNase III mutant strain in *Rhodobacter sphaeroides*.

A much greater number of total 3′ ends was identified to be RNase E-dependent: 5445 at 32 °C and 14,873 at 42 °C (Fig. 6d+g). At the permissive temperature, the majority of differential 3′ ends could be assigned to CDS and UTR (Fig. 6e+f). At the non-permissive temperature more than 2.5% of all differential *pnp* mutant RNA 3′ ends (279 in total, 271 unique *pnp* mutant ends) overlap with those

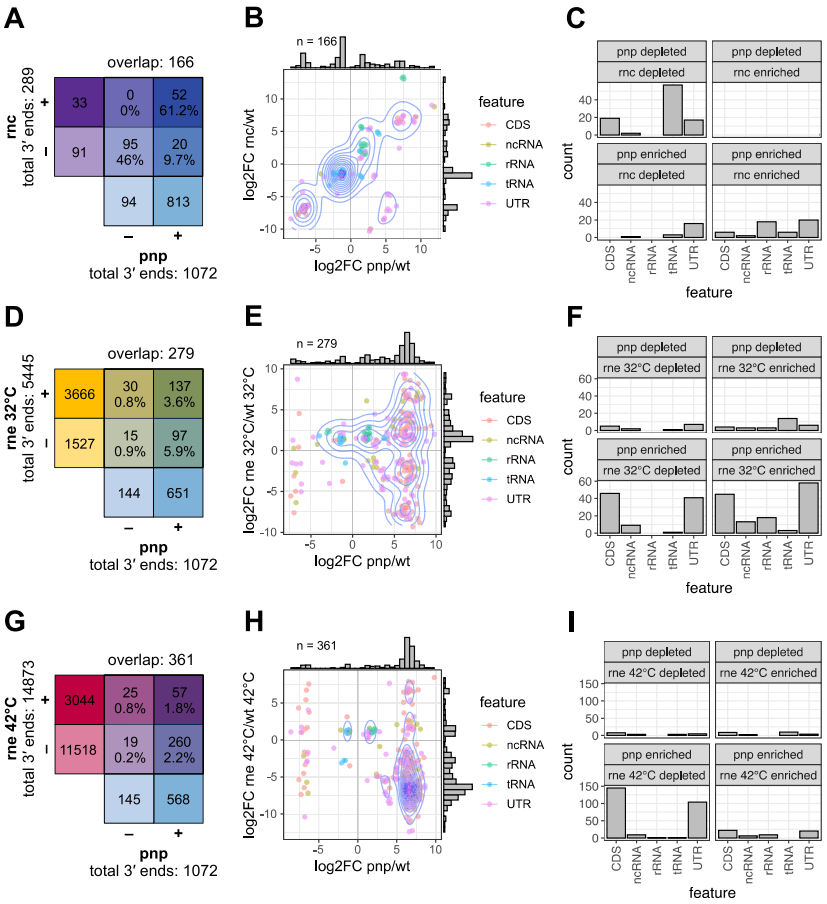


Fig. 6 Many differential 3′ ends found in the PNPase mutant can also be detected in the RNase III or RNase E mutant strain at the same genomic position. **a, d, g** Venn diagrams illustrating the total number of differential 3′ ends found in each mutant strain in comparison to the wild type (+: mutant enriched; −: mutant depleted) and the number differential 3′ ends which could be mapped to the same genomic position. Percentage: number of 3′ ends in respective quadrant devided by the total number of *rnc*/*rne* 32 °C/*rne* 42 °C enriched or depleted 3′ ends. Blue: *pnp* mutant; violett: *rnc* mutant; yellow: *rne* mutant at 32 °C; red: *rne* mutant at 42 °C. **b, e, h** Scatterplots depicting the intersecting analysis, every dot represents one differential 3′ end. x-axis: log_2-fold change *pnp* mutant versus wild type; y-axis: log_2-fold change *rnc*/*rne* mutant versus wild type. Feature classes harboring the observed 3′ ends are color coded. Orange: CDS; yellow: ncRNA; green: rRNA; blue: tRNA; violett: UTR. Blue lines represent a 2D density function. The histograms (grey) illustrate the distribution of intersecting 3′ end log_2-fold changes of the compared strains. **d, f, i** Histograms display the number of intersecting 3′ ends, which are grouped in mutant enriched/depleted and feature classes. x-axis: feature class; y-axis: count

differential 3′ ends identified in the *rne* mutant. Plotting the $\log_2$-fold changes reveals a specific pattern: 75% of all intersecting 3′ ends are *pnp* enriched and *rne* depleted (Fig. 6h), suggesting that these 3′ ends are generated by RNase E cleavage and further removed by PNPase. This group is mainly composed of ends located in CDS and UTRs, but also in several ncRNAs (Fig. 6i). Interestingly only 5.9% of all 3′ ends which are generated by RNase E are further degraded by PNPase. Such a low fraction of around 6% was also observed in *S. pyogenes* [31]. We hypothesize, that a portion of the RNase E generated 3′ ends can at least partly be trimmed by other 3′-to-5′ exonucleases and thus are not detected in the *pnp* mutant strain. Also, stable secondary structures at the newly generated 3′ ends may prevent PNPase degradation. At the permissive temperature, the *pnp* mRNA levels do not differ, whereas at 42 °C the *pnp* mRNA levels are more than doubled ($\log_2$-fold change = 1.75, comparison *rne* mutant versus wild type) strongly supporting an effect of RNase E on *pnp* stability.

We further elucidated the statistical significance of the observed overlapping 3′ ends. For every comparison of detected 3′ ends (*pnp* mutant versus *rnc*/*rne* mutant) BEDtools' subcommand *fisher* was used to first compute the number of possible 3′ ends, taking into account the genome size and the individual distribution sizes of determined 3′ ends. Second, the number of overlaps and non-overlaps was computed and the Fisher's exact test applied. For a possible number of 633,092 (*pnp* versus *rnc*), 748,506 (*pnp* versus *rne* 32 °C) and 667,153 (*pnp* versus *rne* 42 °C) RNA 3′ ends the resulting two-tail *p*-value are close or equal to 0 in each case (0, 7.3×10^{-241}, 2.9×10^{-239}). This strongly suggests a higher number of RNA 3′ ends at the same genomic positions than would be expected, if the given 3′ ends would be randomly distributed within the genome.

The total number of detected 3′ ends as well as the overlapping 3′ ends differ within the strains analyzed in this study. It cannot be excluded, that this observation may be influenced by a different number of uniquely aligned reads: The samples of the RNase E mutant and the corresponding wild type strain previously published [7] showed on average a four times higher number of uniquely aligned reads than the samples sequenced for this study (PNPase mutant, RNase III mutant and wild type strain). On the other hand, although Lécrivain et al. [61] used a somewhat different algorithm and parameters to detect differential 3′ ends in *Streptococcus pyogenes*, the number of identified 3′ ends (wild type enriched: 183; *pnp* mutant enriched: 1255) is strikingly similar to our findings in the Gram-negative organism *R. sphaeroides* (wild type enriched: 189; *pnp* mutant enriched: 885). This similarity strongly supports the reliability of the data obtained in this study and suggests valid differences in between the analyzed strains.

To illustrate these overlapping 3′ ends we chose two of the previously shown sRNAs. Several differential 3′ ends can be found in the sRNAs CcsR1–4, which are derived from one co-transcript by RNase E cleavage [52]. The read coverage depicts a strong enrichment of CcsR1–4 in the *pnp* mutant (Fig. 7a), which agrees with the Northern Blot data (Fig. 4a).

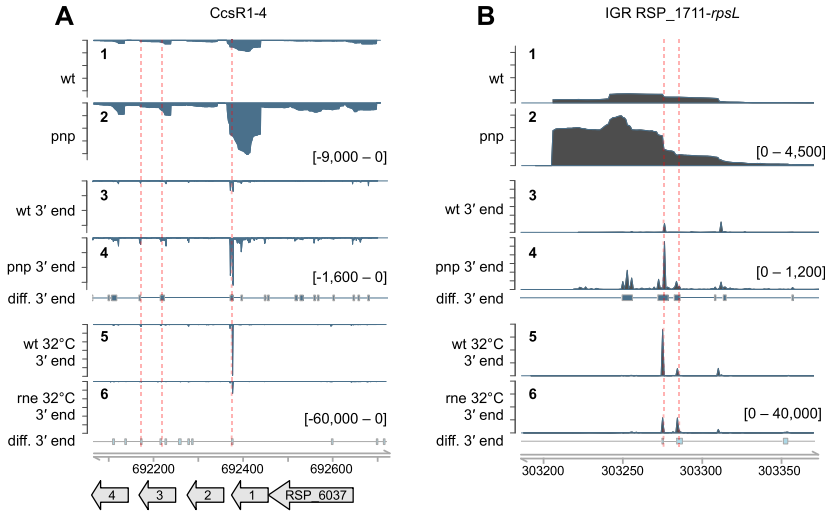


Fig. 7 The sRNA levels of CcsR1–4 and IGR_RSP_1711_*rpsL* are modulated by RNase E cleavage and PNPase processing. XPEAP reveals overlapping differential 3′ ends detected in the RNAs CcsR1–4 (**a**) and IGR_RSP_1711_*rpsL* (**b**). Panel 1 and 2: RNA-Seq full coverage (wild type and *pnp* mutant strain). Panel 3 and 4: 3′ end coverage (wild type and *pnp* mutant strain). Panel 5 and 6: 3′ end coverage (wild type and *rne* mutant strain at 32 °C). Small boxes indicate positions of true differential 3′ ends computed with XPEAP. Red lines: overlapping differential 3′ ends. Y-axis count scales are indicated in square brackets

Several *pnp* mutant enriched 3′ ends have been detected. Three of them overlap with those RNA 3′ ends which are depleted in the *rne* mutant in comparison to the wild type, even at 32 °C (Fig. 7a, highlighted in red color). The read coverage of the sRNA IGR_RSP_1711_*rpsL* reveals, that the first 75 bases of the RNA are enriched in the *pnp* mutant (Fig. 7b). This shorter fragment was also detected by Northern blot analysis (Fig. 4a). Comparing wild type and the RNase E deficient strain, a 3′ end depleted in the mutant could be found at the very same site. Despite some changes in overall abundance, no RNase III-dependent differential 3′ ends were detectable. For both of the two described sRNAs, a similar processing pattern is proposed. The initial transcript is first processed by RNase E, indicated by the depleted 3′ ends in the *rne* mutant. Followed by this cleavage, the remaining RNA molecule is further degraded by PNPase thus leading to enriched 3′ ends in the *pnp* mutant. This observation is in agreement with previous studies in Gram-negative (reviewed in [29, 30]) as well as in Gram-positive bacteria. In *Streptococcus pyogenes* for example, the endonuclease RNase Y generates new RNA 3′ ends, which are subsequently trimmed by PNPase [31]. PNPase seems not to be a component of the degradosome of the related species *R. capsulatus* [11], which most likely also accounts for *R. sphaeroides*. Even though this direct interaction between RNase E and PNPase may be lacking, our data provide strong evidence that stepwise RNA processing is mediated by these enzymes also in *Rhodobacter*.

Conclusion

In this study we characterized the function of PNPase in the Gram-negative alpha-proteobacterium *Rhodobacter sphaeroides* and shed light on the interplay of RNase E, RNase III and PNPase during RNA degradation. A lack of the KH-S1 RNA binding domain leads to severe effects on growth behavior, temperature stress tolerance, pigmentation and the bacteriochlorophyll *a* content. Total RNA sequencing illustrates a high impact of PNPase on levels of diverse RNAs, in particular on tRNAs and rRNAs. We could further demonstrate, that the stability of several regulatory sRNAs relies on PNPase function, thus leading to altered levels during exponential growth phase. Next, the global comparison of differential RNA 3′ ends identified in the *pnp* and *rnc* mutant strain as well as in a strain with reduced RNase E activity is in agreement with a sequential processing of transcripts: 5.9% of all RNase E and 9.7% of all RNase III generated RNA 3′ ends are afterwards trimmed by PNPase.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-021-07409-4.

Additional file 1: Table S1. Strains used in this study. **Table S2.** Overview of all oligonucleotide probes and corresponding sequences that were used in this study. **Figure S1.** Scatterplot of the principal component analysis, which was performed using DESeq2. Each of the colored dots (*pnp* mutant strain: red; wild type: blue) represents one replicate. **Figure S2.** Comparing the *pnp* mutant to the wildtype, the RNAs with decreased (A) or increased abundance (B) could not be assigned to specific orthologous groups of encoded proteins (COG). x-axis: relative number of observations per group [%]; y-axis: COG category. **Figure S3.** The 3′ elongated RNA sequences are similar in the wild type and the *pnp* mutant strain. A) Schematic overview of the sequence extraction procedure. The lengths of all non end-to-end mapped reads (B) and of all tail sequences (C) do not differ. In both strains the first part of the tail (around 20 nt in length) is guanine rich (D). **Figure S4.** Full size Northern blots of the depicted images in Fig. 4. with probes CcsR1 (A), SorY (B), UpsM (C), IGR_1711_*rpsL* (D) and 5′ UTR RSP_6083 (E). Red frames mark the selected sections. **Figure S5.** Full size Northern blots of the depicted half-life images in Fig. 4. with probes CcsR1 (A) and SorY (B). Red frames mark the selected sections. **Figure S6.** Full size Northern blots of the depicted half-life images in Fig. 4, with probes UpsM (A) and IGR_1711_*rpsL* (B). Red frames mark the selected sections. **Figure S7.** Full size Northern blots of the depicted half-life images in Fig. 4 with probe 5′ UTR RSP_6083. Red frames mark the selected sections. **Figure S8.** A) Sequence windows of 20, 30, 40 or 50 nt upstream and downstream of every wild type enriched 3′ end were extracted. B) Every dot in the boxplots depicts the minimal folding energy (MFE) [kcal/mol] of one RNA sequence computed with RNAfold. C) The number of unstructured sequences (MFE = 0.0 kcal/mol) is increased in the group of 20 nt downstream control sequences. **Figure S9.** Representative RNA full coverage and 3′ end coverage profiles of two selected tRNAs in the wild type and the *pnp* mutant strain. The full coverage of tRNAs can exhibit characteristic 3′ ends as shown for $tRNA^{Leu}$ (A) or 3′ ends which are detected because of minor edges and a different slope ($tRNA^{Pro}$, B).

Additional file 2: Supplementary tables of differential gene expression. **Table S3.** DESeq2 analysis results comparing the differential gene expression in the *pnp* mutant strain and the wild type. **Table S4.** BaySeq analysis results comparing the differential gene expression in the *pnp* mutant strain and the wild type.

Additional file 3: Supplementary tables of all differential RNA 3′ ends. **Table S4.** Differential RNA 3′ ends comparing the *pnp* mutant strain and the wild type. **Table S5.** Differential RNA 3′ ends comparing the *rnc* mutant strain and the wild type. **Table S6.** Differential RNA 3′ ends comparing the *rne* mutant strain (32 °C) and the wild type (32 °C). **Table S7.** Differential RNA 3′ ends comparing the *rne* mutant strain (42 °C) and the wild type (42 °C).

Acknowledgements

We thank Lennart Weber for developing the cloning strategy and Kerstin Haberzettl for the cloning. We further thank the Core Unit SysMed at the University of Würzburg for RNA-Seq data generation. This work was supported by Deutsche Forschungsgemeinschaft (GRK 2355) and the IZKF at the University of Würzburg (project Z-6).

Authors' contributions

DTS performed the RNA experiments, developed and programmed XPEAP, contributed to the data analysis and interpretation, and wrote the manuscript. CMR performed the physiological experiments and prepared total RNA for RNA sequencing. GK raised the funds, designed and supervised this study, was involved in data interpretation and manuscript writing. All authors read and approved the final manuscript.

Funding

The work was funded by Deutsche Forschungsgemeinschaft (DFG, GRK 2355). Open Access funding enabled and organized by Projekt DEAL.

Availability of data and materials
The datasets generated and analyzed during the current study are available in the NCBI Gene Expression Omnibus repository (GSE156818 and GSE71844).

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Received: 5 October 2020 Accepted: 26 January 2021
Published online: 06 February 2021

References

1. Feklístov AD, Sharon BA, Darst SA, Gross C. Bacterial sigma factors: a historical, structural, and genomic perspective. Annu Rev Microbiol. 2014; 68(1):357–76.
2. Kazmierczak MJ, Wiedmann M, Boor KJ. Alternative sigma factors and their roles in bacterial virulence. Microbiol Mol Biol Rev. 2005;69(4):527–43.
3. Evguenieva-Hackenberg E, Klug G. New aspects of RNA processing in prokaryotes. Curr Opin Microbiol. 2011;14(5):587–92.
4. Nitzan M, Rehani R, Margalit H. Integration of bacterial small RNAs in regulatory networks. Annu Rev Biophys. 2017;46(1):131–48.
5. Carpousis AJ, Luisi BF, Mcdowall KJ. Chapter 3 Endonucleolytic initiation of mRNA decay in *Escherichia coli*. Prog Mol Biol Transl Sci Mol Biol RNA Process Decay Prokaryotes. 2009:91–135.
6. Chao Y, Li L, Girodat D, Förstner KU, Said N, Corcoran C, et al. In vivo cleavage map illuminates the central role of RNase E in coding and non-coding RNA pathways. Mol Cell. 2017;65(1):39–51.
7. Förstner KU, Reuscher CM, Haberzettl K, Weber L, Klug G. RNase E cleavage shapes the transcriptome of *Rhodobacter sphaeroides* and strongly impacts phototrophic growth. Life Sci Alliance. 2018;1(4):e201800080.
8. Coburn GA, Mackie GA. Reconstitution of the degradation of the mRNA for ribosomal protein S20 with purified enzymes. J Mol Biol. 1998;279(5):1061–74.
9. Mackie GA, Genereaux JL. The role of RNA structure in determining RNase E-dependent cleavage sites in the mRNA for ribosomal protein S20 in vitro. J Mol Biol. 1993;234(4):998–1012.
10. Carpousis AJ. The RNA degradosome of *Escherichia coli*: An mRNA-degrading machine assembled on RNase E. Annu Rev Microbiol. 2007;61(1): 71–87.
11. Jäger S, Fuhrmann O, Heck C, Hebermehl M, Schiltz E, Rauhut R, et al. An mRNA degrading complex in *Rhodobacter capsulatus*. Nucleic Acids Res. 2001;29(22):4581–8.
12. Jäger S, Hebermehl M, Schiltz E, Klug G. Composition and activity of the *Rhodobacter capsulatus* degradosome vary under different oxygen concentrations. J Mol Microbiol Biotechnol. 2004;7(3):148–54.
13. Bermúdez-Cruz R, Ramirez F, Kameyama-Kawabe L, Montañez C. Conserved domains in polynucleotide phosphorylase among eubacteria. Biochimie. 2005;87(8):737–45.
14. Shi ZS, Yang W-ZS, Lin-Chao SS, Chak K-FS, Yuan HS. Crystal structure of *Escherichia coli* PNPase: central channel residues are involved in processive RNA degradation. RNA. 2008;14(11):2361–71.
15. Donovan WP, Kushner SR. Polynucleotide phosphorylase and ribonuclease II are required for cell viability and mRNA turnover in *Escherichia coli* K-12. Proc Natl Acad Sci. 1986;83(1):120–4.
16. Chen R, Weng Y, Zhu F, Jin Y, Liu C, Pan X, et al. Polynucleotide phosphorylase regulates multiple virulence factors and the stabilities of small RNAs RsmY/Z in *Pseudomonas aeruginosa*. Front Microbiol. 2016;7:247.
17. Li ZP, Deutscher MP. The role of individual exoribonucleases in processing at the 3′ end of *Escherichia coli* tRNA precursors. J Biol Chem. 1994;269(8): 6064–71.
18. Mohanty BK, Kushner SR. Polynucleotide phosphorylase functions both as a 3′ → 5′ exonuclease and a poly(a) polymerase in *Escherichia coli*. Proc Natl Acad Sci. 2000;97(22):11966–71.
19. Takata R, Izuhara M, Planta RJ. Differential degradation of the *Escherichia coli* polynucleotide phosphorylase mRNA. Nucleic Acids Res. 1989;17(18): 7441–51.
20. Jarrige A-C, Mathy N, Portier C. PNPase autocontrols its expression by degrading a double-stranded structure in the *pnp* mRNA leader. EMBO J. 2001;20(23):6845–55.
21. Cheng ZF, Deutscher MP. Purification and characterization of the *Escherichia coli* exoribonuclease RNase R. comparison with RNase II. J Biol Chem. 2002; 277(24):21624–9.
22. Cudny H, Deutscher MP. Apparent involvement of ribonuclease D in the 3′ processing of tRNA precursors. Proc Natl Acad Sci. 1980;77(2):837–41.
23. Deutscher MP, Marshall GT, Cudny H. RNase PH: an *Escherichia coli* phosphate-dependent nuclease distinct from polynucleotide phosphorylase. Proc Natl Acad Sci. 1988;85(13):4710–4.
24. Rische-Grahl T, Weber L, Remes B, Förstner KU, Klug G. RNase J is required for processing of a small number of RNAs in *Rhodobacter sphaeroides*. RNA Biol. 2014;11(7):855–64.
25. Rische T, Klug G. The ordered processing of intervening sequences in 23S rRNA of *Rhodobacter sphaeroides* requires RNase J. RNA Biol. 2012;9(3):343–50.
26. Mathy N, Bénard L, Pellegrini O, Daou R, Wen T, Condon C. 5′-to-3′ exoribonuclease activity in bacteria: role of RNase J1 in rRNA maturation and 5′ stability of mRNA. Cell. 2007;129(4):681–92.
27. Rauhut R, Jäger A, Conrad C, Klug G. Identification and analysis of the *rnc* gene for RNase III in *Rhodobacter capsulatus*. Nucleic Acids Res. 1996;24(7): 1246–51.
28. Li Z, Pandit S, Deutscher MP. RNase G (CafA protein) and RNase E are both required for the 5′ maturation of 16S ribosomal RNA. EMBO J. 1999;18(10): 2878–85.
29. Deutscher MP. Degradation of RNA in bacteria: comparison of mRNA and stable RNA. Nucleic Acids Res. 2006;34(2):659–66.
30. Bechhofer DH, Deutscher MP. Bacterial ribonucleases and their roles in RNA metabolism. Crit Rev Biochem Mol Biol. 2019;54(3):242–300.
31. Broglia L, Lécrivain A-L, Renault TT, Hahnke K, Ahmed-Begrich R, Rhun AL, et al. An RNA-seq based comparative approach reveals the transcriptome-wide interplay between 3′-to-5′ exoRNases and RNase Y. Nat Commun. 2020;11(1):1587.
32. van Niel CB. The culture, general physiology, morphology, and classification of the non-sulfur purple and brown bacteria. Bacteriol Rev 1944;8(1):1–118. PubMed PMID: 16350090; PubMed Central PMCID: PMC440875.
33. Remes B, Berghoff BA, Förstner KU, Klug G. Role of oxygen and the OxyR protein in the response to iron limitation in *Rhodobacter sphaeroides*. BMC Genomics. 2014;15(1):794.
34. Glaeser J, Klug G. Photo-oxidative stress in *Rhodobacter sphaeroides*: protective role of carotenoids and expression of selected genes. Microbiology. 2005;151(6):1927–38.
35. Shiozawa JA, Welte W, Hodapp N, Drews G. Studies on the size and composition of the isolated light-harvesting B800-850 pigment-protein complex of *Rhodopseudomonas capsulata*. Arch Biochem Biophys. 1982; 213(2):473–85.
36. Janzon L, Löfdahl S, Arvidson S. Evidence for a coordinate transcriptional control of alpha-toxin and protein a synthesis in *Staphylococcus aureus*. FEMS Microbiol Lett. 1986;33(2–3):193–8.
37. Berghoff BA, Glaeser J, Sharma CM, Vogel J, Klug G. Photooxidative stress-induced and abundant small RNAs in *Rhodobacter sphaeroides*. Mol Microbiol. 2009;74(6):1497–512.
38. Förstner KU, Vogel JM, Sharma C. READemption - a tool for the computational analysis of deep-sequencing-based transcriptome data. Bioinformatics. 2014;30(23):3421–3.
39. Hoffmann S, Otto C, Doose G, Tanzer A, Langenberger D, Christ S, et al. A multi-split mapping algorithm for circular RNA, splicing, trans-splicing and fusion detection. Genome Biol. 2014;15(2):R34.
40. Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biol. 2010;11(10):R106.
41. Hardcastle TJ, Kelly KA. baySeq: Empirical Bayesian methods for identifying differential expression in sequence count data. BMC Bioinformatics. 2010; 11(1):422.
42. Quinlan AR. BEDTools: The Swiss-Army tool for genome feature analysis. Curr Protoc Bioinformatics. 2014;47(1):11.
43. Yamanaka K, Inouye M. Selective mRNA degradation by polynucleotide phosphorylase in cold shock adaptation in *Escherichia coli*. J Bacteriol. 2001; 183(9):2808–16.
44. Polissi A, Laurentis WD, Zangrossi S, Briani F, Longhi V, Pesole G, et al. Changes in *Escherichia coli* transcriptome during acclimatization at low temperature. Res Microbiol. 2003;154(8):573–80.

45. Goverde RL, Huis Int Veld JHJ, Kusters JG, Mooi FR. The psychrotrophic bacterium *Yersinia enterocolitica* requires expression of *pnp*, the gene for polynucleotide phosphorylase, for growth at low temperature (5°C). Mol Microbiol. 1998;28(3):555–69.
46. Clarke DJ, Dowds BC. The gene coding for polynucleotide phosphorylase in *Photorhabdus sp.* strain K122 is induced at low temperatures. J Bacteriol. 1994;176(12):3775–84.
47. Pobre V, Arraiano CM. Next generation sequencing analysis reveals that the ribonucleases RNase II, RNase R and PNPase affect bacterial motility and biofilm formation in *E. coli*. BMC Genomics. 2015;16(1):72.
48. Płociński P, Macios M, Houghton J, Niemiec E, Płocińska R, Brzostek A, et al. Proteomic and transcriptomic experiments reveal an essential role of RNA degradosome complexes in shaping the transcriptome of *Mycobacterium tuberculosis*. Nucleic Acids Res. 2019;47(11):5892–905.
49. Andrade JM, Pobre VM, Matos AM, Arraiano C. The crucial role of PNPase in the degradation of small RNAs that are not associated with Hfq. RNA. 2012; 18(4):844–55.
50. Cameron TA, Lay NRD. The phosphorolytic exoribonucleases polynucleotide phosphorylase and RNase PH stabilize sRNAs and facilitate regulation of their mRNA targets. J Bacteriol. 2016;198(24):3309–17.
51. Cameron TA, Matz LM, Sinha D, De Lay NR. Polynucleotide phosphorylase promotes the stability and function of Hfq-binding sRNAs by degrading target mRNA-derived fragments. Nucleic Acids Res. 2019;47(16):8821–37.
52. Billenkamp FA, Peng TA, Berghoff BA, Klug G. A cluster of four homologous small RNAs modulates C1 metabolism and the pyruvate dehydrogenase complex in *Rhodobacter sphaeroides* under various stress conditions. J Bacteriol. 2015;197(10):1839–52.
53. Adnan F, Weber L, Klug G. The sRNA SorY confers resistance during photooxidative stress by affecting a metabolite transporter in *Rhodobacter sphaeroides*. RNA Biol. 2015;12(5):569–77.
54. Weber L, Thoelken C, Volk M, Remes B, Lechner M, Klug G. The conserved *dcw* gene cluster of *R. sphaeroides* is preceded by an uncommonly extended 5′ leader featuring the sRNA UpsM. PLoS One. 2016;11(11): e0165694.
55. Berghoff BA, Glaeser J, Sharma CM, Zobawa M, Lottspeich F, Vogel J, et al. Contribution of Hfq to photooxidative stress resistance and global regulation in *Rhodobacter sphaeroides*. Mol Microbiol. 2011;80(6):1479–95.
56. Reuven NB, Zhou Z, Deutscher MP. Functional overlap of tRNA nucleotidyltransferase, poly(a) polymerase I, and polynucleotide Phosphorylase. J Biol Chem. 1997;272(52):33255–9.
57. Mohanty BK, Maples VF, Kushner SR. Polyadenylation helps regulate functional tRNA levels in *Escherichia coli*. Nucleic Acids Res. 2012;40(10): 4589–603.
58. Mohanty BK, Kushner SR. Processing of the *Escherichia coli* leuX tRNA transcript, encoding tRNA Leu5, requires either the 3′→5′ exoribonuclease polynucleotide phosphorylase or RNase P to remove the rho-independent transcription terminator. Nucleic Acids Res. 2010;38(2):597–607.
59. Walter M, Kilian J, Kudla J. PNPase activity determines the efficiency of mRNA 3-end processing, the degradation of tRNA and the extent of polyadenylation in chloroplasts. EMBO J. 2002;21(24):6905–14.
60. Portier C, Dondon L, Grunberg-Manago M, Régnier P. The first step in the functional inactivation of the *Escherichia coli* polynucleotide phosphorylase messenger is a ribonuclease III processing at the 5′ end. EMBO J. 1987;6(7): 2165–70.
61. Lécrivain A-L, Rhun AL, Renault TT, Ahmed-Begrich R, Hahnke K, Charpentier E. In vivo 3′-to-5′ exoribonuclease targetomes of *Streptococcus pyogenes*. Proc Natl Acad Sci. 2018;115(46):11814–9.

Publisher's Note

Chapter 3

Maturation of UTR-derived sRNAs is modulated during adaptation to different growth conditions

International Journal of
Molecular Sciences

MDPI

Article

Maturation of UTR-Derived sRNAs Is Modulated during Adaptation to Different Growth Conditions

Daniel-Timon Spanka and Gabriele Klug *

Institute of Microbiology and Molecular Biology, Justus Liebig University Giessen, Heinrich-Buff-Ring 26-32, 35392 Giessen, Germany; daniel-timon.spanka@mikro.bio.uni-giessen.de
* Correspondence: Gabriele.Klug@mikro.bio.uni-giessen.de

Abstract: Small regulatory RNAs play a major role in bacterial gene regulation by binding their target mRNAs, which mostly influences the stability or translation of the target. Expression levels of sRNAs are often regulated by their own promoters, but recent reports have highlighted the presence and importance of sRNAs that are derived from mRNA 3′ untranslated regions (UTRs). In this study, we investigated the maturation of 5′ and 3′ UTR-derived sRNAs on a global scale in the facultative phototrophic alphaproteobacterium *Rhodobacter sphaeroides*. Including some already known UTR-derived sRNAs like UpsM or CcsR1-4, 14 sRNAs are predicted to be located in 5′ UTRs and 16 in 3′ UTRs. The involvement of different ribonucleases during maturation was predicted by a differential RNA 5′/3′ end analysis based on RNA next generation sequencing (NGS) data from the respective deletion strains. The results were validated in vivo and underline the importance of polynucleotide phosphorylase (PNPase) and ribonuclease E (RNase E) during processing and maturation. The abundances of some UTR-derived sRNAs changed when cultures were exposed to external stress conditions, such as oxidative stress and also during different growth phases. Promoter fusions revealed that this effect cannot be solely attributed to an altered transcription rate. Moreover, the RNase E dependent cleavage of several UTR-derived sRNAs varied significantly during the early stationary phase and under iron depletion conditions. We conclude that an alteration of ribonucleolytic processing influences the levels of UTR-derived sRNAs, and may thus indirectly affect their mRNA targets.

Keywords: UTR-derived sRNA; sRNA processing and maturation; ribonucleases; RNase E; *Rhodobacter sphaeroides*; Alphaproteobacteria

Citation: Spanka, D.-T.; Klug, G. Maturation of UTR-Derived sRNAs Is Modulated during Adaptation to Different Growth Conditions. *Int. J. Mol. Sci.* **2021**, *22*, 12260. https://doi.org/10.3390/ijms222212260

Academic Editor: Kunio Takeyasu

Received: 6 October 2021
Accepted: 8 November 2021
Published: 12 November 2021

1. Introduction

Most bacteria live in environments that are subjected to changes in available nutrients, oxygen concentration, light conditions, temperature, and other parameters. In order to survive these changes and defend harmful stress conditions, bacteria have to adjust several layers of gene regulation and thus their physiology. This mostly goes along with massive changes of the transcriptome (e.g.,: [1–3]), but it is now well recognized that regulation at the post-transcriptional level also plays a crucial role in bacterial adaptation (e.g.,: [4,5]). Small RNAs affect regulation mostly at the post-transcriptional level, and make important contributions to the adaptation to stress conditions in bacteria ([6,7]; as reviewed in [8]). The function of several sRNAs in stress adaptation was investigated in the facultative phototroph *Rhodobacter sphaeroides*. This bacterium can perform aerobic or anaerobic respiration, fermentation or anoxygenic photosynthesis. Since the formation of photosynthetic complexes in the presence of high oxygen levels generates singlet oxygen, the formation of photosynthetic complexes and of defense systems is regulated by oxygen concentration, light intensity and the concentration of reactive oxygen species (e.g., [9–13]; reviewed in [14]). Some of the sRNAs involved in this regulation are derived from the 5′ or 3′ UTRs of precursor transcripts by processing. SorX is derived from the 3′ UTR of the *ompR-1* mRNA (*RSP_0847*) by RNase E-mediated cleavage [15,16]. Its level strongly increases in

response to several stresses, and by targeting the mRNA for a subunit of a spermidine transporter it counteracts oxidative stress [15]. The oxidative stress-induced CcsR1-4 RNAs are generated by processing of the 3′ UTR of the *ccaF1* mRNA (*RSP_6037*) and modulate the C1 metabolism and the pyruvate dehydrogenase complex [17]. RNase E as well as the small CcaF1 protein have important roles in maturation of the CcsR RNAs [18]. PcrX is derived by RNase E mediated cleavage from the 3′ UTR of the polycistronic *pufBALMX* mRNA. The *pufBALMX* mRNA encodes proteins of the photosynthetic complexes and PcrX targets *pufX* and modulates the expression of *puf* genes [19]. UpsM is derived from the 5′ UTR of the *dcw* (division and cell wall) gene cluster by RNase E mediated cleavage [20] and strongly influences growth of *R. sphaeroides* by influencing the *dcw* mRNA levels in *cis* and in *trans* [21]. The expression of all these sRNAs depends on the activity of a promoter, which in case of the 3′ UTR-derived RNAs is located in front of the upstream gene. Since in addition the levels of these sRNAs depend on maturation steps, we wondered whether these maturation steps may also be influenced by environmental factors and thus contribute to regulation. To address this, we first identified further 5′ or 3′ UTR-derived sRNAs and elucidated their generation by processing. Analysis of the sRNA levels in various mutants allowed us to decipher the involvement of the RNA-binding protein Hfq and of various RNases.

Several different RNases can be found in *R. sphaeroides*, among them the highly conserved and essential endoribonuclease E (RNase E). It has a major influence on large portions of the transcriptome, since more than 15,000 cleavage sites could be identified in *R. sphaeroides* [16]. RNase E cleaves mainly at AU-rich regions in *E. coli* [22], but also in *R. sphaeroides* [16] and binds to monophosphorylated 5′ ends [23]. Furthermore, RNase E is involved in processing of sRNAs from 3′ UTRs [19,24–26]. RNase III is another important endonuclease that plays a role during rRNA processing [27] and mRNA turnover in *E. coli* [28,29] and *Rhodobacter* [30–32]. Another important enzyme in the RNA life cycle is PNPase, which acts as a 3′-to-5′ exonuclease and degrades mRNAs [33] and sRNAs [34,35]. Recent studies showed that PNPase often attacks RNA 3′ ends that were generated by endonucleases such as RNase Y in the Gram-positive *Streptococcus pyogenes* [36] or by RNase E and RNase III in the Gram-negative organism *R. sphaeroides* [37]. Besides RNase E and RNase III, RNase J1 also takes part in rRNA processing by cleaving intervening sequences of the 23S rRNA in *R. sphaeroides* [38]. Only a very few other transcripts could be identified that are processed by RNase J [39]. Furthermore, we analyzed the involvement of RppH (an RNA pyrophosphohydrolase in *E. coli*; [40]) and YbeY (involved in maturation of 16S rRNA in *E. coli*; [41]) in regard to their role in UTR-derived sRNA maturation in *R. sphaeroides*. Alongside the mentioned RNases, the RNA chaperone Hfq is necessary in many cases to ensure a functional post-transcriptional gene regulation. Hfq can assist the hybrid formation between sRNAs and their target RNAs, thus influencing the RNA stability or translation [42,43]. This mechanism is also highly relevant under stress conditions [44,45].

Our results confirmed that the levels of UTR-derived sRNAs are not only determined by the levels of transcription, but that maturation is also influenced by environmental conditions and therefore needs to be considered as another important step of regulation.

2. Results

2.1. Identification of Five Novel UTR-Derived sRNAs

Several sRNAs were shown to play a role in stress responses in *Rhodobacter sphaeroides*, including few UTR-derived sRNAs [15,17,19,20]. In this study we depict a general picture of UTR-derived sRNAs in *Rhodobacter sphaeroides*. Based on RNA-Seq datasets, we especially searched for sRNAs that are located in the 5′- or 3′-UTR of mRNA transcripts. We were able to predict the presence of five novel UTR-derived sRNAs, which could all be validated via northern blot analysis. The lengths were predicted using RNA-Seq data and confirmed on northern blots using known RNAs as size markers (Figure 1A). In addition to the five novel sRNAs, we also included three sRNAs in our analysis, which have been described previously: IGR_RSP_1711_rpsL ([37]; here renamed UdsA), the sRNA downstream of

Int. J. Mol. Sci. 2021, 22, 12260

RSP_7527 ([16]; here named UdsC), and RSs2778 ([46]; here renamed to UdsE). These eight sRNAs were named UdsA to UdsH in regard to their maturation (UTR-derived sRNA). The described Uds' can be found on both chromosomes (chromosome 1: 5, chromosome 2: 3), but not on any plasmid (Figure 1B). Comparing the genomic location with the predicted 5′- and 3′-UTRs [16] reveals that four of the UTR-derived sRNAs are located in 5′-UTRs (UdsA, UdsE, UdsF and UdsG) and four in 3′-UTRs (UdsB, UdsC, UdsD and UdsH). An RT-PCR approach with specific primers for each sRNA and its corresponding mRNA was used to verify that these sRNAs are truly UTR-derived and stem from sRNA-mRNA cotranscripts (Figure S1). According to the Rfam database (version 14.6), none of the 5′ UTR-derived sRNAs is a riboswitch.

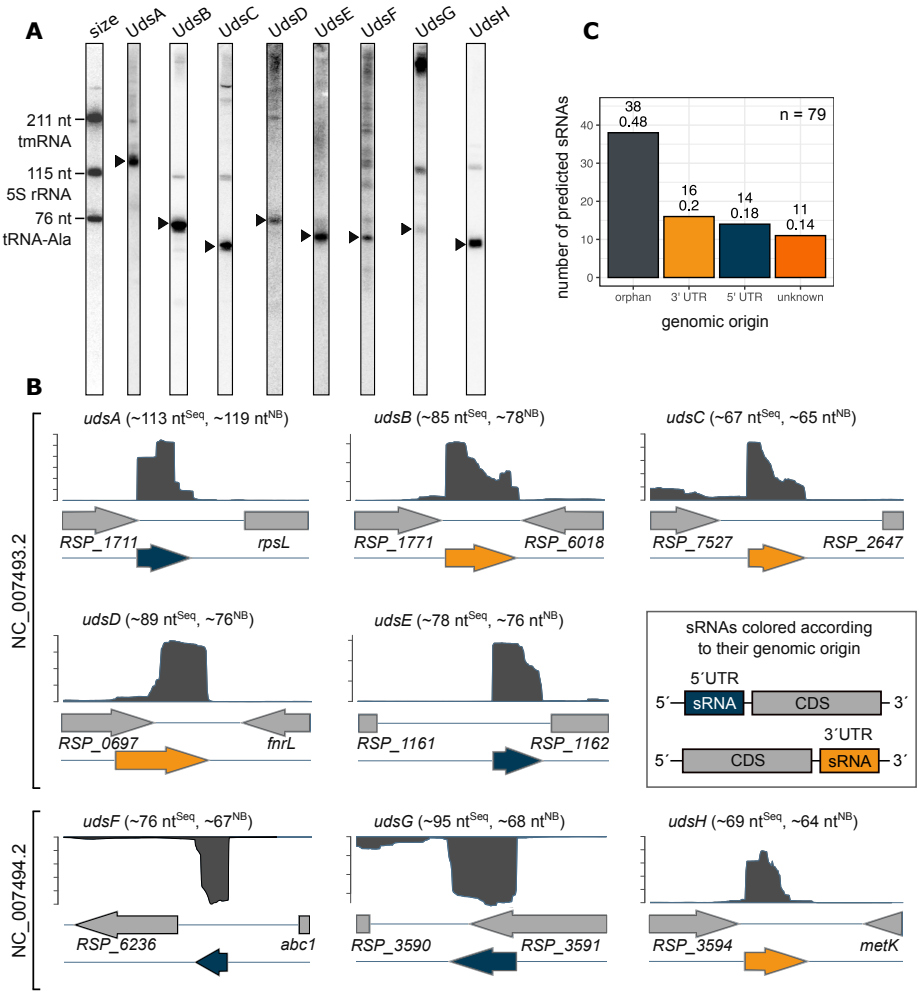


Figure 1. Identification of novel UTR-derived sRNAs in *Rhodobacter sphaeroides*. (**A**) 10 µg of total RNA from exponentially growing *R. sphaeroides* wild type cultures were separated on a denaturating 10% PAA gel and subsequentially blotted. Probes were directed against predicted sRNA sequences. Black triangle marks the mature sRNA. The tmRNA (211 nt), 5S rRNA (115 nt) and tRNA-Ala (76 nt) were used as an internal size standard. 5S rRNA served as loading control. Three sRNAs have previously been mentioned: UdsA (formerly IGR_RSP_1711_rpsL; [37]), UdsC (formerly the sRNA downstream of *RSP_7527*; [16]), and UdsE (formerly RSs2778; [46]). (**B**) Total read coverage of the Uds' loci. Axis' not to scale. Lengths were predicted by RNA sequencing (^{Seq}) and northern blot analysis (^{NB}). (**C**) The genomic origin of all known and predicted sRNAs in *R. sphaeroides* was prediced. Nearly 50 % of all sRNAs originate from orphan genes. total $n = 79$.

Next, a master dataset of sRNAs from *R. sphaeroides* was generated using (a) previously predicted sRNAs ($n = 50$), (b) a list of already described and validated sRNAs ($n = 23$), and (c) novel predicted and validated UTR-derived sRNAs form this study ($n = 5$; general feature file of master dataset see Supplementary Materials). Subsequently, all sRNAs were classified according to their genomic origin using BEDtools window (version 2.25.0, [47]) with predicted 5′/3′ untranslated regions [16]) and the general feature file as input. The majority of sRNAs is classified as orphan (38 of 79 in total), 18 % originate from RNA 5′-UTRs and 20 % from 3′-UTRs (Figure 1C). 14 % of all sRNAs listed in the master database could not properly assigned to any distinct origin.

2.2. Global Prediction of UTR-Derived sRNA Generation Mechanisms

Several different enzymes and features can be involved in the generation of UTR-derived sRNAs. Independently of the origin (5′ or 3′ UTR), an sRNA 5′ end can be generated by a transcription start site (TSS), an endonucleolytic cleavage or by a 5′-to-3′ exoribonuclease (Figure 2A). Options for the generation of the sRNA 3′ ends include a terminator, endonucleolytic cleavage or pausing of 3′-to-5′ processing by an exonuclease such as PNPase. To predict those mechanisms on a genome-wide scale, we first searched for overlaps between all UTR-derived sRNAs, and predicted TSS [16] as well as Rho-independent terminators. Second, all RNA 5′/3′ ends that are dependent on RNase E, RNase III or PNPase were computed with XPEAP (version 1.0.1) as described earlier [37]. Next, all overlaps with these RNA 5′/3′ ends were computed with BEDtools function window (version 2.25.0; [47]). Every feature that is located within a window of −5 nt to +5 nt of every sRNA 5′ or 3′ end is considered as a potential generation mechanism for this particular RNA end. If more than one feature is assigned to that window, all are considered as generation mechanisms in the downstream analysis, since as a matter of principle even two features which are in close proximity can both contribute to the corresponding RNA end generation. RNA 5′/3′ ends lacking any overlap with the input features were classified as unknown regarding the respective mechanism by which they were generated. The analysis reveals that all 5′ ends of 5′ UTR-derived sRNAs are generated by transcription start sites, whereas the 3′ ends are mainly generated by a so far unknown mechanism (Figure 2B). Nevertheless, RNase E, RNase III and also PNPase and a Rho-independent terminator are each responsible for at least one RNA 3′ end. The picture changes when looking at the 5′ ends of 3′ UTR-derived sRNAs: endonucleolytic cleavage by RNase E is predicted to play a major role and accounts for eleven 5′ ends. The second major part is predicted to be generated by TSS. This finding may be biased, because the TSS prediction performed by Remes et al. [48] is based on a comparison between RNA samples that were treated or untreated with TEX (terminator 5′-phosphate dependent 5′-to-3′ exoribonuclease). In the past, we observed that the transcription start site prediction for sRNAs sometimes resulted in false positive hits which may be linked to the high quantity of sRNAs compared to mRNAs. Furthermore, sRNAs are highly structured, which may protect them from degradation by TEX. The 3′ ends of 3′ UTR-derived sRNAs are predicted to depend mainly on terminator structures and also on PNPase. In total, seven of these 3′ ends could not be assigned to any feature.

2.3. Several Enzymes Account for The Maturation and Processing of UTR-Derived sRNAs In Vivo

To compare the previously described predictions for the RNA end formation with in vivo data, strains with deletions of the genes coding for RNase III, RNase J, YbeY, RppH and Hfq were used. As RNase E is essential in *Rhodobacter sphaeroides*, a mutant strain was generated using a thermosensitive RNase E from *E. coli* [20]. This enzyme shows a reduced catalytic activity at 32 °C and is even more impeded at 42 °C. In *R. sphaeroides* PNPase is essential too, so the RNA binding domains KH/S1 were removed by insertion of an in-frame stop codon [37]. Total RNA from the mutant strains *pnp*, $rne^{E.\ coli\ (ts)}$, Δ*rnc*, Δ*rnj*, Δ*ybeY*, Δ*rppH* and Δ*hfq* was isolated from exponentially growing cultures. Next, a northern blot analysis was performed to compare the sRNA levels between the different

Int. J. Mol. Sci. 2021, 22, 12260

mutant strains and the wild type. Most remarkably, the 3′-to-5′exononuclease PNPase is involved in the maturation or processing of all analyzed Uds' (Figure 3A,B). The mature sRNA levels can either be increased (UdsA, Figure 3C) or decreased (UdsB, Figure 3D). PNPase is known to have an important role in degradation of sRNAs that are associated with Hfq [35], a role in sRNA maturation that, to the best of our knowledge, was not reported. The second enzyme with major impact on maturation/processing of the Uds' is the endoribonuclease RNase E that influences the maturation of four of the analyzed sRNAs (UdsB, UdsC, UdsE und UdsH). The mature sRNAs UdsB and UdsC can be detected in the $rne^{E.\ coli\ (ts)}$ mutant strain at the permissive temperature (32 °C) but are not detectable at the non-permissive growth temperature of 42 °C. Instead, precursor molecules are strongly enriched in the mutant strain, indicating that RNase E is required for maturation of these sRNAs (Figure 3D,E). RNase E is known to be important for sRNA maturation in *R. sphaeroides* [16], *Vibrio cholerae* [24,25] and *Salmonella enterica* [26]. Moreover, the endonuclease RNase III accounts for the processing of three Uds' (UdsA, UdsC, UdsH; Figure 3). The RNase J is known to act as 5′-to-3′ exonuclease involved in 23S rRNA maturation in *R. sphaeroides* [38]. It is also involved in the maturation of the sRNA UdsH: precursor RNAs are enriched, whereas the mature sRNA levels are decreased in the Δ*rnj* mutant strain (Figure 3F).

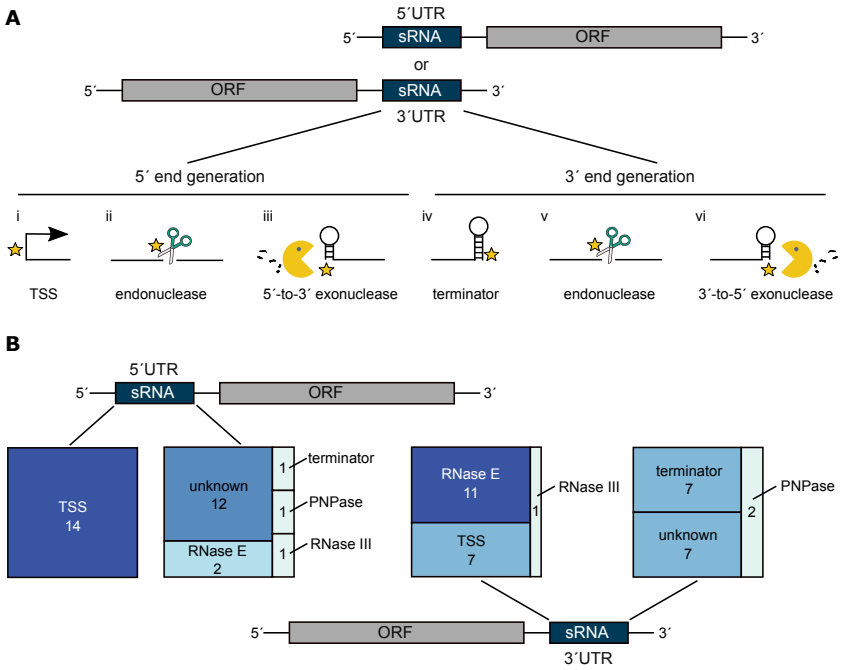


Figure 2. Predicted generation mechanisms of all UTR-derived sRNAs. (**A**) Independently of the location (5′ or 3′ UTR), the sRNA ends can be generated via different mechanisms. The options for the 5′ end generation include a transcription start site (i), an endonucleolytic cleavage (ii) or a 5′-to-3′ processing (iii). A terminator (iv), an endonucleolytic cleavage or an 3′-to-5′ exonuclease can contribute to the RNA 3′ end formation. (**B**) The generation mechanism of all UTR-derived sRNA ends in *R. sphaeroides* was determined using predictions of transcription start sites, Rho-independent terminators and RNase III/RNase E/PNPase-dependent 5′/3′ ends. All 5′ ends of the 5′ UTR-derived sRNAs are generated by transcription start sites, the 3′ ends are mostly formed by an unknown mechanism. RNase E (5′ ends) and Rho-independent terminators (3′ ends) account for the formation and processing of 3′ UTR-derived sRNAs. Nevertheless, unknown factors are likely to contribute to the 3′ end generation.

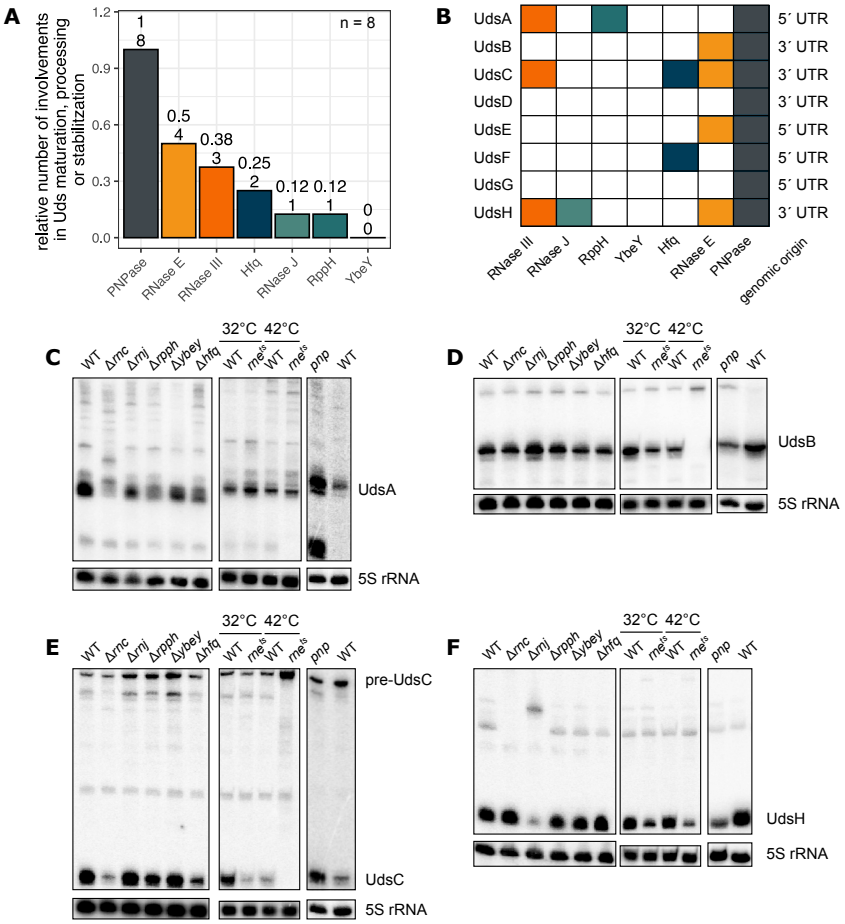


Figure 3. The processing and maturation of UTR-derived sRNAs in *R. sphaeroides* is influenced by various different enzymes. (**A**) Total RNA was isolated from the depicted strains and analyzed by northern blots. If the processing pattern, the abundance of a precursor RNA or the final sRNA varied comparing the wild type and a mutant strain, the respective enzyme was classified as "involved in Uds maturation, processing or stabilization". $n = 8$. Full blots with samples from biological triplicates are shown in Figures S2–S4. (**B**) Summary of the involved enzymes subdivided by individual UTR-derived sRNAs. Colors represent the enzymes also depicted in Figure 3A. (**C**–**F**) Northern blots illustrating the processing and maturation of UdsA, UdsB, UdsC and UdsH. 10 µg of total RNA per lane. 5S rRNA serves as loading control. Northern blot of *pnp* and wild type in (**C**) was first published by Spanka et al. [37] under Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/, accessed on 7 November 2021).

Overall, the predicted generation mechanisms and the enzymes which are involved in the processing reactions in vivo agree in most cases (Table 1). The northern blots for UdsA and UdsG also hint to an involvement of RNase E, the results are, however, not as clear as for other sRNAs (Figures S2A and S4A).

2.4. Growth Conditions Impact the UTR-Derived sRNA Levels

2.4.1. Uds' Abundances Are Growth Phase Dependent

To analyze the impact of growth phase on maturation of UTR-derived sRNAs, *R. sphaeroides* liquid cultures were incubated for 72 h, and total RNA was isolated during the exponential (5 h), early stationary (24 h) and late stationary phases (72 h). More-

over, outgrowth (OG) cultures were inoculated after 24 h and 72 h and cultivated for 1 h (Figure 4). The Uds sRNA levels were strongly dependent on growth phase: UdsA and UdsC were highly abundant in the exponential phase (5 h after inoculation) and during the outgrowth after 24 h but could hardly be detected in the samples from the stationary phase (Figure 4B,D). Other sRNAs such as UdsF and UdsH are highly enriched during the early stationary phase ($\log_2$fold change > 1.25; Figure 4E,F). Furthermore, a general trend could be observed in the samples taken during the late stationary phase, when nearly all sRNA levels showed lower abundances compared to the exponential phase (Figure 4I). The sRNAs UdsB, UdsC and UpsM are processed by RNase E (this study and [20]). Since the precursor RNAs can be detected on northern blots, we quantified these signals and calculated the ratio sRNA/pre-sRNA (Figure 4H). Remarkably this ratio increases after 24 h for the sRNAs UdsB and UpsM (two-sided Student's *t*-test, *p*-value ≤ 0.05), whereas the ratio is not significantly changed comparing the exponential phase and the 24 h outgrowth (two-sided Student's *t*-test, *p*-value > 0.05). This observation suggests an increased processing of the sRNAs UdsB and UpsM by RNase E during the early stationary phase.

Table 1. Comparison of the predicted generation mechanisms for the described UTR-derived sRNAs and the determined enzymes, which are involved in processing and maturation in vivo.

sRNA	Origin	Predicted 5′ End	Predicted 3′ End	Involvement of RNases/Hfq In Vivo
UdsA	5′ UTR	TSS	RNase E	RNase III, RppH, PNPase
UdsB	3′ UTR	RNase E	terminator	RNase E, PNPase
UdsC	3′ UTR	RNase E	terminator	RNase III, RNase E, PNPase, Hfq
UdsD	3′ UTR	RNase E	terminator/PNPase	PNPase
UdsE	5′ UTR	TSS	unknown	RNase E, PNPase
UdsF	5′ UTR	TSS	unknown	Hfq, PNPase
UdsG	5′ UTR	TSS	RNase E	PNPase
UdsH	3′ UTR	RNase E	terminator	RNase III, RNase J, RNase E, PNPase

To investigate the influence of an altered promoter activity on the sRNA levels, transcriptional promoter fusions were constructed. For the 5′ Uds', only the promoter sequence was fused to mVenus. In contrast to that, two sequences were used for each of the 3′ Uds': the upstream coding sequence (CDS) and a longer fragment containing the promoter of the upstream gene and the CDS (Figure 5A,B). This strategy allows the detection of putative internal promoters in case of 3′ UTR-derived sRNAs. According to the upstream sequences, the promoters of *udsA* and *udsH* depend on the alternative sigma factors $RpoH_I/RpoH_{II}$ and the promoter of *RSP_7527-udsC* is $RpoH_{II}$ dependent (Table S1). Promoter sequences of the other UTR-derived sRNAs do not accord with the published $RpoH_I/RpoH_{II}$ or $RpoH_{II}$ consensus motifs [49]. The growth experiment was repeated with wild type strains harbouring the described plasmids. To avoid misleading results caused by the high protein stability followed by an accumulation of mVenus, samples were only taken after 5 h and 24 h of cultivation and from the outgrowth cultures. The normalized fluorescence intensities varied substantially among the different promoter constructs, ranging from an F/OD of 100 (UdsH) to 4000 (UdsB) after 5 h of cultivation (Figure S6). For those two sRNAs in particular, the constructs harbouring only the CDS exhibited a detectable fluorescence signal, which was nevertheless lower than the signal from the promoter + CDS constructs (Figure S6). This indicates a transcription of the sRNA by two promoters, one belonging to the cotranscribed gene and one located in the open reading frame. Except for the UdsG promoter, the activity of all other promoters was increased or decreased when comparing the early stationary phase and respective outgrowth to the exponential phase. The 3′ derived UdsC and UdsD are exclusively transcribed by the promoter of the upstream gene, while for UdsB and UdsH additional promoter activity within the upstream coding region was detected.

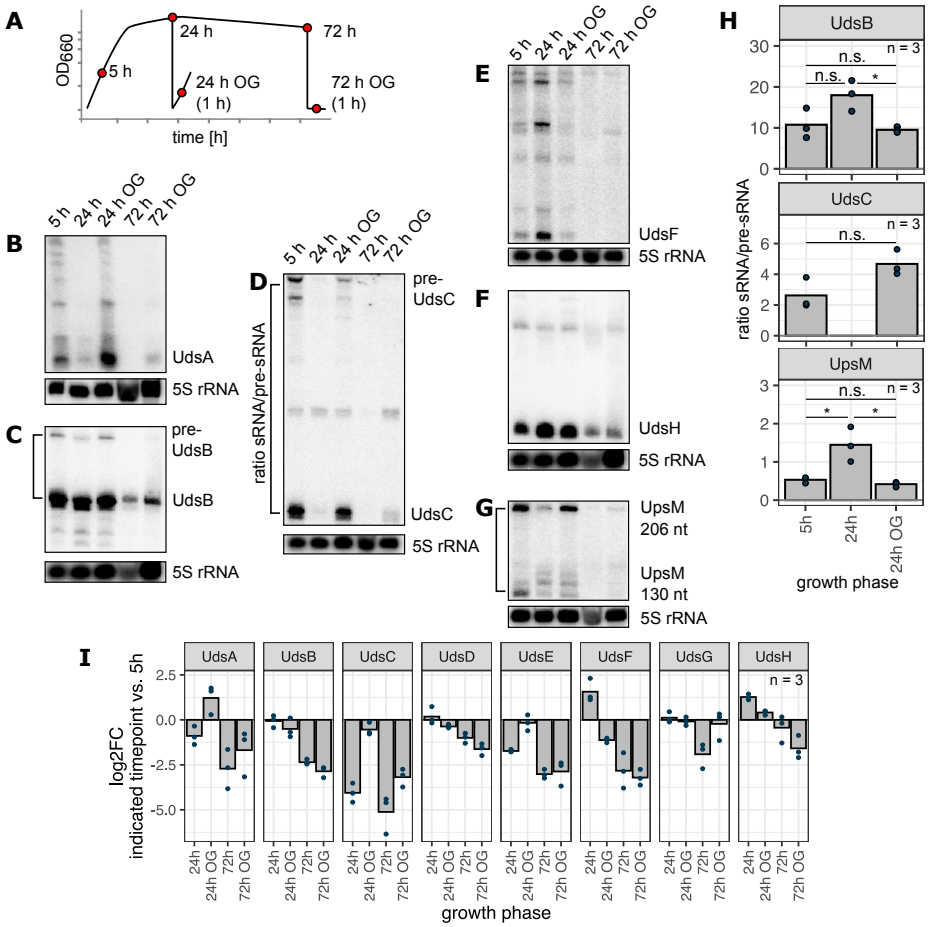


Figure 4. The UTR-derived sRNA levels strongly depend on the growth phase. (**A**) *R. sphaeroides* wild type cultures were incubated for 72 h under microaerobic conditions. Cells were harvested at the depicted times (red dots). After 24 h and 72 h, outgrowth cultures were inoculated. (**B**–**G**) Total RNA from the depicted samples was analyzed by northern blot using specific probes against UdsA to UdsH and UpsM. The sRNA levels vary throughout the different growth phases and can be increased or decreased when compared to the levels during the exponential phase. Loading control: 5S rRNA. Full blots with samples from biological triplicates are shown in Figure S5. Membranes were used with multiple probes: UdsA, UdsD and UdsE; UdsB, UdsG,UdsH and UpsM; UdsC and UdsF. (**H**) Signals of the sRNAs UdsB, UdsC and UpsM and their corresponding pre-sRNAs were quantified and the ratio was calculated (y-axis). Bars indicate the mean value, every dot represents one biological replicate, $n = 3$. Groups were compared with the two-sided Student's *t*-test: * *p*-value < 0.05; n.s. not significant. (**I**) $\log_2$fold changes were computed comparing the sRNA abundances during the growth phases with the respective signal in the 5 h sample. Quantification based on northern blot data, $n = 3$.

Next, we compared the observed sRNA levels and corresponding promoter activities using a trend heatmap (Figure 5C). For every sRNA and promoter construct the relative change in signal intensity comparing the 24 h and outgrowth sample to the exponential phase was computed. Samples with a $\log_2$fold change > 0.65 were classified as "increased", $\log_2$fold change < -0.65 as "decreased" and all others as "no change". In case of UdsA the changes of sRNA level and promoter activity show the same trend in the 24 h outgrowth cultures. For all other sRNAs the changes of sRNA levels cannot be solely due to changed promoter activity. This points to an important role of sRNA maturation in the growth phase dependent expression of sRNAs.

Int. J. Mol. Sci. 2021, 22, 12260

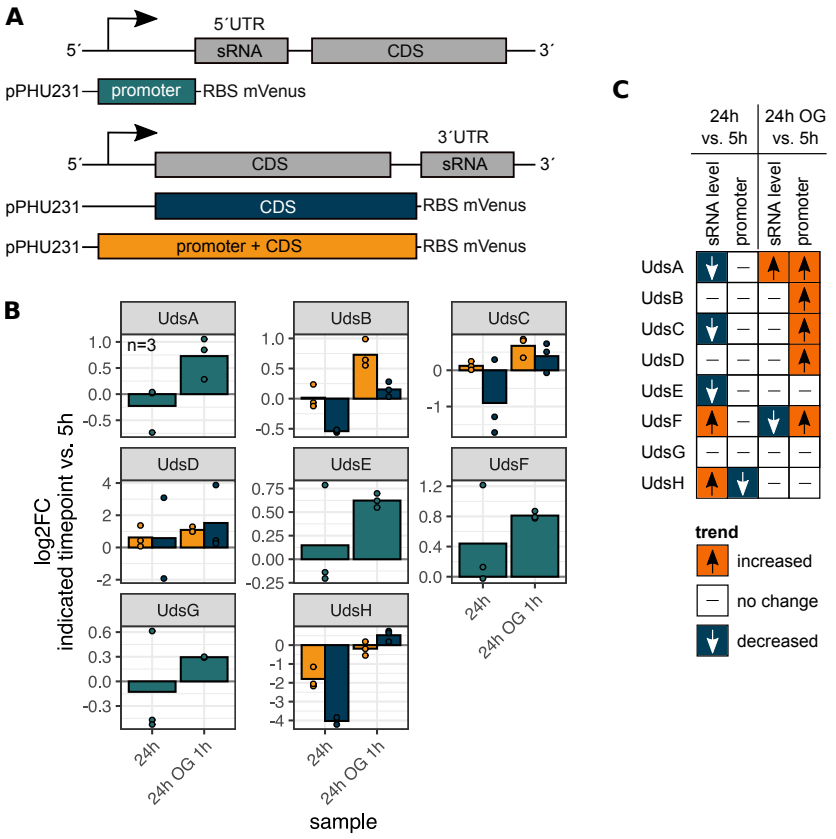


Figure 5. The UTR-derived sRNA promoter activity is growth phase dependent but does not represent the observed sRNA level in all cases. (**A**) Overview of the transcriptional promoter fusions used in this study. (**B**) The plasmids were conjugated in the wild type strain and the fluorescence intensity was measured after 5 h, 24 h and from the outgrowth culture. x-axis: sample. y-axis: $\log_2$fold change of the indicated timepoint vs. signal during exponential growth phase (5 h). Green: promoter sequence. Blue: coding sequence (CDS). Yellow: promoter + coding sequence. $n = 3$. Every dot represents the mean value of two technical replicates. Signals F/OD_{660} are shown in Figure S6. (**C**) Classification of the sRNA level and the promoter activity, based on northern blot data and fluorescence intensities. Red: increased compared to the exponential phase. Blue: decreased compared to the exponential phase. White with horizontal dash: no change.

2.4.2. External Stressors Affect the sRNA Abundances

Previous studies highlighted the important role of sRNAs during the oxidative stress response in *R. sphaeroides* (e.g., [46,50,51]). We asked whether stress conditions also affect the levels of UTR-derived sRNAs by influencing their maturation. Wild type cultures were incubated with 1 mM H_2O_2, grown in $^1O_2$ generating conditions or exposed to a heat shock at 42 °C. RNA samples were taken before and after the treatment and subsequentially analyzed via northern blot (Figure 6). The abundances of nearly all Uds' are influenced by at least one external stressor; only UdsD and UdsH showed a more or less stable signal independently of the growth condition (Figure 6A). The sRNA UdsC showed a strong dependence on oxidative stress that was induced by hydrogen peroxide and singlet oxygen (mean $\log_2$fold change > 0.65, Figure 6B). A general trend could be observed after heat shock induction, since all sRNA abundances were reduced except of the UdsD and UdsH levels.

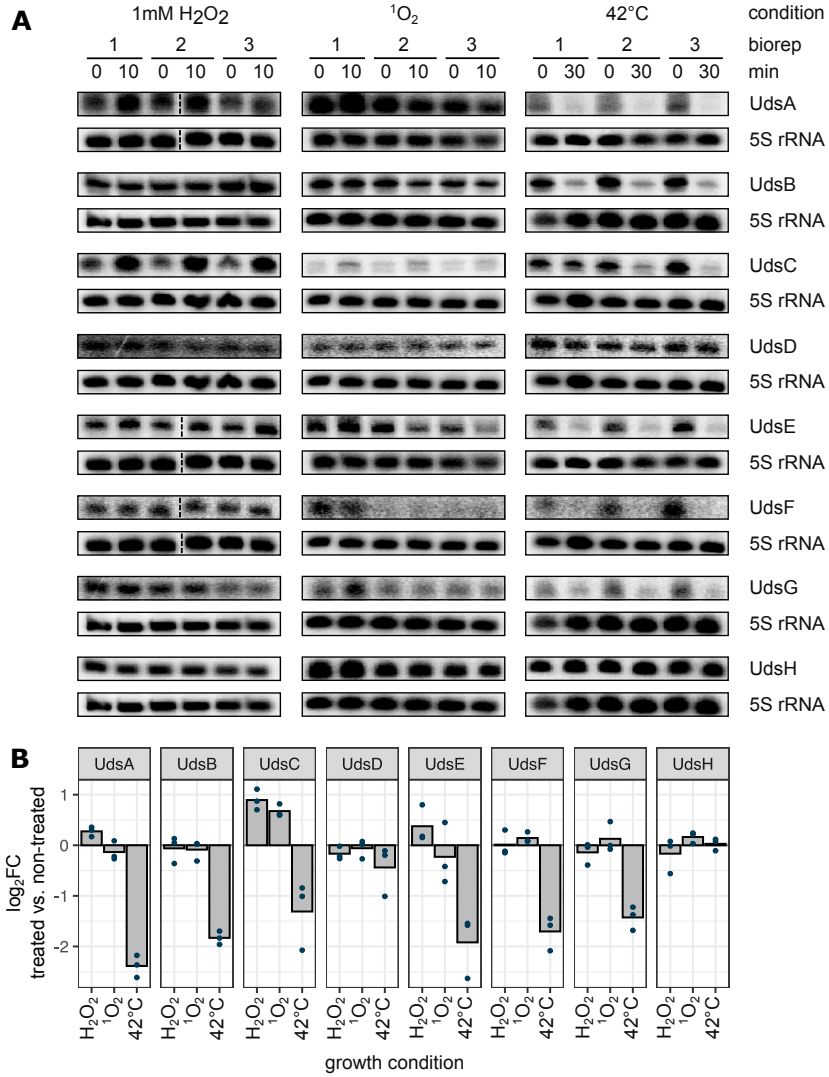


Figure 6. UTR-derived sRNAs are influenced by various stress conditions. (**A**) Exponentially growing *R. sphaeroides* liquid cultures were exposed to oxidative (1 mM H_2O_2, 10 min), singlet oxygen ($^1O_2$, 10 min) or heat stress (42 °C, 30 min). Samples for RNA isolation were harvested before and after the indicated time. Northern blot analysis of biological triplicates, 10 µg total RNA per lane. 5S rRNA served as loading control. Uncut northern blots are shown in Figures S7 and S8. (**B**) $\log_2$fold changes (treated vs. non-treated) were computed based on the northern blot data. Every dot represents one biological replicate. $n = 3$.

All promoter activities under stress conditions were tested for those sRNAs which exhibited $\log_2$fold changes > 0.65 or <-0.65 on the northern blot under stress conditions (eight combinations of stress condition and construct in total, Figure 7). We only observed an increased fluorescence intensity for the promoter construct of UdsC under H_2O_2 stress and a decreased signal of the UdsC promoter during singlet oxygen stress (Figure 7A,B). In contrast to that, a shift to 42 °C led mainly to constant signals and only the strain harbouring the promoter fusion of *RSP_0557* (positive control) showed an increasing fluorescence intensity over time (Figure 7C). Next, the change in fluorescence signal was categorized

to visualize the trends as described above (Figure 7D). Similar to the analyzed growth experiments, only one comparison showed the same trend between the sRNA levels and the corresponding promoter activities. These results strongly suggest that UTR-derived sRNA levels rely on transcription rate and factors like processing events or altered degradation rates which contribute to the mature sRNA abundances. Nevertheless, no specific precursor RNAs or enriched degradation products could be found that might be linked to the function of one specific ribonuclease responsible for that particular processing reaction.

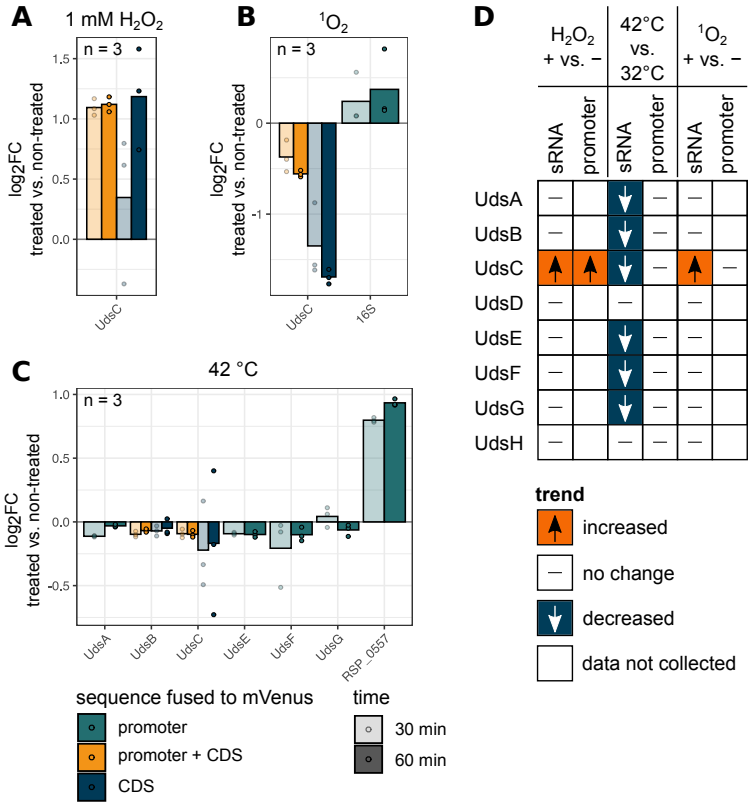


Figure 7. Oxidative and heat stress have a major impact on several UTR-derived sRNA promoter activities. Fluorescence intensity of the wild type harbouring the indicated plasmids was measured before and after treatment with 1 mM H_2O_2 (**A**), $^1O_2$ stress (**B**), 42 °C heat stress (**C**) and the $\log_2$fold changes were calculated. x-axis: sample. y-axis: $\log_2$fold change treated vs. non-treated. Green: promoter sequence. Blue: coding sequence (CDS). Yellow: promoter + coding sequence. Color intensity indicates duration of induction. $n = 3$. Every dot represents the mean value of two technical replicates. The promoters of *RSP_0557* (unpublished data) and 16S rRNA (McIntosh et al., 2019) were used as positive controls for the indicated growth conditions. Signals F/OD_{660} are shown in Figure S9. (**D**) Classification of the sRNA level and the promoter activity, based on northern blot data and fluorescence intensities. Red: increased compared to non-treated sample. White with horizontal dash: no change. Blue: decreased compared to non-treated sample. White without dash: data not collected.

Analyzing RNA-Seq data from *R. sphaeroides* grown under iron limitation revealed that the RNase E generated UdsB was slightly less abundant compared to the control grown in media with supplemented iron. However, the *RSP_1771* part of the *RSP_1771-udsB* cotranscript was more abundant. This prompted us to investigate if the RNase E mediated processing of UTR-derived sRNAs is influenced by iron availability during the exponential growth phase. Total RNA was isolated from *R. sphaeroides* cultures grown in malate minimal

media supplemented with iron or under iron depleted conditions [52] and analyzed via northern blot. RNase E is involved in maturation of SorX, UpsM, CcsR1-4, UdsB and UdsC from precursors ([15,17,20]; this study). We therefore quantified the mature sRNA levels and, if possible, the respective precursor RNAs from northern blots (Figure 8A–E). Next, the $\log_2$fold changes between iron replete and iron deplete conditions and the signal ratios were computed (Figure 8F,G). We observed that the ratio of sRNA to precursor sRNA is significantly reduced for SorX and UdsC when the cultures were grown in iron depleted medium, indicating a reduced processing rate by RNase E for these sRNAs (two-sided Student's *t*-test, *p*-value < 0.001). In contrast to that, the ratio of UdsB/pre-UdsB was not significantly decreased, and in the case of UpsM, the ratio even increased (Figure 8C,G). Our results indicate that RNase E processing of UTR-derived sRNAs is modulated under iron limiting growth conditions in a substrate dependent manner.

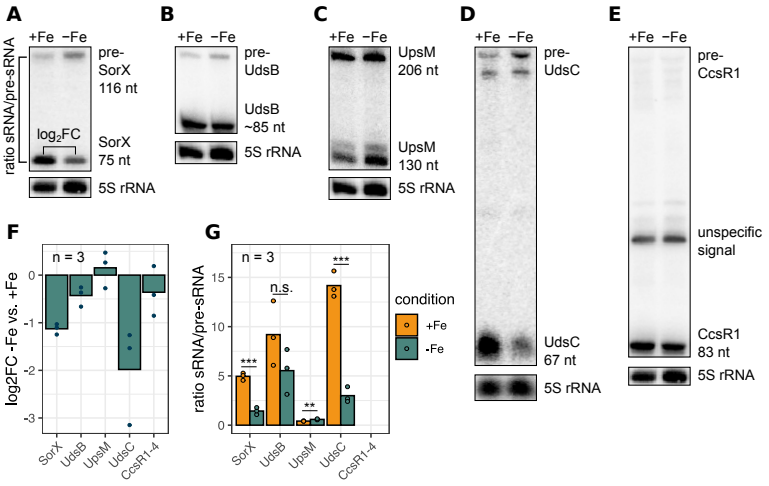


Figure 8. Iron availability influences the RNase E dependent processing of UTR-derived sRNAs. Total RNA from *R. sphaeroides* wild type cultures grown in media supplemented with iron (+Fe) or after iron depletion (−Fe) was isolated and analyzed via northern blot. Membranes were hybridized with probes directed against sRNAs which are processed by RNase E: SorX (**A**), UdsB (**B**), UpsM (**C**), UdsC (**D**) and CcsR1 (**E**). The experiment was performed in biological triplicates. 5S rRNA served as loading control. Full blots with samples from biological triplicates are shown in Figure S10. (**F**) $\log_2$fold changes of the mature sRNA species were calculated comparing the −Fe and +Fe conditions. Grey bars indicate the mean value, every dot represents one biological replicate, $n = 3$. (**G**) Signals of the sRNAs and pre-sRNAs were quantified and the ratio was calculated (y-axis). Bars indicate the mean value (+Fe: yellow, −Fe: green), every dot represents one biological replicate, $n = 3$. Groups were compared with the two-sided Student's *t*-test: *** *p*-value < 0.001; ** *p*-value < 0.01; n.s. not significant. Signal of the CcsR1 precursor was too low for quantification.

3. Discussion

Our data demonstrate that the maturation of sRNAs from UTRs is an important step for the control of sRNA levels. As a result, levels of the co-transcribed mRNA and sRNA can respond differently to environmental changes. Why may such differential regulation be appropriate? To address this question, it is important to know the function of the mRNA and the function of the sRNA, which is unfortunately the case for only a few examples. Transcription of mRNA and the UTR-derived sRNA from the same promoter leads to the production of similar levels of both, and to the same transcriptional regulation. This seems reasonable if both RNAs have a function in the same pathway and/or affect the same physiological process as already shown for some UTR-derived sRNAs. e.g., CpxQ is derived from the 3′ UTR of *cpxP* and both RNAs are involved in the inner membrane

stress response in *Salmonella enterica* [26]. In the same bacterium, NarS is derived from the 3′ UTR of *narK* that encodes a nitrate transporter. NarS is involved in the cross-regulation of nitrate and nitrite transport [53]. In enterohemorrhagic *E. coli*, StxS is derived from the 5′ UTR of *stx1AB* for Shiga toxin 1 by premature transcriptional termination. StxS represses Shiga toxin 1 production under lysogenic conditions [54]. In *Pseudomonas aeruginosa rhlI* encodes the enzyme for AHL synthesis. RhlS is derived from its 5′ UTR of *rhlI* and is required for the production of normal levels of AHL [55]. In *E. coli*, MalH is derived from the 3′ UTR of the maltose uptake operon *malEFG* and contributes to alternative carbon source utilization by affecting maltoporin expression [56]. The product of *argR* and the 3′ UTR-derived ArgX, both regulate the arginine deiminase pathway in *Lactococcus lactis* [57]. Some UTR-derived sRNAs were also characterized in *R. sphaeroides*: the 3′ UTR-derived SorX and the *ompR1* mRNA, both function in the oxidative stress response [15,58]. The 3′ UTR-derived PcrX RNA and the *puf* operon are required for formation of photosynthetic complexes and their regulation [19]. The CcsR1-4 RNAs are derived from the 3′ UTR of *ccaF1* (*RSP_6037*), which encodes a small RNA-binding protein required for the maturation of CcsR and other sRNAs [17,18]. It seems reasonable that all these sRNAs are under the control of the same promoter as the related mRNA. Why have another level of regulation at the step of sRNA maturation?

Many bacterial genes are organized in polycistronic operons and consequently regulated by the same promoter. Nevertheless, an additional regulation at the posttranscriptional level can result in a different abundance of mRNA segments that determine the stoichiometry of the resulting proteins. In case the of e.g., the *puf* operon of *R. capsulatus*, this is due to segmental differences in mRNA stability [59]. Differences in initiation of translation for individual genes of an operon were demonstrated e.g., for the *atp* operon [60] or the *gal* operon [61] in *E. coli*. Partial transcriptional termination leads to differential expression of the genes in e.g., the *E. coli rpsO-pnp* operon [62]. It is conceivable that in case of UTR-derived sRNAs the change of the ratio of the sRNA and mRNA is favorable under certain environmental conditions, but this needs to be tested in the future.

Although our data demonstrate that environmental factors can influence maturation of UTR-derived sRNA and also point to some of the mechanisms involved in the maturation of the individual sRNAs, the exact mechanisms underlying the regulation need further investigation. If a second promoter is contributing to the generation of a 3′ UTR-derived sRNA, this may of course lead to an expression pattern that is different from that of the mRNA. But transcriptional regulation may also account for different levels of RNases or indirectly affect the level of an UTR-derived sRNA. In case of the *R. sphaeroides* sRNA UpsM, growth phase-dependent levels are mediated by base pairing to another sRNA, StsR [21]. UpsM is derived from the 5′ UTR of the *dcw* (division and cell wall synthesis) gene cluster in *R. sphaeroides* by partial transcriptional termination [20]. The orphan sRNA StsR is induced during the stationary growth phase by the alternative sigma factors $RpoH_{I/II}$ [48]. StsR base pairs to UpsM and the 5′ UTR of the *dcw* genes, resulting in a structural change which gives access to an RNase E cleavage site within the *upsM* sequence. Interestingly, the interaction to StsR and the subsequent cleavage of the *dcw* 5′ UTR also affects read through into the *dcw* genes [21].

Furthermore, the amount or activity of the RNases may be altered in response to environmental cues. Changing amounts of RNases can be due to transcriptional regulation or proteolysis. Quantitative mass spectrometry data from a previous study [63] revealed that the protein levels of several RNases vary through the different growth phases, which includes significant changes in RNase E, III, P, PNPase and also Hfq levels (Figure 9). This may influence both the processing and stability of UTR-derived sRNAs and thus contribute to a modulation of sRNA level dependent on the growth phase. Moreover, post-translational modifications can alter the stability of RNases (reviewed in [64]). e.g., higher levels of RNase R were reported under stress or in stationary phase in *E. coli* [65] and attributed to stress-dependent reduction of acetylation [66]. The activity of RNases can be affected by post-transcriptional modifications or by cellular localization (reviewed in [64]).

This might also be the case for RNase E: We observed an increased RNase E mediated processing of two UTR-derived sRNAs during the early stationary growth phase (Figure 4), although the protein level was about 35% reduced (Figure 9) and the promoter activities of these UTR-derived sRNAs remained constant (Figure 5). In case of heat stress, it is also conceivable that changes in RNA structure can lead to altered maturation. Our study reported an influence of iron depletion in the maturation of several sRNAs by RNase E. However, the RNase E dependent maturation of other sRNAs was not influenced by iron availability, excluding a general effect of iron availability on RNase E activity.

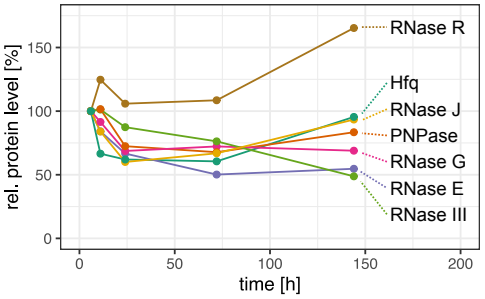


Figure 9. The levels of several RNases and of the RNA chaperon Hfq vary at different stages of growth as determined by quantitative mass spectrometry. *R. sphaeroides* wild type cultures were grown under microaerobic conditions and sampled in mid exponential phase ($OD_{660} = 0.5$), in transition to stationary phase (11 h after inoculation, $OD_{660} \approx 1.0$), in early stationary phase (24 h after inoculation, $OD_{660} \approx 1.8$) and in late stationary phase (72 and 144 after inoculation, $OD_{660} \approx 1.1$) and a quantitative proteome analysis was performed as described in Bathke et al. [63]. Values for the RNases and for Hfq are taken from the data set of this publication.

Most likely, it will not be possible to address the exact mechanisms underlying regulated sRNA maturation at a global scale. A better understanding of the importance and the role of UTR-derived RNAs will need a closer look at the maturation processes and their regulation in the future.

4. Material and Methods

4.1. Bacterial Strains and Growth Conditions

The strains used in this study are described in Table S2. Erlenmeyer flasks with a volume of 50 mL were filled with 40 mL of malate minimal media. Microaerobic *Rhodobacter sphaeroides* (recently renamed *Cereibacter sphaeroides* [67]) cultures were incubated at 32 °C under continuous shaking in the dark, resulting in a dissolved oxygen concentration of 25 µM to 30 µM [68]. To apply organic peroxide stress, H_2O_2 (1 mM final concentration) was added to the liquid cultures. Aerobic cultivation with induction of photooxidative stress was performed as described by Glaeser and Klug [69]. Briefly, microaerobic liquid cultures were shifted to aerobic growth conditions (approximately 180 µM). They were cultivated in the dark at 32 °C in air-gassed Meplat flasks. Methylen blue acts as a photosensetizer and was added in a final concentration of 0.2 µM. During the exponential growth phase, the cultures were exposed to white light (800 W m^2) to induce the generation of $^1O_2$. For the heat shock experiments, pre- and main cultures were incubated at 32 °C in the dark under microaerobic conditions. During the exponential growth phase, cultures were shifted to a 42 °C preheated water bath where they were incubated for 30 min under continuous shaking. To generate iron limitation, cultures were treated as described in Remes et al. [68]. Cultures were grown in medium without supplemented iron with 2,2′dipyridyl (30 µM, Merck) for three times. In the last pre-culture and in the experimental culture, no 2,2′dipyridyl was added.

4.2. Construction of a rppH and a ybeY Deletion Strain

Deletion of the gene *rppH* (*RSP_0931*) in the *Rhodobacter sphaeroides* 2.4.1 wild type strain [70] was carried out by homologous recombination and insertion of a kanamycin resistance gene. The up and down fragments were amplified by PCR using the primer pairs KO_RSP0931_up_f /KO_RSP0931_up_r and KO_RSP0931_dw_f /KO_RSP0931_dw_r. Both fragments were cloned in pPHU281 with EcoRI/BamHI and BamHI/HindIII. The kanamycin resistance gene was inserted between the fragments with BamHI. The plasmid was transformed to *E. coli* S17-1 and then transferred to *Rhodobacter sphaeroides* 2.4.1 by diparental conjugation. Positive clones were selected on malate minimal agar containing 25 µg mL^{-1} kanamycin.

The same procedure was also applied to delete the gene *ybeY* (*RSP_3598*) in the *Rhodobacter sphaeroides* 2.4.1 wild type strain. For the up and down fragment amplification the primer pairs KO_3598_ybeY_up_f/KO_3598_ybeY_up_r and KO_3598_ybeY_dw_f /KO_3598_ybeY_dw_r were used. Plasmid construction was carried out as described above but with a gentamicin resistance gene instead (taken from pPHU45Ω). Clones were selected on malate minimal agar containing 10 µg mL^{-1} gentamicin.

4.3. Promoter Activity Assay

DNA fragments harbouring the putative promoter sequences of every UTR-derived sRNA were amplified and fused to the mVenus gene using plasmid pPHU231 as described by Charoenpanich et al. [71] and McIntosh et al. [72]. Restriction enzyme cleavage sites (HindIII/XbaI) were incorporated via the primer sequences as well as a strong ribosome binding site (AGGGGAGAAG). Final plasmids were conjugated to the *Rhodobacter sphaeroides* wild type using the *E. coli* S17-1 strain. Liquid cultures were incubated as described above and prediluted to an OD_{660} of 0.15. Volumes of 100 µL liquid culture were transferred to transparent 96-well plates and fluorescence was subsequently measured in the Tecan Infinite M Nano (Tecan Group AG). Primer sequences and cloned constructs are provided in Tables S3 and S4.

4.4. Reverse Transcription (RT) PCR

RT-PCR was performed using the Brilliant III Ultra-Fast SYBR Green QRT-PCR Master Mix (Agilent #600886) according to the manufacturer's manual. DNA free total RNA extracted from exponentially growing wildtype cultures served as template for the reaction. The RT-PCR products were separated on 10% polyacrylamide gels and visualized with ethidium bromide staining.

4.5. Northern Blot Analysis

Total RNA was isolated with the hot phenol method [73]. DNase treatment was performed according to the manufacturer's instructions (Invitrogen #AM1907). The electrophoretic separation on denaturing PAA urea gels was conducted as described by Berghoff et al. [46]. The oligonucleotides were end-labeled using T4 polynucleotide kinase (T4-PNK, Thermo Scientific, Waltham, MA, USA) with [γ^{32}P]-ATP (SRP-301, Hartmann Analytic) according to the manufacturer's protocol. Oligonucleotides used in this study are listed in Table S4. The membranes were washed in 5x SSC buffer after overnight incubation with the labeled oligonucleotides. Sealed membranes were then exposed to a screen for 48 h. The QuantityOne 1-D Analysis Software (BioRad, version 4.6.6) was used to quantify the signals. All sRNA signals were normalized to the 5S rRNA signal which was used as a loading control.

4.6. Bioinformatical Analysis

All differential RNA 5′ and 3′ ends which are RNase E-, RNase III- or PNPase-dependent were identified with XPEAP as described earlier [37]. Parameters used for the computation were: $\log_2$fold change cutoff ≤ -1 or $\geq +1$; adjusted p-value ≤ 0.05 (Benjamini-Hochberg algorithm). Rho-independent transcription terminator prediction

was performed using TransTherm HP [74]. Prediction of 5′/3′UTRs and TSS was carried out by Remes et al. [48] based on differential NGS RNA-Seq data from *Rhodobacter sphaeroides* 2.4.1. All known and predicted sRNAs in *Rhodobacter sphaeroides* 2.4.1 were first classified according to their genomic origin. Overlaps between annotated sRNAs and predicted 5′/3′UTRs were computed with BEDtools window (version 2.25.0, options -s -wa -wb; [47]). Further, for every annotated sRNA all overlaps with (a) predicted transcription start sites (TSS), (b) Rho-independent transcription terminators and (c) differential RNA 5′/3′ ends that depend on RNase E, RNase III or PNPase were identified using the same function with a window size of 5 nt (10 nt for Rho-independent terminators). Next, windows from position −5 nt to +5 nt at 5′ and 3′ ends of all sRNAs were defined. All overlapping features that could be assigned to these windows were considered as the putative generation mechanism of this particular sRNA. Multiple overlaps per site were allowed. The read data of all described mutant strains are deposited on NCBI Gene Expression Omnibus: PNPase and RNase III mutant strains (NCBI GEO accession number: GSE156818) and thermosensitive RNase E mutant strain (NCBI GEO accession number: GSE71844, published in Förstner et al. [16]).

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/ijms222212260/s1.

Author Contributions: Conceptualization, D.-T.S. and G.K.; methodology, D.-T.S. and G.K.; software, D.-T.S.; validation, D.-T.S. and G.K.; formal analysis, D.-T.S.; investigation, D.-T.S.; resources, G.K.; data curation, D.-T.S.; writing—original draft preparation, D.-T.S. and G.K.; writing—review and editing, G.K. and D.-T.S.; visualization, D.-T.S.; supervision, G.K.; project administration, G.K.; funding acquisition, G.K. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by Deutsche Forschungsgemeinschaft (DFG, GRK 2355).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All RNA sequencing data is published at the NCBI Gene Expression Omnibus: PNPase and RNase III mutant strains (NCBI GEO accession number: GSE156818) and thermosensitive RNase E mutant strain (NCBI GEO accession number: GSE71844, published in Förstner et al. [16]).

Acknowledgments: We thank Tom Rische-Grahl and Kerstin Haberzettl for cloning the Δ*ybeY* and Δ*rppH* mutant strains. Moreover, we are greatful for the help provided by Andreas Jäger during RNA isolation. We further thank Andrea Weisert for cloning the promoter fusions. The construct pPHU231-P_{0557}-mVenus was kindly provided by Matthew McIntosh.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CDS	coding sequence
OD	optical density
OG	outgrowth
RT	reverse transcription
TEX	terminator 5′-phosphate dependent 5′-to-3′ exoribonuclease
TSS	transcription start site
UTR	untranslated region
Uds	UTR-derived sRNA

References

1. Kim, S.; Kim, Y.; Suh, D.H.; Lee, C.H.; Yoo, S.M.; Lee, S.Y.; Yoon, S.H. Heat-responsive and time-resolved transcriptome and metabolome analyses of *Escherichia coli* uncover thermo-tolerant mechanisms. *Sci. Rep.* **2020**, *10*, 17715. [CrossRef] [CrossRef]
2. Fei, Y.Y.; Bhat, J.A.; Gai, J.Y.; Zhao, T.J. Global Transcriptome Profiling of *Enterobacter* Strain NRS-1 in Response to Hydrogen Peroxide Stress Treatment. *Appl. Biochem. Biotechnol.* **2020**, *191*, 1638–1652. [CrossRef] [CrossRef]
3. Bronowski, C.; Mustafa, K.; Goodhead, I.; James, C.E.; Nelson, C.; Lucaci, A.; Wigley, P.; Humphrey, T.J.; Williams, N.J.; Winstanley, C. *Campylobacter jejuni* transcriptome changes during loss of culturability in water. *PLoS ONE* **2017**, *12*, e0188936. [CrossRef] [PubMed] [CrossRef]
4. Nogueira, T.; Springer, M. Post-transcriptional control by global regulators of gene expression in bacteria. *Curr. Opin. Microbiol.* **2000**, *3*, 154–158. [CrossRef] [CrossRef]
5. Wassarman, K.M.; Zhang, A.; Storz, G. Small RNAs in *Escherichia coli*. *Trends Microbiol.* **1999**, *7*, 37–45. [CrossRef] [CrossRef]
6. Masse, E.; Gottesman, S. A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 4620–4625. [CrossRef] [CrossRef]
7. Lalaouna, D.; Baude, J.; Wu, Z.; Tomasini, A.; Chicher, J.; Marzi, S.; Vandenesch, F.; Romby, P.; Caldelari, I.; Moreau, K. RsaC sRNA modulates the oxidative stress response of *Staphylococcus aureus* during manganese starvation. *Nucleic Acids Res.* **2019**, *47*, 9871–9887. [CrossRef] [CrossRef] [PubMed]
8. Hör, J.; Matera, G.; Vogel, J.; Gottesman, S.; Storz, G. Trans-Acting Small RNAs and Their Effects on Gene Expression in *Escherichia coli* and *Salmonella enterica*. *EcoSal Plus* **2020**, *9*. [CrossRef] [CrossRef]
9. Braatsch, S.; Gomelsky, M.; Kuphal, S.; Klug, G. A single flavoprotein, AppA, integrates both redox and light signals in *Rhodobacter sphaeroides*. *Mol. Microbiol.* **2002**, *45*, 827–836. [CrossRef] [CrossRef]
10. Metz, S.; Haberzettl, K.; Frühwirth, S.; Teich, K.; Hasewinkel, C.; Klug, G. Interaction of two photoreceptors in the regulation of bacterial photosynthesis genes. *Nucleic Acids Res.* **2012**, *40*, 5901–5909. [CrossRef] [CrossRef]
11. Mank, N.N.; Berghoff, B.A.; Hermanns, Y.N.; Klug, G. Regulation of bacterial photosynthesis genes by the small noncoding RNA PcrZ. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 16306–16311. [CrossRef] [PubMed] [CrossRef] [PubMed]
12. Happ, H.N.; Braatsch, S.; Broschek, V.; Osterloh, L.; Klug, G. Light-dependent regulation of photosynthesis genes in *Rhodobacter sphaeroides* 2.4.1 is coordinately controlled by photosynthetic electron transport via the PrrBA two-component system and the photoreceptor AppA. *Mol. Microbiol.* **2005**, *58*, 903–914. [CrossRef] [CrossRef] [PubMed]
13. Berghoff, B.A.; Konzer, A.; Mank, N.N.; Looso, M.; Rische, T.; Förstner, K.U.; Krüger, M.; Klug, G. Integrative "Omics"-Approach Discovers Dynamic and Regulatory Features of Bacterial Stress Responses. *PLoS Genet.* **2013**, *9*, e1003576. [CrossRef] [CrossRef]
14. Zeilstra-Ryalls, J.H.; Kaplan, S. Oxygen intervention in the regulation of gene expression: the photosynthetic bacterial paradigm. *Cell. Mol. Life Sci. (CMLS)* **2004**, *61*, 417–436. [CrossRef] [PubMed] [CrossRef]
15. Peng, T.; Berghoff, B.A.; Oh, J.I.; Weber, L.; Schirmer, J.; Schwarz, J.; Glaeser, J.; Klug, G. Regulation of a polyamine transporter by the conserved 3′ UTR-derived sRNA SorX confers resistance to singlet oxygen and organic hydroperoxides in *Rhodobacter sphaeroides*. *RNA Biol.* **2016**, *13*, 988–999. [CrossRef] [CrossRef] [PubMed]
16. Förstner, K.U.; Reuscher, C.M.; Haberzettl, K.; Weber, L.; Klug, G. RNase E cleavage shapes the transcriptome of *Rhodobacter sphaeroides* and strongly impacts phototrophic growth. *Life Sci. Alliance* **2018**, *1*, e201800080. [CrossRef] [CrossRef]
17. Billenkamp, F.; Peng, T.; Berghoff, B.A.; Klug, G. A Cluster of Four Homologous Small RNAs Modulates C1 Metabolism and the Pyruvate Dehydrogenase Complex in *Rhodobacter sphaeroides* under Various Stress Conditions. *J. Bacteriol.* **2015**, *197*, 1839–1852. [CrossRef] [CrossRef]
18. Grützner, J.; Billenkamp, F.; Spanka, D.T.; Rick, T.; Monzon, V.; Förstner, K.U.; Klug, G. The small DUF1127 protein CcaF1 from *Rhodobacter sphaeroides* is an RNA-binding protein involved in sRNA maturation and RNA turnover. *Nucleic Acids Res.* **2021**, *49*, 3003–3019. [CrossRef] [PubMed] [CrossRef]
19. Eisenhardt, K.M.; Reuscher, C.M.; Klug, G. PcrX, an sRNA derived from the 3′- UTR of the *Rhodobacter sphaeroides puf* operon modulates expression of *puf* genes encoding proteins of the bacterial photosynthetic apparatus. *Mol. Microbiol.* **2018**, *110*, 325–334. [CrossRef] [CrossRef]
20. Weber, L.; Thoelken, C.; Volk, M.; Remes, B.; Lechner, M.; Klug, G. The Conserved Dcw Gene Cluster of *R. sphaeroides* Is Preceded by an Uncommonly Extended 5′ Leader Featuring the sRNA UpsM. *PLoS ONE* **2016**, *11*, e0165694. [CrossRef] [CrossRef]
21. Grützner, J.; Remes, B.; Eisenhardt, K.M.H.; Scheller, D.; Kretz, J.; Madhugiri, R.; McIntosh, M.; Klug, G. sRNA-mediated RNA processing regulates bacterial cell division. *Nucleic Acids Res.* **2021**, *49*, 7035–7052. [CrossRef] [PubMed] [CrossRef]
22. McDowall, K.; Lin-Chao, S.; Cohen, S. A+U content rather than a particular nucleotide order determines the specificity of RNase E cleavage. *J. Biol. Chem.* **1994**, *269*, 10790–10796. [CrossRef] [CrossRef]
23. Mackie, G.A. Ribonuclease E is a 5′-end-dependent endonuclease. *Nature* **1998**, *395*, 720–724. [CrossRef] [PubMed] [CrossRef] [PubMed]
24. Hoyos, M.; Huber, M.; Förstner, K.U.; Papenfort, K. Gene autoregulation by 3′ UTR-derived bacterial small RNAs. *eLife* **2020**, *9*, e58836. [CrossRef] [CrossRef]
25. Davis, B.M.; Waldor, M.K. RNase E-dependent processing stabilizes MicX, a *Vibrio cholerae* sRNA. *Mol. Microbiol.* **2007**, *65*, 373–385. [CrossRef] [CrossRef] [PubMed]

26. Chao, Y.; Vogel, J. A 3′ UTR-Derived Small RNA Provides the Regulatory Noncoding Arm of the Inner Membrane Stress Response. *Mol. Cell* **2016**, *61*, 352–363. [CrossRef] [CrossRef] [PubMed]
27. Nikolaev, N.; Schlessinger, D.; Wellauer, P.K. 30 S pre-ribosomal RNA of *Escherichia coli* and products of cleavage by ribonuclease III: Length and molecular weight. *J. Mol. Biol.* **1974**, *86*, 741–748. [CrossRef] [CrossRef]
28. Lim, B.; Sim, S.H.; Sim, M.; Kim, K.; Jeon, C.O.; Lee, Y.; Ha, N.C.; Lee, K. RNase III Controls the Degradation of corA mRNA in *Escherichia coli*. *J. Bacteriol.* **2012**, *194*, 2214–2220. [CrossRef] [PubMed] [CrossRef]
29. Gordon, G.C.; Cameron, J.C.; Pfleger, B.F. RNA Sequencing Identifies New RNase III Cleavage Sites in *Escherichia coli* and Reveals Increased Regulation of mRNA. *mBio* **2017**, *8*, e00128-17. [CrossRef] [CrossRef]
30. Kordes, E.; Jock, S.; Fritsch, J.; Bosch, F.; Klug, G. Cloning of a gene involved in rRNA precursor processing and 23S rRNA cleavage in *Rhodobacter capsulatus*. *J. Bacteriol.* **1994**, *176*, 1121–1127. [CrossRef] [PubMed] [CrossRef]
31. Rauhut, R.; Andreas, J.; Conrad, C.; Klug, G. Identification and Analysis of the rnc gene for RNase III in *Rhodobacter capsulatus*. *Nucleic Acids Res.* **1996**, *24*, 1246–1251. [CrossRef] [PubMed] [CrossRef]
32. Reuscher, C.M.; Klug, G. Antisense RNA asPcrL regulates expression of photosynthesis genes in *Rhodobacter sphaeroides* by promoting RNase III-dependent turn-over of *puf* mRNA. *RNA Biol.* **2021**, *18*, 1445–1457. [CrossRef] [CrossRef] [PubMed]
33. Donovan, W.P.; Kushner, S.R. Polynucleotide phosphorylase and ribonuclease II are required for cell viability and mRNA turnover in *Escherichia coli* K-12. *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 120–124. [CrossRef] [CrossRef]
34. Andrade, J.M.; Arraiano, C.M. PNPase is a key player in the regulation of small RNAs that control the expression of outer membrane proteins. *RNA* **2008**, *14*, 543–551. [CrossRef] [CrossRef]
35. Andrade, J.M.; Pobre, V.; Matos, A.M.; Arraiano, C.M. The crucial role of PNPase in the degradation of small RNAs that are not associated with Hfq. *RNA* **2012**, *18*, 844–855. [CrossRef] [CrossRef]
36. Broglia, L.; Lécrivain, A.L.; Renault, T.T.; Hahnke, K.; Ahmed-Begrich, R.; Rhun, A.L.; Charpentier, E. An RNA-seq based comparative approach reveals the transcriptome-wide interplay between 3′-to-5′ exoRNases and RNaseY. *Nat. Commun.* **2020**, *11*, 1587. [CrossRef] [PubMed] [CrossRef]
37. Spanka, D.T.; Reuscher, C.M.; Klug, G. Impact of PNPase on the transcriptome of *Rhodobacter sphaeroides* and its cooperation with RNase III and RNase E. *BMC Genom.* **2021**, *22*, 106. [CrossRef] [CrossRef]
38. Rische, T.; Klug, G. The ordered processing of intervening sequences in 23S rRNA of *Rhodobacter sphaeroides* requires RNase J. *RNA Biol.* **2012**, *9*, 343–350. [CrossRef] [PubMed] [CrossRef]
39. Rische-Grahl, T.; Weber, L.; Remes, B.; Förstner, K.U.; Klug, G. RNase J is required for processing of a small number of RNAs in *Rhodobacter sphaeroides*. *RNA Biol.* **2014**, *11*, 855–864. [CrossRef] [CrossRef]
40. Deana, A.; Celesnik, H.; Belasco, J.G. The bacterial enzyme RppH triggers messenger RNA degradation by 5′ pyrophosphate removal. *Nature* **2008**, *451*, 355–358. [CrossRef] [CrossRef]
41. Jacob, A.I.; Köhrer, C.; Davies, B.W.; RajBhandary, U.L.; Walker, G.C. Conserved Bacterial RNase YbeY Plays Key Roles in 70S Ribosome Quality Control and 16S rRNA Maturation. *Mol. Cell* **2013**, *49*, 427–438. [CrossRef] [CrossRef]
42. Møller, T.; Franch, T.; Højrup, P.; Keene, D.R.; Bächinger, H.P.; Brennan, R.G.; Valentin-Hansen, P. Hfq: a bacterial Sm-like protein that mediates RNA-RNA interaction. *Mol. Cell* **2002**, *9*, 23–30. [CrossRef] [CrossRef]
43. Zhang, A.; Wassarman, K.M.; Ortega, J.; Steven, A.C.; Storz, G. The Sm-like Hfq Protein Increases OxyS RNA Interaction with Target mRNAs. *Mol. Cell* **2002**, *9*, 11–22. [CrossRef] [CrossRef]
44. Berghoff, B.A.; Glaeser, J.; Sharma, C.M.; Zobawa, M.; Lottspeich, F.; Vogel, J.; Klug, G. Contribution of Hfq to photooxidative stress resistance and global regulation in *Rhodobacter sphaeroides*. *Mol. Microbiol.* **2011**, *80*, 1479–1495. [CrossRef] [CrossRef]
45. Glaeser, J.; Zobawa, M.; Lottspeich, F.; Klug, G. Protein Synthesis Patterns Reveal a Complex Regulatory Response to Singlet Oxygen in *Rhodobacter*. *J. Proteome Res.* **2007**, *6*, 2460–2471. [CrossRef] [CrossRef]
46. Berghoff, B.A.; Glaeser, J.; Sharma, C.M.; Vogel, J.; Klug, G. Photooxidative stress-induced and abundant small RNAs in *Rhodobacter sphaeroides*. *Mol. Microbiol.* **2009**, *74*, 1497–1512. [CrossRef] [PubMed] [CrossRef]
47. Quinlan, A.R. BEDTools: The Swiss-Army Tool for Genome Feature Analysis. *Curr. Protoc. Bioinform.* **2014**, *47*, 11.12.1–11.12.34. [CrossRef] [CrossRef] [PubMed]
48. Remes, B.; Rische-Grahl, T.; Müller, K.M.H.; Förstner, K.U.; Yu, S.H.; Weber, L.; Jäger, A.; Peuser, V.; Klug, G. An RpoHI-Dependent Response Promotes Outgrowth after Extended Stationary Phase in the Alphaproteobacterium *Rhodobacter sphaeroides*. *J. Bacteriol.* **2017**, *199*, e00249-17. [CrossRef] [PubMed] [CrossRef]
49. Nuss, A.M.; Glaeser, J.; Berghoff, B.A.; Klug, G. Overlapping Alternative Sigma Factor Regulons in the Response to Singlet Oxygen in *Rhodobacter sphaeroides*. *J. Bacteriol.* **2010**, *192*, 2613–2623. [CrossRef] [CrossRef]
50. Müller, K.M.H.; Berghoff, B.A.; Eisenhardt, B.D.; Remes, B.; Klug, G. Characteristics of Pos19—A Small Coding RNA in the Oxidative Stress Response of *Rhodobacter sphaeroides*. *PLoS ONE* **2016**, *11*, e0163425. [CrossRef] [PubMed] [CrossRef]
51. Adnan, F.; Weber, L.; Klug, G. The sRNA SorY confers resistance during photooxidative stress by affecting a metabolite transporter in*Rhodobacter sphaeroides*. *RNA Biol.* **2015**, *12*, 569–577. [CrossRef] [CrossRef]
52. Peuser, V.; Metz, S.; Klug, G. Response of the photosynthetic bacterium *Rhodobacter sphaeroides* to iron limitation and the role of a Fur orthologue in this response. *Environ. Microbiol. Rep.* **2011**, *3*, 397–404. [CrossRef] [CrossRef]
53. Wang, C.; Chao, Y.; Matera, G.; Gao, Q.; Vogel, J. The conserved 3′ UTR-derived small RNA NarS mediates mRNA crossregulation during nitrate respiration. *Nucleic Acids Res.* **2019**, *48*, 2126–2143. [CrossRef] [PubMed] [CrossRef]

54. Sy, B.M.; Lan, R.; Tree, J.J. Early termination of the Shiga toxin transcript generates a regulatory small RNA. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 25055–25065. [CrossRef] [CrossRef] [PubMed]
55. Thomason, M.K.; Voichek, M.; Dar, D.; Addis, V.; Fitzgerald, D.; Gottesman, S.; Sorek, R.; Greenberg, E.P. A rhlI 5′ UTR-Derived sRNA Regulates RhlR-Dependent Quorum Sensing in *Pseudomonas aeruginosa*. *mBio* **2019**, *10*, e02253-19. [CrossRef] [CrossRef] [PubMed]
56. Iosub, I.A.; Marchioretto, M.; van Nues, R.W.; McKellar, S.; Viero, G.; Granneman, S. The mRNA derived MalH sRNA contributes to alternative carbon source utilization by tuning maltoporin expression in *E. coli*. *RNA Biol.* **2020**, *18*, 914–931. [CrossRef] [PubMed] [CrossRef] [PubMed]
57. van der Meulen, S.B.; Hesseling-Meinders, A.; de Jong, A.; Kok, J. The protein regulator ArgR and the sRNA derived from the 3′-UTR region of its gene, ArgX, both regulate the arginine deiminase pathway in *Lactococcus lactis*. *PLoS ONE* **2019**, *14*, e0218508. [CrossRef] [CrossRef] [PubMed]
58. Zhao, Z.; Peng, T.; Oh, J.I.; Glaeser, J.; Weber, L.; Li, Q.; Klug, G. A response regulator of the OmpR family is part of the regulatory network controlling the oxidative stress response of *Rhodobacter sphaeroides*. *Environ. Microbiol. Rep.* **2018**, *11*, 118–128. [CrossRef] [CrossRef] [PubMed]
59. Klug, G.; Adams, C.W.; Belasco, J.; Doerge, B.; Cohen, S.N. Biological consequences of segmental alterations in mRNA stability: effects of deletion of the intercistronic hairpin loop region of the *Rhodobacter capsulatus puf* operon. *EMBO J.* **1987**, *6*, 3515–3520. [CrossRef] [PubMed] [CrossRef]
60. McCarthy, J.; Schairer, H.; Sebald, W. Translational initiation frequency of *atp* genes from *Escherichia coli:* identification of an intercistronic sequence that enhances translation. *EMBO J.* **1985**, *4*, 519–526. [CrossRef] [CrossRef]
61. Queen, C.; Rosenberg, M. Differential translation efficiency explains discoordinate expression of the galactose operon. *Cell* **1981**, *25*, 241–249. [CrossRef] [CrossRef]
62. Régnier, P.; Portier, C. Initiation, attenuation and RNase III processing of transcripts from the *Escherichia coli* operon encoding ribosomal protein S15 and polynucleotide phosphorylase. *J. Mol. Biol.* **1986**, *187*, 23–32. [CrossRef] [CrossRef]
63. Bathke, J.; Konzer, A.; Remes, B.; McIntosh, M.; Klug, G. Comparative analyses of the variation of the transcriptome and proteome of *Rhodobacter sphaeroides* throughout growth. *BMC Genom.* **2019**, *20*, 358. [CrossRef] [PubMed] [CrossRef] [PubMed]
64. Deutscher, M.P. Regulation of Bacterial Ribonucleases. *Annu. Rev. Microbiol.* **2021**, *75*, 71–86. [CrossRef] [PubMed] [CrossRef] [PubMed]
65. Chen, C.; Deutscher, M.P. Elevation of RNase R in Response to Multiple Stress Conditions. *J. Biol. Chem.* **2005**, *280*, 34393–34396. [CrossRef] [PubMed] [CrossRef]
66. Liang, W.; Deutscher, M.P. Post-translational modification of RNase R is regulated by stress-dependent reduction in the acetylating enzyme Pka (YfiQ). *RNA* **2011**, *18*, 37–41. [CrossRef] [CrossRef] [PubMed]
67. Hördt, A.; López, M.G.; Meier-Kolthoff, J.P.; Schleuning, M.; Weinhold, L.M.; Tindall, B.J.; Gronow, S.; Kyrpides, N.C.; Woyke, T.; Göker, M. Analysis of 1000+ Type-Strain Genomes Substantially Improves Taxonomic Classification of Alphaproteobacteria. *Front. Microbiol.* **2020**, *11*, 468. [CrossRef] [CrossRef]
68. Remes, B.; Berghoff, B.A.; Förstner, K.U.; Klug, G. Role of oxygen and the OxyR protein in the response to iron limitation in *Rhodobacter sphaeroides*. *BMC Genom.* **2014**, *15*, 794. [CrossRef] [CrossRef]
69. Glaeser, J.; Klug, G. Photo-oxidative stress in *Rhodobacter sphaeroides*: protective role of carotenoids and expression of selected genes. *Microbiology* **2005**, *151*, 1927–1938. [CrossRef] [CrossRef]
70. van Niel, C.B. The Culture, General Physiology, Morphology, and Classification of the Non-Sulfur Purple and Brown Bacteria. *Microbiol. Mol. Biol. Rev.* **1944**, *8*, 1–118. [CrossRef] [CrossRef]
71. Charoenpanich, P.; Meyer, S.; Becker, A.; McIntosh, M. Temporal Expression Program of Quorum Sensing-Based Transcription Regulation in *Sinorhizobium meliloti*. *J. Bacteriol.* **2013**, *195*, 3224–3236. [CrossRef] [PubMed] [CrossRef] [PubMed]
72. McIntosh, M.; Eisenhardt, K.; Remes, B.; Konzer, A.; Klug, G. Adaptation of the Alphaproteobacterium *Rhodobacter sphaeroides* to stationary phase. *Environ. Microbiol.* **2019**, *21*, 4425–4445. [CrossRef] [PubMed] [CrossRef] [PubMed]
73. Janzon, L.; Loefdahl, S.; Arvidson, S. Evidence for a coordinate transcriptional control of alpha-toxin and protein a synthesis in *Staphylococcus aureus*. *FEMS Microbiol. Lett.* **1986**, *33*, 193–198. [CrossRef] [CrossRef]
74. Kingsford, C.L.; Ayanbule, K.; Salzberg, S.L. Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. *Genome Biol.* **2007**, *8*, R22. [CrossRef] [PubMed] [CrossRef]

Chapter 4

The small DUF1127 protein CcaF1 from *Rhodobacter sphaeroides* is an RNA-binding protein involved in sRNA maturation and RNA turnover

Published online 12 March 2021

Nucleic Acids Research, 2021, Vol. 49, No. 6 ***3003–3019***
doi: 10.1093/nar/gkab146

NAR Breakthrough Article

The small DUF1127 protein CcaF1 from *Rhodobacter sphaeroides* is an RNA-binding protein involved in sRNA maturation and RNA turnover

Julian Grützner[1], Fabian Billenkamp[1,2], Daniel-Timon Spanka[1], Tim Rick[1], Vivian Monzon[3], Konrad U. Förstner[3,4] and Gabriele Klug[1,*]

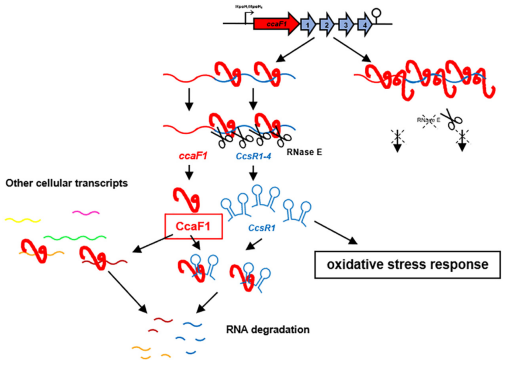
[1]Institute of Microbiology and Molecular Biology, Justus Liebig University Giessen, IFZ, Heinrich-Buff-Ring 26–32, D-35292 Giessen, Germany, [2]Institute of Animal Nutrition, Friedrich Loeffler Institute, Bundesalle 37, D-38116 Braunschweig, Germany, [3]ZB MED-Information Center of Life Science, Germany and [4]Institute of Information Science, TH Köln, University of Applied Science, Gustav-Heinemann-Ufer 54, D-50968 Köln, Cologne, Germany

Received June 17, 2020; Revised February 10, 2021; Editorial Decision February 19, 2021; Accepted February 26, 2021

ABSTRACT

Many different protein domains are conserved among numerous species, but their function remains obscure. Proteins with DUF1127 domains number >17 000 in current databases, but a biological function has not yet been assigned to any of them. They are mostly found in alpha- and gammaproteobacteria, some of them plant and animal pathogens, symbionts or species used in industrial applications. Bioinformatic analyses revealed similarity of the DUF1127 domain of bacterial proteins to the RNA binding domain of eukaryotic Smaug proteins that are involved in RNA turnover and have a role in development from *Drosophila* to mammals. This study demonstrates that the 71 amino acid DUF1127 protein CcaF1 from the alphaproteobacterium *Rhodobacter sphaeroides* participates in maturation of the CcsR sRNAs that are processed from the 3′ UTR of the *ccaF* mRNA and have a role in the oxidative stress defense. CcaF1 binds to many cellular RNAs of different type, several mRNAs with a function in cysteine / methionine / sulfur metabolism. It affects the stability of the CcsR RNAs and other non-coding RNAs and mRNAs. Thus, the widely distributed DUF1127 domain can mediate RNA-binding, affect stability of its binding partners and consequently modulate the bacterial transcriptome, thereby influencing different physiological processes.

GRAPHICAL ABSTRACT



INTRODUCTION

Bacterial genomes typically harbor many small open reading frames that have not been annotated in the past due to their small size. More recently it has emerged that small proteins participate in a multitude of cellular processes (1,2). Although only a minor fraction of the small proteins could be analyzed up to now, they exhibit a great diversity in their mechanisms of action and their physiological functions. Important roles for small proteins in e.g. cell division, transport, spore formation and signal transduction have been unraveled (2).

This study uncovers a new function for a small protein from *Rhodobacter sphaeroides* and, to the best of our knowl-

*To whom correspondence should be addressed. Tel: +49 641 99 355 42; Fax: +49 641 99 355 49; Email: gabriele.klug@mikro.bio.uni-giessen.de
Present address: Vivian Monzon, European Molecular Biology Laboratory, European Bioinformatics Institute, CB 10, 1SD Hinxton, UK.

edge, represents the first functional assignment to a member of the DUF1127 proteins. *Rhodobacter sphaeroides* is a facultative phototrophic alphaproteobacterium living in fresh and brackish water habitats. If sufficient oxygen is available, it can perform aerobic respiration. When oxygen becomes limiting or under anaerobic conditions, ATP is produced by anoxygenic photosynthesis, anaerobic respiration or fermentation. Since the simultaneous presence of (bacterio-) chlorophylls, light and oxygen leads to the production of the harmful singlet oxygen, the formation of photosynthetic complexes is tightly controlled by redox and light signals (3–6). Furthermore, *R. sphaeroides* has developed a complex regulatory network consisting of proteins and sRNAs to defend against singlet oxygen stress (7–11). To date several different sRNAs that are induced by oxidative stress and have roles in oxidative/singlet oxygen stress response have been investigated in *R. sphaeroides* (12–17). By interacting to the mRNA for the transcriptional regulator FlhR, the four homologous sRNAs CcsR1–4 modulate the C1 metabolism and the pyruvate dehydrogenase complex in response to various stresses (18). As a consequence, the pool of the reductant glutathione is increased and aerobic electron transport, a main source of reactive oxygen species (ROS), is reduced. The CcsR1–4 sRNAs are derived from the 3′ UTR of the RSP_6037 mRNA (Figure 1A). Transcription of the RSP_6037-CcsR genes is initiated at a RpoHI/RpoHII-dependent promoter (18).

The two alternative sigma factors RpoHI and RpoHII are known to activate many genes in *R. sphaeroides* under a variety of stress conditions (8,19,20) and are also important for outgrowth after long stationary phase (21). Each CcsR RNA harbors two hairpin–loop structures and each loop contains a CCUCCUCCC anti-Shine Dalgarno sequence (7) that prompted Reinkensmeier and Giegerich (22) to name them 'Cuckoo' RNAs. RSP_6037 encodes a small protein of unknown function of 71 amino acids. Amino acids 23–62 constitute a DUF1127 domain (18). More than 17,000 bacterial protein sequences with DUF1127 domains in about 4000 bacterial species are listed in InterPro and the number is steadily increasing (23). The DUF1127 domain consists of 45–50 amino acids and often covers almost the entire protein. Alternatively, the DUF1127 domain is located at the C-terminus of slightly larger proteins with 60–75 amino acids. DUF1127 proteins are widely distributed among Alpha- and Gammaproteobacteria and mostly found in the orders Rhizobiales, Rhodobacterales, Enterobacteriales and Pseudomonales. Our previous work demonstrated that RSP_6037 influences the amounts of the CcsR RNAs, that are produced from the 3′ UTR of the RSP_6037 transcript (18).

According to Reinkensmeier and Giegerich (22) adjacency to ORFs with DUF1127 domains defines an orthologous subgroup of Cuckoo RNAs that is characterized by this genomic context. This subgroup, labeled CIN1 (conserved intergenic neighborhood 1), is present only in the *Rhodobacteraceae*, *Brucellaceae*, *Rhizobiaceae* and *Phylobacteraceae (22)*. Barnett *et al.* (24) found that the RSP_6037 ortholog SMc02051 (47 aa) from *Sinorhizobium meliloti* 1021 has a RpoHI and RpoHII responsive expression as described for the CcsR locus in *Rhodobacter sphaeroides* (Figure 1B). Interestingly, a similar responsiveness was observed for an adjacent DUF1127-containing ORF (SMc02052). SmMc02051 also has adjacent Cuckoo RNAs with a predicted σ^{70}-dependent promoter and is classified as a CIN1 member (25).

Here, we present major advancements in the investigation of RSP_6037 as a model for the DUF1127 domain containing ORFs that are adjacent to Cuckoo RNAs and demonstrate that this DUF1127-containing protein binds RNAs and can affect RNA stability.

MATERIALS AND METHODS

Bacterial strains and growth conditions

Rhodobacter sphaeroides strains (listed in S1 Table) were cultivated in a malate minimal-salt medium or on solid medium containing 1.6% (w/v) agar at 32°C in the dark (26). For microaerobic growth conditions (25–30 μM of dissolved oxygen) Erlenmeyer flasks filled up to 80% of the maximum volume were shaken at 140 rpm. When necessary, tetracycline (2 μg ml^{-1}) or spectinomycin (10 μg ml^{-1}) was added to liquid and solid growth media. Stress conditions were generated by a final concentration of 300 μM tBOOH, 1 mM H_2O_2 or 250 μM paraquat (O^{2-}) or by temperature shift to 42°C under microaerobic conditions.

To culture *Escherichia coli* strains (listed in S1 Table), cells were continuously shaken at 180 rpm in Luria–Bertani (LB) medium at 37°C or grown on solid growth media containing 1.6% (w/v) agar. When necessary, tetracycline (20 μg ml^{-1}) or spectinomycin (10 μg ml^{-1}) was added to the media.

S. meliloti 1021 was cultivated similar to *R. sphaeroides* strains using TY medium (27).

Construction of overexpression plasmids

For construction of a plasmid for constitutive overexpression of the small DUF1127 protein CcaF1 (RSP_6037) from *R. sphaeroides* 2.4.1 without the sRNA cluster CcsR1–4 but with the terminator structure of this gene-locus at the 3′ end, a 310 bp fragment of the *ccaF1* (RSP_6037) gene locus (primers: CcaF1_f and CcaF1int_r; Supplementary Table S2) and a 135 bp fragment of the terminator sequence (primers: CcaF1int_f and CcaF1_r; Tabel S2) were amplified by PCR using chromosomal *R. sphaeroides* 2.4.1 DNA as template. Both fragments have an overlapping region. By a second PCR step both fragments were fused resulting in a 445 bp fragment of the *ccaF1* (RSP_6037) gene sequence and the terminator sequence. The corresponding fragment was sub-cloned into the BamHI and EcoRI sites of the pJET1.2 cloning vector (Thermo Fisher Scientific, Waltham, MA, USA) and, after digestion with the corresponding restriction enzymes, ligated into the expression vector pRK4352 (28).

The constitutive overexpression of the *ccaF1* (RSKD131_0402) gene from *R. sphaeroides* KD131 (Primers are listed in Supplementary Table S2) was performed in the same way like for *ccaF1* (RSP_6037) from *R. sphaeroides* 2.4.1, but using chromosomal *R. sphaeroides* KD131 DNA as template. The constitutive overexpression of the sRNA cluster CcsR1–4 and the whole gene locus from *R. sphaeroides* KD131 (primers are listed in Supplementary Table S2) was performed in the

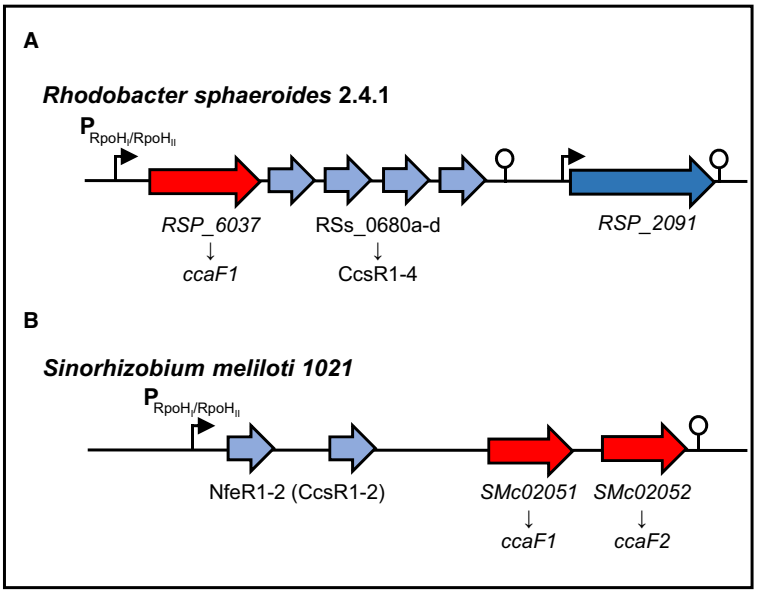


Figure 1. Schematic representations of the CcsR-RNA/DUF1127-protein loci in *R. sphaeroides* 2.4.1 and *S. meliloti* 1021. (**A**) Genomic context of the DUF1127 protein RSP_6037/CcaF1 (red) and the CcsR1–4 sRNAs (light blue) from *R. sphaeroides* 2.4.1. The protein–sRNA operon is preceded by a RpoHI/RpoHII promoter (black arrow) and a Rho-independent terminator structure is located at the 3′ end (modified from Billenkamp *et al.* (18)). (**B**) Corresponding locus of CcsR(NfeR)-RNAs and DUF1127 proteins in *S. meliloti* 1021. Open reading frames of the DUF1127 proteins are colored red, while sRNAs are colored in light blue. An RpoHI/RpoHII dependent promoter and a terminator are indicated by an arrow and a hairpin structure.

same way as described in (18), but using chromosomal *R. sphaeroides* KD131 DNA as template. The resulting overexpression plasmids were conjugated from *E. coli* S17-1 to *R. sphaeroides* 2.4.1 (29).

Construction of 3xFLAG-tagged CcaF1 and RSP_0557 expression strains

The *R. sphaeroides* RSP_6037 and RSP_0557 loci were amplified by PCR of pRK6037 and pBBR0557 (17) plasmid DNA using primers CcaF1FLAG_NT_f and CcaF1FLAG_NT_r or RSP0557FLAG_NT_f and RSP0557FLAG_NT_r (primers are listed in Supplementary Table S2). The amplified fragments start with an ATG start-codon at the 5′ end followed by the 3xFLAG sequence at the N-terminus and the *ccaF1* (RSP_6037) or RSP_0557 gene. The fragment was sub-cloned into the BamHI and EcoRI sites of the pJET1.2 cloning vector (Thermo Fisher Scientific, Waltham, MA, USA) and ligated into the pRK4352 overexpression vector (28) after digestion with suitable restriction enzymes. The resulting overexpression plasmids pRKCcaF1FLAG_NT and pRK0557FLAG_NT were transferred from *E. coli* S17-1 to *R. sphaeroides* 2.4.1 by biparental conjugation (29).

Construction of a His_6-MBP-TEV-CcaF1 overexpression plasmid

The Gibson assembly method was used to construct the plasmids for heterologous protein expression. The used oligonucleotides (primers are listed in Supplementary Table S2) provided overlaps of at least 20 nucleotides. Purified PCR fragments were used in equimolar concentrations and the reaction was incubated for 60 minutes at 50°C. Subsequently the reaction mixture was cooled to 4°C and transformed into chemically competent cells of strain *E. coli* DH5α λpir (30).

Zone of inhibition, survival assay and spot assay

Zone of inhibition assay was performed as described in Li *et al.* (31). The 5 mm filter-paper disks contain 5 μl of oxidative agent (200 μM paraquat or 700 μM tBOOH). The plates were incubated for 48 h at 32°C in the dark and the diameter of the zone of inhibition indicates the sensitivity of the cells against the agent.

For determination of survival rates, cultures were grown under microaerobic conditions. 300 μM tBOOH were added and after 30 min, 60 min or 90 min dilutions were plated on solid malate minimal-salt medium. The plates were incubated for 48 h at 32°C in the dark. The number of colonies of a control culture grown without the addition of any oxidative stress agents was defined as 100% survival.

To test the growth behavior of the different strains in presence of various stresses we performed a spot assay. Cultures were grown at 32°C in microaerobic condition. At an OD_{660} of 0.5 10 μl of different dilutions (10^0–10^{-5}) were spotted on an agar plate containing the stress agents (10 μM $CdCl_2$, 250 mM NaCl, 100 μM tBOOH). The plates were incubated for 48 h at 32°C in the dark. For heat shock, plates were in-

cubated at 42°C over-night and afterwards at 32°C. The intensity of the spots was quantified by the 1D-Quantity One software (Bio-Rad). The control spot (10^0) was defined as 100% survival.

Determination of RNA half-life

The *R. sphaeroides* cultures of interest were incubated under the desired growth condition (see bacterial growth conditions) to an OD_{660} 0.5. After taking sample t_0, rifampicin was added to a final concentration of 0.2 mg/ml. The cells were harvested by centrifugation (10 000 rpm, 10 min, 4°C) at defined time points and RNA was isolated.

RNA isolation

For RNA isolation *R. sphaeroides* cultures were grown to an $OD_{660\ nm}$ 0.5 under the different growth conditions (see bacterial growth conditions). Cells were harvested by centrifugation at 10 000 rpm for 10 min at 4°C. RNA was isolated for Northern Blot analysis, RT-PCR and RNA sequencing analysis using the hot phenol method (32) and precipitated with 1/10× vol. 3 M sodium acetate pH 4.5 and 2.5× vol. 96% ethanol. For RNA sequencing the remaining DNA was removed by TURBO-DNase treatment (Invitrogen).

Northern blot

For detection of small RNAs 7.5 μg total RNA were separated on a 10% polyacrylamide gel containing 7 M urea. Afterwards RNA was transferred to Nylon membranes (Roth) by semi-dry electroblotting. Oligodeoxynucleotides (listed in S2 Table) for detection were labeled with [γ-^{32}P]-ATP (Hartmann Analytic) by T4 polynucleotide kinase (Fermentas; #EK0031) and were hybridized overnight. Membranes were exposed on phosphoimaging screens (Bio-Rad) and analyzed by the 1D-Quantity One software (Bio-Rad). For determination of mRNAs and precursor transcripts 10 μg total RNA were separated on a 1% (w/v) agarose 2.2 M formaldehyde gel and transferred to nylon membrane by vacuum pressure blotting. DNA fragments of the mRNAs and precursor transcripts were labeled with [α-^{32}P]-dCTP (Hartmann Analytic) using nick translation (nick translation kit; Amersham Pharmacia Biotech). Membranes were hybridized overnight, exposed on phosphoimaging screens (Bio-Rad) and analyzed by the 1D-Quantity One software (Bio-Rad).

Co-immunoprecipitation

For co-immunoprecipitation *R. sphaeroides* pRKCcaF1FLAG_NT and *R. sphaeroides* pRKCcaF1 or *R. sphaeroides* pRK0557FLAG_NT and *R. sphaeroides* pRK0557 were grown under microaerobic conditions and harvested in exponential growth phase (culture volume of 400 ml) by centrifugation at 10 000 rpm at 4°C. Pellets were resuspended in 2 ml of cold lysis buffer (20 mM Tris pH 7.5, 150 mM KCl, 1 mM $MgCl_2$, 1 mM DTT) and disrupted by sonication (33). Cell lysate was centrifuged for 10 min at 13 000 rpm and 4°C, followed by an ultracentrifugation step (100 000 rpm, 1 h, 4°C). Afterwards the supernatant was mixed with 40 μl of ANTI-FLAG M2 Magnetic Beads (Sigma-Aldrich) and incubated for 2 h, at 4°C under rotation. Following five washing steps with 500 μl of lysis buffer, magnetic beads were resuspended in 500 μl of lysis buffer and RNA was isolated with phenol and chloroform–isoamyl alcohol followed by precipitation with 1/10× vol. 3 M sodium acetate pH 4.5 and 2.5× vol. 96% ethanol overnight. The precipitated CoIP RNA was treated by DNase I (Invitrogen) to remove any DNA contaminations. The isolated RNA was analyzed by RNA sequencing and RT-PCR.

Reverse transcription (RT) PCR

CoIP RNA was analyzed after DNase-treatment by a reverse transcription (RT) PCR using the One-Step Brilliant III QRT-PCR Master Mix Kit (Agilent). Each 10 μl reaction mixture contained 5 μl Master Mix (supplied), 0.1 μl DTT (100 mM, supplied), 0.5 μl Ribo-Block solution (supplied), 0.4 μl water, 1 μl of each primer (10 pmol/μl) listed in Supplementary Table S2, and 2 μl RNA (20 ng/μl). The reactions were performed in a spectrofluorometric thermal cycler (BioRad) and analyzed by BioRad CFXManager 3.0. Afterwards the RT-PCR products were separated on a 10% polyacrylamide gel and analyzed by ethidium-bromide staining.

Protein production and purification

Strain *E. coli* BL21 DE3 (New England Biolabs, Germany) carrying a pET24c plasmid was used for recombinant protein expression. Cells were grown in LB medium supplemented with kanamycin (50 mg/l) at 37°C under vigorous shaking until an OD_{600} of 0.75 was reached. Subsequently the culture was cooled for 10 min in an ice bath before protein expression was induced by addition of D-(+)-lactose-monohydrate (12.5 g/l). The culture was then incubated at 16°C under vigorous shaking for 16 h. Cells were harvested by centrifugation (5000 rpm, 10 min, 4°C), the resulting cell pellet flash frozen in liquid nitrogen and stored at –20°C until use.

For protein purification cells were resuspended in lysis buffer (50 mM Tris–HCl, 500 mM NaCl, 50 mM KCl, 10 mM $MgCl_2 \cdot 6H_2O$, 20 mM imidazole, 1 mM DTT, 0.02% Tween20, pH 8) and lysed by sonication (Bandelin Sonoplus). Cell debris and intact cells were removed by centrifugation (20 000 rpm, 30 min, 4°C) and filtration. The obtained lysate was then loaded onto a 5 ml HisTrap HP (GE healthcare) column and equilibrated with lysis buffer using the ÄKTA PURE25 system. The column was then washed with 10 CV (column volume) lysis buffer. A linear gradient of 3 CV elution buffer (50 mM Tris–HCl, 500 mM NaCl, 50 mM KCl, 10 mM $MgCl_2 \cdot 6H_2O$, 600 mM imidazole, 1 mM DTT, 0.02% Tween20, pH 8) from 10 to 100% was used to elute the protein, followed by 2 CV elution buffer. Elution fractions were analyzed by 12% SDS-PAGE. Samples containing the protein of interest were then combined and centrifuged (13 000 rpm, 1 min, 4°C) prior to loading 500 μl onto a Superdex 200 increase 10/300 GL SEC column equilibrated with SEC buffer (50 mM TrislHCl, 500 mM NaCl, 50 mM KCl, 10 mM $MgCl_2 \cdot 6H_2O$, 1 mM DTT, pH 8) and

connected to the ÄKTA PURE25 system. After isocratic elution with SEC buffer (0.5 ml/min) fractions containing the protein of interest were identified by 12% SDS PAGE.

100 U (w/w) of TEV protease (New England Biolabs, Germany) were applied over night at 4°C for cleavage of the fusion protein. The His_6-TEV protease and the His_6-MBP-tag were removed by nickel NTA agarose beads (Qiagen). Fractions containing the protein of interest were identified by 12% SDS PAGE.

Electrophoretic mobility shift assay (EMSA)

RNA was transcribed *in vitro* using T7 polymerase (NEB) and PCR products as template, which contain the T7 promoter region at the 5′ ends. 150 fmol of the radio-labelled RNA was denatured separately for 1 min at 95°C, cooled down for 2 min on ice and renatured for 5 min at 32°C. After these de- and renaturing steps, 5x structure buffer (25 mM $MgCl_2$ and 300 mM KCl) and the purified protein CcaF1 in different molar ratios were added in a final volume of 10 μl. For formation of the RNA–protein complex, the samples were incubated for 30 min at 32°C. Afterwards, the reactions were mixed with 3 μl of loading dye (50% glycerol, 0.5× TBE, 0.2% bromophenol blue) and loaded onto a 6% non-denaturing polyacrylamide gel containing 0.5× TBE. Gels were pre-run in 0.5× TBE running buffer at 100 V for 60 min at room temperature before loading. Electrophoresis was performed at room temperature by applying 200 V for 4 h. Gels were dried, exposed on phosphoimaging screens (Bio-Rad) and analyzed by the 1D-Quantity One software (Bio-Rad).

Library construction, RNA sequencing and data analysis

RNA sequencing data are based on triplicates and the RNA for each triplicate stemmed from three independent cultures. After harvesting the respective aliquots total RNA was extracted followed by DNase treatment. RNA quality was checked using a 2100 Bioanalyzer with the RNA 6000 Nano kit (Agilent Technologies). The RNA integrity number (RIN) for all samples was between 2.2 and 5.1. 300 ng of total RNA were used for the preparation of a cDNA library with the NEBNext Multiplex Small RNA Library Prep kit for Illumina (NEB) in accordance with the manufacturers' instructions with modifications: RNA was dephosphorylated at the 3′ end, phosphorylated at the 5′ end and decapped using 10 U T4-PNK ± 40 nmol ATP and 5 U RNA 5′ pyrophosphohydrolase (RppH), respectively (NEB). After each enzymatic treatment RNA was purified with the Zymo RNA Clean & Concentrator kit. The RNA fragments were ligated for cDNA synthesis to 3′ SR adapter and 5′ SR adapter diluted 1:3 with nuclease-free water before use. PCR amplification to add Illumina adaptors and indices to the cDNA was performed for 14 cycles with 1:3 diluted primer. Barcoded DNA Libraries were purified using magnetic MagSi-NGS^{PREP} Plus beads (AMSBIO) at a 1.8 ratio of beads to sample volume. Libraries were quantified with the Qubit 3.0 fluorometer (ThermoFisher) and the library quality and size distribution were checked using a 2100 Bioanalyzer with the DNA-1000 kit (Agilent). Sequencing of pooled libraries, spiked with 10% PhiX control library, was performed in single-end mode on the NextSeq 500 platform (Illumina) with the High Output Kit v2.5 (75 Cycles). Demultiplexed FASTQ files were generated with bcl2fastq2 v2.20.0.422 (Illumina). The sequencing data are available at NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo) under the accession number GSE144523 and GSE145045. The adapter sequences were removed from the sequence reads in Fastq format. Read processing, generation of statistics, gene-wise read counting, coverage calculations and normalization were performed using READemption version 0.4.3 (34) using segemehl version 0.2.0 (35,36) for read alignments. Gene expression analysis was computed via DESeq version 1.22.1 (37). Downstream processing and statistical analysis were performed using the statistical language R (http://www.r-project.org).

Phylogenetic tree of conserved CcaF1 regions

Ninety-five CcaF1 amino acid sequences based on the CIN1 loci from Reinkensmeier and Giegerich 2015 (22) were collected from the NCBI database and aligned based on the DUF1127 domain in MEGA X (38). A phylogenetic tree was generated using the UPMGA method (39). A bootstrap consensus tree was derived from 500 replicates (40). Branches corresponding to partitions reproduced in <50% bootstrap replicates were collapsed. All positions with <95% site coverage were eliminated, i.e., fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position (partial deletion option). There was a total of 46 positions in the final dataset consisting of the DUF1127 domain and flanking amino acids, while further extensions were removed from analysis.

RESULTS

Comparison of different CcsR loci in Alphaproteobacteria

The genomic context of Cuckoo RNAs associated with a DUF1127-containing orthologous gene was labeled CIN1 (22). CIN1 includes the RNA family RSs0680 (now CcsR) and several members of the Rfam RNA family ar14. ar14 from *Sinorhizobium meliloti* 1021 includes the DUF1127-containing ORF SMc02051. The Cuckoo RNAs from this locus have been named NfeR (nodule formation efficiency RNA) and are expressed in root nodules and under salt stress, possibly from an RpoHI/HII-dependent promoter (25). Our data demonstrate that the CcsR RNAs from this CIN1 locus of *S. meliloti* 1021 are also induced by heat stress, but not by oxidative stress. The same result was obtained for CcsR1 from *S. fredii* HH103, while CcsR1 from *R. capsulatus* SB1003 was induced by superoxide and heat (Supplementary Figure S1). Therefore, the name CcsR (conserved CCUCCUCCC ('cuckoo')-motif stress-induced RNA) will be generally used for Cuckoo RNAs in CIN1 loci throughout the manuscript following the name of their first characterized example (18).

In *R. sphaeroides* 2.4.1 the CcsR RNAs are derived from the 3′ UTR of the RSP_6037 mRNA by RNase E-dependent processing (41). In contrast, the CcsR RNAs in

S. meliloti 1021 appear to be derived from the 5′ UTR of SMc02051 (Figure 1B). This leads to two common CIN1-locus orientations that are partially exclusive to taxonomic groups (Supplementary Figure S2).

A phylogenetic tree of the conserved regions of 95 selected amino acid sequences of DUF1127-proteins from CIN1 loci correlates with a phylogenetic tree based on 16S rDNA and with the family level assignments of the organisms harboring the CcsR RNA loci (Supplementary Figure S2). The typical *R. sphaeroides* locus with multiple CcsR RNAs derived from the 3′ UTR of the DUF1127-containing ORF is predominant in the *Rhodobacteraceae*, *Phylobacteraceae* and most *Rhizobiaceae*, while the *Brucellaceae* typically harbor two distinct CIN1 loci, each with only one CcsR RNA. The amino acid sequences of the associated DUF1127-proteins form separate clusters in this case (Supplementary Figure S2).

In the CIN1 loci of *Sinorhizobium*, one DUF1127-coding ORF is directly adjacent to the CcsR-RNAs and clusters together with the DUF1127-coding ORFs of the CIN1 loci in other Alphaproteobacteria. This ORF is followed by a second DUF1127-coding ORF, which might be an extension to the CIN1 locus based on expression in *S. meliloti* 1021. However, the representatives of this ORF form a separate cluster in the phylogenetic tree that is distinct from the other CIN1 related DUF1127-proteins (Supplementary Figure S2).

Since the DUF1127-containing ORFs form signature ORFs for the definition of CIN1 loci, we named RSP_6037 and its orthologs 'CcaF1' (conserved CcsR associated factor).

In silico characterization of the CcaF proteins

The 71 aa CcaF1 protein from *R. sphaeroides* 2.4.1 comprises an arginine-rich DUF1127 domain in its C-terminal part. The PHYRE2 webserver predicts a structure for this CcaF1 protein, which strongly resembles (71% confidence) the RNA-binding domain of the Smaug protein of *Drosophila melanogaster* (Supplementary Figure S3). The Smaug protein represses translation and induces mRNA decay in *Drosophila* embryos (42,43). PHYRE2 also detected structural homology to SAM (sterile alpha motif) pointed domain containing proteins with up to 68% confidence. SAM pointed domains are involved in protein-protein interaction and occur in eukaryotic and some bacterial proteins (www.ebi.ac.uk/interpro/entry/ IPR013761). Interestingly, the RNA binding domain of Smaug also comprises a SAM sub-domain, which was shown to interact with RNA (44). This was the first report that SAM domains can also interact with RNA. Some residues of this small domain are conserved between the Smaug domain and CcaF1, while other residues are only conserved among the bacteria (Supplementary Figure S3). CcaF1 from *R. sphaeroides* was shown to influence CcsR levels (18). Hence, CcaF1 from *R. sphaeroides* 2.4.1 can serve as a hub for a more detailed analysis of the CcaF protein group. Due to the small sizes of CcaF proteins, characterization of this protein group also goes along with functional characterization of the DUF1127 domain.

The CcaF1 protein affects CcsR levels *in trans* and alters stress resistance

We did not achieve deletion of the chromosomal *ccaF1-ccsR* locus indicating that these genes are essential (18). Therefore, we chose an overexpression strategy to study the role of CcaF1. Figure 2A shows the complete amino acid sequences of CcaF1 orthologs from the *R. sphaeroides* strains 2.4.1 (71 aa) and KD131 (50 aa). A schematic overview of the plasmid constructs we used to study the effect of CcaF1 (RSP_6037) on the CcsR levels is shown in Figure 2B. All combinations of *ccaF1* and *ccsR* genes were cloned under the control of the strong 16S promoter on a plasmid. Plasmid pRK4352 with 16S promoter but no cloned genes from *R. sphaeroides* served as control (EVC). As demonstrated previously (18), expression of the CcsR RNAs from the 16S promoter leads to a strong increase in CcsR levels, while there was no visible increase when the *ccaF*1 gene was co-expressed together with the *ccsR* genes (Figure 2C, left panel, lanes 2–4). When only the *ccaF1* gene was present on the plasmid, the level of CcsR RNAs expressed from the chromosome, was clearly decreased (Figure 2C, left panel, lanes 2 and 5). This demonstrates that CcaF1 can also act on CcsR levels *in trans*. To verify that the effect on the CcsR level is mediated by the CcaF1 small protein and not by the *ccaF1* mRNA, we exchanged the ATG start codon of the *ccaF1* gene to TGA. No other in frame ATG is present in the *ccaF1* gene. Expression of this mutated gene together with the CcsR RNAs from the plasmid had the same effect on CcsR levels as overexpression of the CcsR RNAs alone (Figure 2C, left panel, lanes 3 and 6) strongly supporting the assumption that CcsR levels are affected by the CcaF1 protein, not the *ccaF1* mRNA. The effects of the different plasmids were identical for all individual CcsR RNAs and were also observed for the CcsR1–4 precursor. We also performed a Northern blot with a *ccaF1* (RSP_6037) specific probe to confirm higher *ccaF1* transcript levels in presence of plasmid pRKCcaF1(Supplementary Figure S4). This Northern also shows increased levels of the *ccaF1*-CcsR1–4 precursor transcripts in strains expressing *ccaF1* together with CcsR1–4, but no *ccaF1* specific transcript is present in those strains suggesting that longer precursors are translated to produce CcaF1.

Overexpression of CcsRs results in considerably slower growth (doubling time 5.4 h compared to the wild type with 3.5 h), while the doubling time of a strain having CcaF1 with CcsR1–4 together on a plasmid is only slightly increased (4.0 h) compared to that of the wild type. Overexpression of *ccaF*1 alone results in an extended lag phase but growth in exponential phase is not affected (doubling time of 3.4 h). All strains reached the same OD in stationary phase (Figure 3A).

The 71 aa CcaF1 protein of *R. sphaeroides* 2.4.1 consists mostly of the DUF1127 domain (bold in Figure 2A) but harbors 27 additional amino acids in the N-terminal domain and 11 additional amino acids at the C-terminus. In contrast, the *R. sphaeroides* strain KD131 encodes a 50 aa CcaF1 protein (RSKD131–0402) with only six additional amino acids N-terminal of the DUF1127 domain (Figure 2A). Otherwise, the amino acid sequences of the two small proteins are identical. The genomic context around

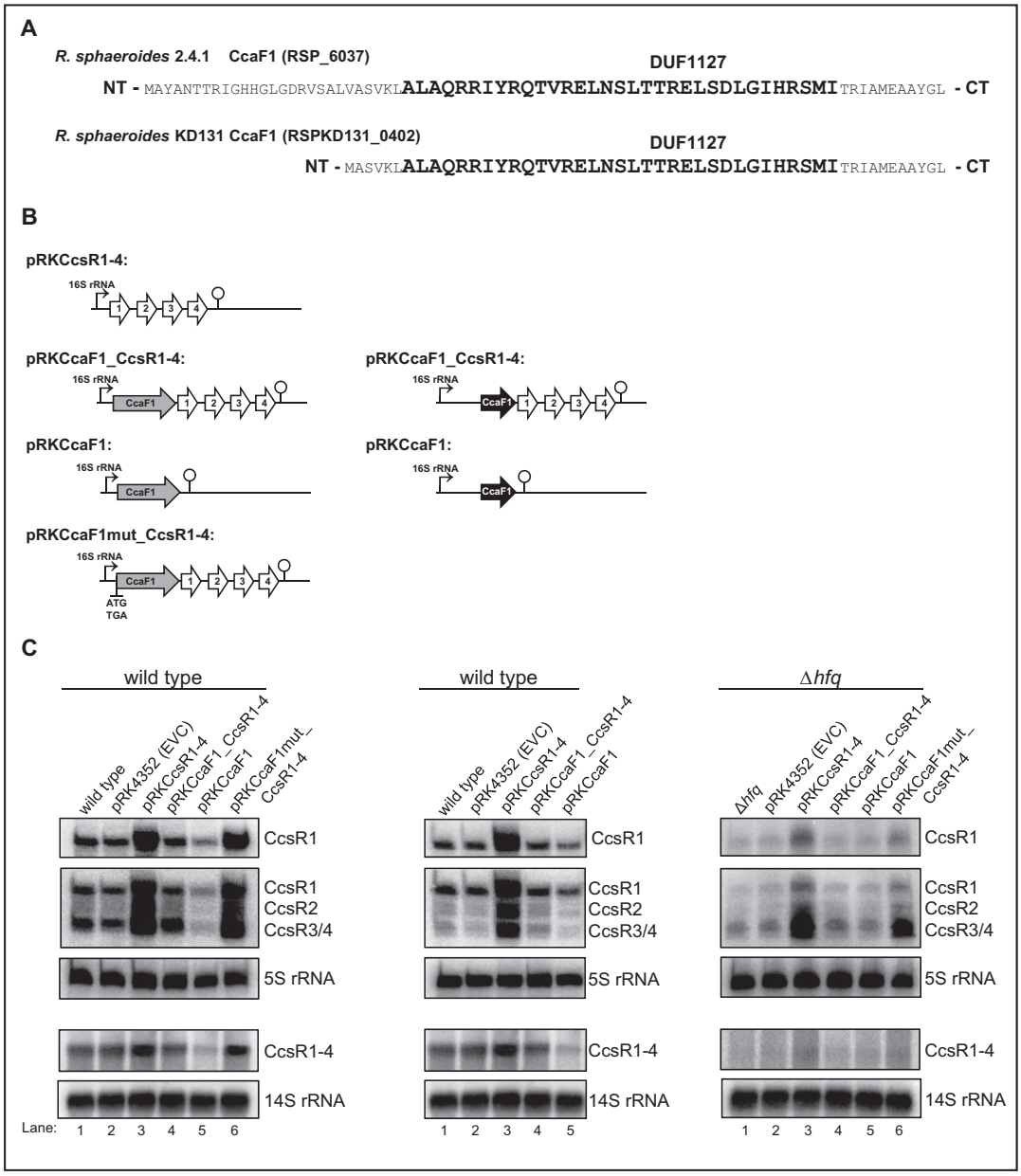


Figure 2. Small DUF1127 proteins affect CcsR level in *R. sphaeroides*. (**A**) Amino acid sequence of the small DUF1127 protein CcaF1 from *R. sphaeroides* 2.4.1 and the corresponding homologue RSPKD131_0402 from *R. sphaeroides* KD131. The conserved DUF1127 domain as shown in Supplementary Figure S3 is high-lighted in bold letters. (**B**) Schematic overview of the plasmids introduced into the wild type strain 2.4.1 or the mutant lacking the *hfq* gene. The *ccaf1* gene is shown in light grey, the RSKD131_0402 gene in black, CcsR RNAs in white. In plasmid pRKCcaF1mut_CcsR1–4 the ATG of the *ccaF1* gene was changed to TGA. (**C**) Northern blots of total RNA from strains containing an empty vector control (pRK4352, EVC) with just the 16S promoter, or plasmids as shown in (B). DNA fragments specific for CcsR1, CcsR2, CcsR3/CcsR4 or CcsR1–4 were used as probes. Signals for 5S RNA or 14S RNA were used as loading controls. *R. sphaeroides* cleaves the 23S RNA into fragments of 16S and 14S (70). The upper three panels stem from 10% denaturing polyacryamide gels, the lower two panels from 1% formaldehyde agarose gels.

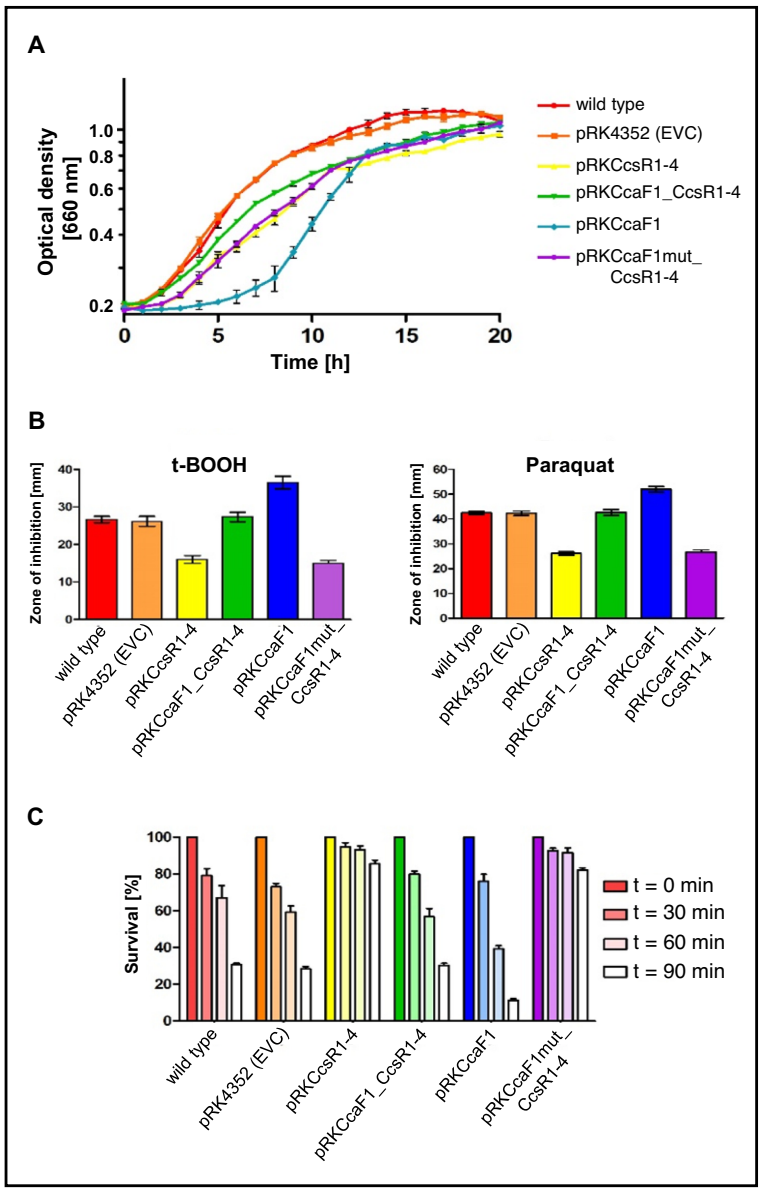


Figure 3. DUF1127 protein CcaF1 (RSP_6037) affects stress resistance in *R. sphaeroides*. (**A**) Growth curves of the wild type or the wild type with the empty vector (EVC) or the plasmids as shown in Figure 2B. All strains were cultivated under microaerobic conditions. The average of three independent measurements and the standard deviation are plotted. The color code for the different strains was also applied in (B) and (C). (**B**) Zone of inhibition assay of strains overexpressing the plasmids shown in Figure 2B in comparison to the wild type and wild type with empty vector control (pRK4352) under organic peroxide (700 mM tBOOH) and superoxide stress (300 mM paraquat). The plotted values represent the mean of at least three independent experiments and the standard deviation is indicated. (**C**) Survival assay of strains overexpressing the plasmids shown in Figure 2B in comparison to the wild type and wild type with empty vector control (pRK4352) under organic peroxide stress (300 mM tBOOH). The number of colonies of a control culture grown without the addition of any oxidative stress agents was defined as 100% survival. The bars represent the mean of three independent plating assays and the standard deviation is indicated.

the *ccaF1-ccsR* locus is also identical in the two strains, the overall synteny is highly conserved between these two *R. sphaeroides* strains. When we expressed the shorter CcaF1 protein of strain KD131 from the 16S promoter on a plasmid (pRKCcaF1) in strain *R. sphaeroides* 2.4.1, we saw a similar reduction in the abundance of the CcsR RNAs (Figure 2C, middle panel, lanes 2 and 5) as that observed with CcaF1 from strain 2.4.1 (Figure 2C, left panel, lanes 2 and 5). This demonstrates that the 50 aa CcaF1 from *R. sphaeroides* KD131 is sufficient for the effect on CcsR levels.

A previous study demonstrated lower amounts of CcsR1–4 in a strain lacking Hfq (7). To test, whether Hfq affects CcaF1 function, we also transferred the different plasmids into the *hfq* mutant and analyzed CcsR levels. As observed in the wild type, CcaF1 counteracted the stronger expression of CcsR1–4 when present on the plasmid together with the *ccsR1–4* genes (Figure 2C, right panel, compare lanes 3 and 4). Overexpression *in trans* of the *ccaF1* gene alone had little effect on the CcsR levels (Figure 2C, right panel, compare lanes 2 and 5). The exact role of Hfq in CcsR maturation and possibly further CcaF1-dependenten processes needs to be addresses in the future.

Increased levels of CcsR1–4 were previously shown to lead to increased resistance of *R. sphaeroides* to the superoxide generating paraquat and to tertiary butyl-alcohol (t-BOOH) (18). t-BOOH represents organic peroxides that are produced from cellular components due to singlet oxygen exposure. Figure 3B confirms the increase of resistance to the two chemicals when CcsR levels are increased. When CcsR1–4 are overexpressed together with the *ccaF1* gene, the resistance level resembles those of the controls, but only if the ATG is not mutated. Overexpression of CcaF1 alone, resulted in decreased resistance to paraquat and t-BOOH compared to the control. This effect of overexpression of CcaF1 on the oxidative stress response was also confirmed in survival assays (Figure 3C). Thus, the effect of the different plasmid constructs on resistance reflects the CcsR amounts that were detected in the Northern blots.

In addition, we performed spot assays to test the growth behavior of the different strains in the presence of various stresses. These assays (Supplementary Figure S4) confirmed the results we obtained by t-BOOH zone of inhibition and survival assays: overexpression of CcsR1–4 (pRKCcsR1–4) resulted in significantly increased survival, while overexpression of CcaF1 (pRKCcaF1) decreased survival. The same correlation between expression and survival was observed for heat stress (42°C) or stress by $CdCl_2$. The CcaF1-CcsR1–4 overexpression strain showed similar survival as the wild type and the control strain harboring an empty vector in presence of salt (NaCl), while the strain overexpressing CcaF1 alone showed reduced survival (Supplementary Figure S5).

RNase E and CcaF1 are involved in processing of the *ccaF1-ccsR* transcript and maturation of the CcsR RNAs

When the CcsR1–4 RNAs (originally designated RSs0680 a-d) were detected as photooxidative stress-induced sRNAs in a RNAseq data set, their co-transcription was already proposed based on 5′ RACE and RT-PCR (7). This strongly suggested that the individual sRNAs are generated by RNA processing.

To exclude that overexpression of *ccaf1* affects CcsR levels by altering promoter activity, we applied reporter constructs that have the *ccaF1* promoter including 100 nt or 200 nt upstream of the promoter transcriptionally fused to eCFP (16) and monitored fluorescence. As shown in Supplementary Figure S6 *Rhodobacter* strains without eCFP exhibit autofluorescence, which was set to 100% relative fluorescence. Presence of the promoter:eCFP fusions elevated the fluorescence to 120–135% during exponential growth at 32°C. After heat stress, fluorescence was increased to 160–190% independently of overexpression of *ccaF1* (Supplementary Figure S6). This excludes an effect of CcaF1 on the activity of the *ccaF1* promoter and supports the assumption that RNA processing/degradation is responsible for the effect of CcaF1 on CcsR levels.

The endoribonuclease E is involved in the maturation of several sRNAs from the 3′ or 5′ UTR of mRNAs in *E. coli* and *R. sphaeroides* (45,41). Despite the different GC contents (51% for *E. coli*, 69% for *R. sphaeroides*) RNase E recognizes AU rich sequences in both organisms (41). The RNAseq data set that compares total RNA reads and RNA 5′ ends in the control strain and in a mutant that expresses a temperature sensitive RNase E from *E. coli* (41) reveals RNase E cleavage sites at the 5′ ends of the individual CcsR RNAs (Supplementary Figure S7). A main 5′ end is detected in the wild type at 32°C and due to induction of the RpoHI/HII-dependent promoter shows higher abundance at 42°C. In the strain expressing the temperature-sensitive RNase E the 5′ end is more abundant at 32°C compared to the wild type and this difference is more pronounced at the non-permissive temperature of 42°C. Our previous study demonstrated that RNase E cleavage is already partly impaired in the mutant at 32°C due to the different RNase E enzyme (41). RNase E-dependent 5′ ends are also detected for the individual CcsR RNAs (Supplementary Figure S7). Cleavages at the 5′ ends occur at the sequences G**UUU**CC (for CcsR1, nucleotides adjacent to cleavage sites in bold), C**UCU**UC (for CcsR2), A**CUU**C (for CcsR3) and A**CU**UC (for CcsR4). To further confirm the important role of RNase E in CcsR maturation, we performed Northern blots with probes directed against the CcsR1 RNA or against precursor transcripts (Figure 4). At time point 0 the cultures were shifted to 42°C which leads to induction of the RpoHI/HII-dependent promoter (Supplementary Figure S7) and also to inactivation of RNase E. Figure 4 demonstrates increased levels of the precursor transcripts (a precursor harbouring *ccaF1* and CcsR1–4 would comprise 680 nt, a precursor including CcsR1–4 about 400 nt) as well as strongly increased CcsR1 levels after the shift to 42°C in the wild type. In the rne^{ts} mutant however, strong accumulation of precursor transcripts occurs but the level of CcsR1 is strongly decreasing, supporting a major role for RNase E in CcsR maturation.

We also shifted strain 2.4.1 (pRKCcaF1) to 42°C, which overexpresses *ccaF1*. The accumulation of CcsR1 upon the temperature shift was clearly reduced compared to the wild type (Figure 4). At the same time, precursor transcripts

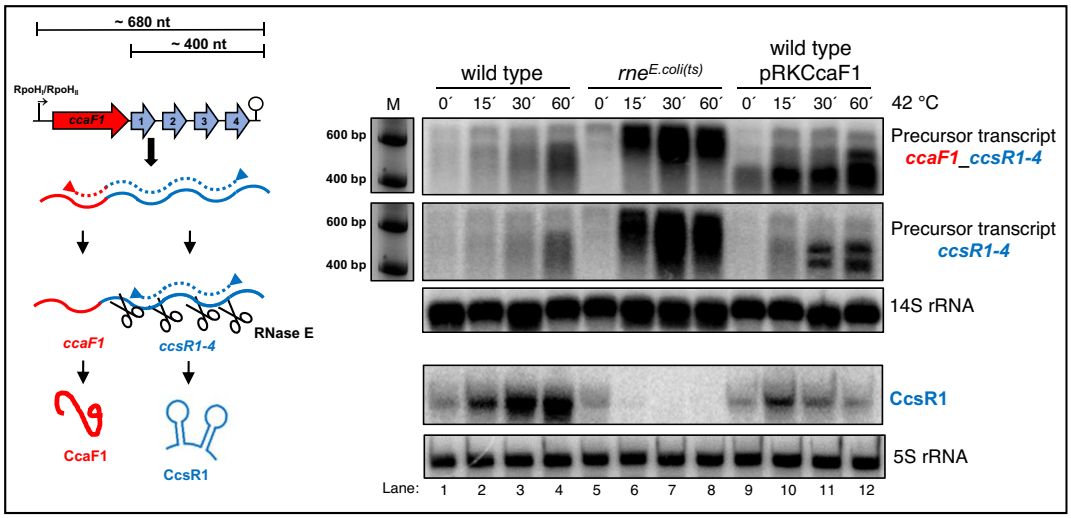


Figure 4. RNase E and CcaF1 are involved in CcsR maturation. Analysis of *ccaF1-ccsR1–4* and *ccsR1–4* precursor and CcsR1 RNA by Northern blot in *R. sphaeroides* 2.4.1 wild type, *rne*^{ts} mutant and strain 2.4.1 that overexpresses *ccaF1* (pRKCcaF1). Cells were harvested at 32°C and at different time points after shift to 42°C. Total RNA was isolated and either run on a 1% formaldehyde agarose gel and, after blotting, hybridized against probes spanning the CcaF1-CcsR1–4 or the CcsR1–4 region, or on a 10% denaturing polyacryamide gel for detection of CcsR1. 5S and 14S rRNA serve as loading controls.

accumulate that show however a different pattern than the precursors accumulating in the mutant with reduced RNase E activity (Figure 4). This demonstrates that not only RNase E but also CcaF1 is involved in maturation of the CcsR RNAs.

CcaF1 directly binds to CcsR RNAs

The high arginine content of the DUF1127 domain, the structural similarity to the Smaug domain and its role in CcsR maturation suggest that CcaF1 may have RNA-binding capacity. To test this hypothesis, we applied the purified protein for *in vitro* RNA binding assays. CcaF1 was purified as His_6-MBP-fusion protein, which was cleaved by TEV protease to release CcaF1 (Supplementary Figure S8). Figure 5 shows gel retardation assays with radiolabeled CcsR1 or CcsR1–4. While addition of CcaF1 to CcsR1 leads to a single retarded band (A), two distinct bands were visible with the longer CcsR1–4 transcript. This suggests that more than one CcaF1 protein (or protein complex) can bind to CcsR1–4. Addition of an excess of unlabeled CcsR1 to the reaction abolished the retardation, indicating that the binding is specific (Supplementary Figure S9 A). When we added unlabeled RSs0827 RNA, even at high molar excess complex formation between CcsR1 and CcaF1 was not abolished (Supplementary Figure S9B). RSs0827 is the most abundant sRNA in stationary phase in *R. sphaeroides* (21) and was not found in the CoIP analysis with CcaF1 (see below). Thus, we selected it as unspecific competitor in our assay. Furthermore, we used a CcsR1 variant that has four nucleotides in loop 1 exchanged, which is predicted to also change the local structure. This mutant variant was not bound by CcaF1 (Figure 5B).

Effect of CcaF1 overexpression on the transcriptome of *R. sphaeroides*

To analyze the global effect of CcaF1 on the transcriptome, we performed RNAseq on total RNA from a strain harboring plasmid pRKCcaF1 (Figure 2B) and a control strain only harboring the vector. Strains were cultivated under microaerobic conditions at 32°C. The reads for the individual genes and non-coding RNAs were compared for the two strains. PCA plots verified the high reproducibility of the technical triplicates, each from biological triplicates (Figure 6A). For few selected genes we also performed real time RT-PCR with the RNA samples used for RNAseq analysis to verify the results (not shown). Although the fold-changes showed some variation between the two methods, the direction of change was consistent for the tested RNAs.

A volcano plot displays the strongest differences in read numbers between the strain overexpressing CcaF1 and the wild type strain (Figure 6B). In order to avoid mis-leading results due to low read numbers, we set a basemean of $\geq$100. Thus, we consider only genes with a normalized average read count over all samples above this threshold to dismiss potential false positive genes and to increase reliability. With this cut-off only 1487 of total 4411 genes remained in our analysis. We further chose a cut off $\log_2$fold $\leq$–1 or $\log_2$fold $\geq$1 and an adjusted *P*-value of $\leq$0.05 (Benjamini-Hochberg algorithm). The strongest fold change (marked in blue) was observed for *ccaF1*, which was overexpressed from the plasmid. In addition to *ccaF1*, 42 RNAs showed increased levels of $\log_2$fold $\geq$ 1.0 (Supplementary Table S3). Among these genes were RSP_3095 and RSP_3094 for a sigma-70 factor and an anti-sigma factor that are involved in the adaptation to stationary phase (46), and *sitA*, *sitB, sitC* for an Mn^{2+} ABC transporter. The enriched protein-

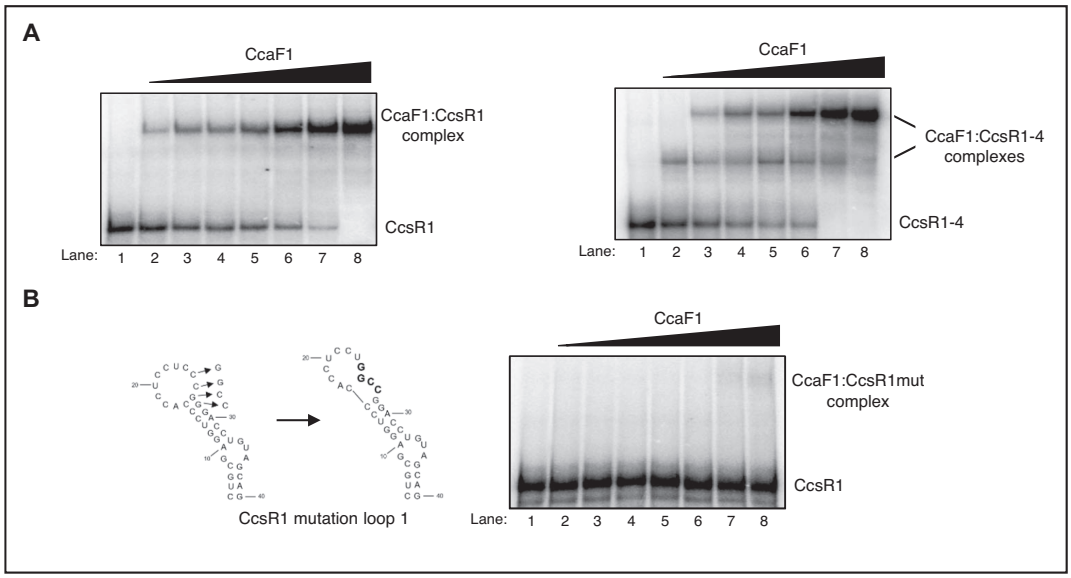


Figure 5. CcaF1 binds to CcsR1 and the CcsR1–4 precursor. (**A**) Gel motility shift RNA-binding experiments with 150 fmol radioactively labelled CcsR1 or CcsR1–4 incubated with increasing amounts (2, 5, 10, 50, 100, 200 or 500 nM) of purified CcaF1. The reactions were run on a 6% native polyacrylamide gel. (**B**) Predicted structures of the 5′ hairpin in the CcsR1 RNA with and without a mutation of four nucleotides (CCGG→GGCC) in the loop (left side) and gel motility shift experiments with 150 fmol radioactively labelled, mutated CcsR1 RNA incubated with increasing amounts (2, 5, 10, 50, 100, 200 or 500 nM) of purified CcaF1 (right side). The reactions were run on a 6% native polyacrylamide gel.

encoding RNAs belong to different COG (cluster of orthologous groups) functions, 8 of them encode hypothetical proteins. Three non-coding RNAs were enriched, among them SorX (formerly RSs2461). SorX is derived from the 3′ UTR of the *ompR*-1 gene by RNase E cleavage and affects resistance to singlet oxygen and organic hydroperoxides by interacting with the mRNA for a polyamine transporter (16,41,47).

Nineteen RNAs showed lower levels in the overexpression strain compared to the wild type, among them one non-coding RNA of unknown function, 23S and 16S rRNAs and six mRNAs encoding hypothetical proteins (Supplementary Table S3). Three of these 19 RNAs encode cold shock proteins (RSP_3620, RSP_1952 and RSP_3621). Several tRNAs were also decreased in the overexpression strain but the *P*-values were above our cut-off. Likewise, in our analysis the *P*-values for the CcsR RNAs and for other known non-coding RNAs were too high for reliable conclusions.

The RNAseq results demonstrate that CcaF1 does not only affect CcsR levels (as demonstrated by northern blots) but several other cellular RNAs with different physiological functions. These effects may be direct, by interaction of CcaF1 with those RNAs, or indirect through other RNAs that interact with CcaF1.

Co-immunoprecipitation identifies RNA targets of CcaF1

To discriminate between direct and indirect effects of CcaF1 on RNA levels, we expressed a CcaF1 variant with an N-terminal FLAG-tag and performed co-immunoprecipitation with total RNA from *R. sphaeroides*. We confirmed that the tagged CcaF1 protein has the same effect on CcsR levels as the non-tagged-version (Supplementary Figure S10) and is thus functional. The total CoIP sample was analyzed on an SDS gel and by high resolution MS (LS-ESI-HRMS). Silver stain of the SDS gel shows as faint band that was confirmed as CcaF1FLAG by western blot (Supplementary Figure S11). Faint bands at higher molecular weights indicate the formation of stable multimers. By far the most abundant protein detected by MS of this sample was CcaF1, no other protein was present in similar amounts (Supplementary Table S4). Due to the high sensitivity of the MS, minor amounts of other proteins were detected. Most of these proteins are abundant in the cell and are known to interact with many proteins (GroES, GroEL) or to interact with RNA (ribosomal proteins, Rho, TufA) that was also present in the analyzed CoIP sample. Neither Hfq nor RNase E were detected in the CoIP sample, excluding their direct association with CcaF1.

The co-immunoprecipitated RNA was used for RNAseq (Rip-seq) and compared to RNAseq from total RNA and to a negative control, a CoIP with RNA from cells not expressing the FLAG-tagged CcaF1. From the RNAs identified in the CoIP, a subset was selected for further analysis using northern blot and real time RT-PCR. Supplementary Table S5 provides a quantification of the RNAs that were strongly enriched in the co-immunoprecipitation. This table considers only RNAs with >10 reads in the CoIP sample, and at least 10-fold higher read number in the CoIP sample compared to the control (CcaF1 overexpression without FLAG-tag). The strongest enrich-

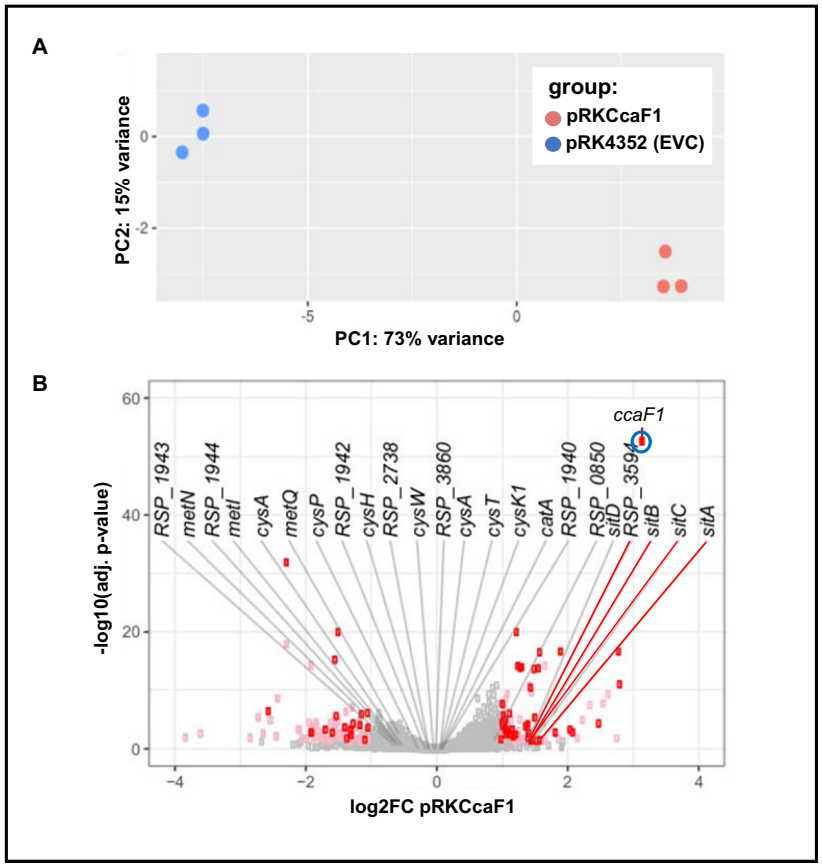


Figure 6. Overexpression of CcaF1 impacts the *R. sphaeroides* transcriptome. (**A**) A principal component analysis was performed as part of the downstream DESeq2 analysis. The scatterplot shows two distinct groups, each one harbors the replicates belonging to one of the individual strains used in this study. (**B**) Volcano plot for the comparison of the *ccaF1* overexpressing strain and the wild type, based on RNAseq data. Genes with significant change in abundance are colored red (adjusted *P*-value ≤ 0.05, $\log_2$fold change ≤–1 or ≥1, basemean > 100) and pink (adjusted *P*-value ≤0.05, $\log_2$fold change ≤–1 or ≥1, basemean ≤ 100). Gray dots: adjusted *P*-value >0.05. The 24 RNAs with highest enrichment in the CoIP are labelled.

ment was observed for the mRNA of the catalase gene (*catA*, RSP_2779, factor 219), the *sit* genes (ABC Mn transporter), and for RSP_1943 (hypothetical protein) and RSP_1944 (Uroporphyrin-III-methyltransferase / siroheme synthase). Interestingly, many RNAs which were strongly enriched in the CoIP have known or predicted roles in cystein / methionine / sulfur metabolism: RSP_1944 (methyltransferase), RSP_1942 (sulfite/nitrite reductase), *cysH* (RSP_1941, phosphoadenosine phosphosulfate reductase), *cysK1* (RSP_1109, cystein synthase), *cysA*, *cysP*, *cysT*, *cysW* (RSP_3696–3699), and RSP_3861 (ABC sulfate/ thiosulfate transporter), RSP_3860 (probable rhodanese-related sulfurtransferase), RSP_3859 (ABC sulfate/molybdate transporter), *metN*, *met*Q and *metI* (RSP_0129, 0130, 0132, methionine uptake transporter). The RNAseq data indicate that the adjacently located genes are transcribed into long polycistronic transcripts.

Read coverage plots for the CoIP results for selected non-coding RNAs and mRNAs are shown in Figure 7 together with results from RT-PCR. Quantitative data from real time RT-PCR from the CoIP are shown in Supplementary Figure S12. We observed accumulation of CcsR1, UpsM, and PcrZ, 6S RNA, SRP RNA, tmRNA, and tRNA Gly when the FLAG-tagged CcaF1 was present. UpsM is a highly abundant sRNA that is derived from the 5′ UTR of the *dcw* gene cluster mRNA by RNase E cleavage (41,48) and PcrZ has an important role in the regulation of photosynthesis genes (28,49). *puf* and *puc* genes encode pigment-binding proteins of the photosynthetic apparatus. For the small RNA UpsM (enrichment in CoIP about 12-fold, Supplementary Figure S12) we also demonstrated direct binding to CcaF1 by gel retardation (Supplementary Figure S9C). As expected from the enrichment factors in the CoIP (Supplementary Figure S12) higher amounts of CcaF1 were required for complex formation with UpsM than with CcsR1.

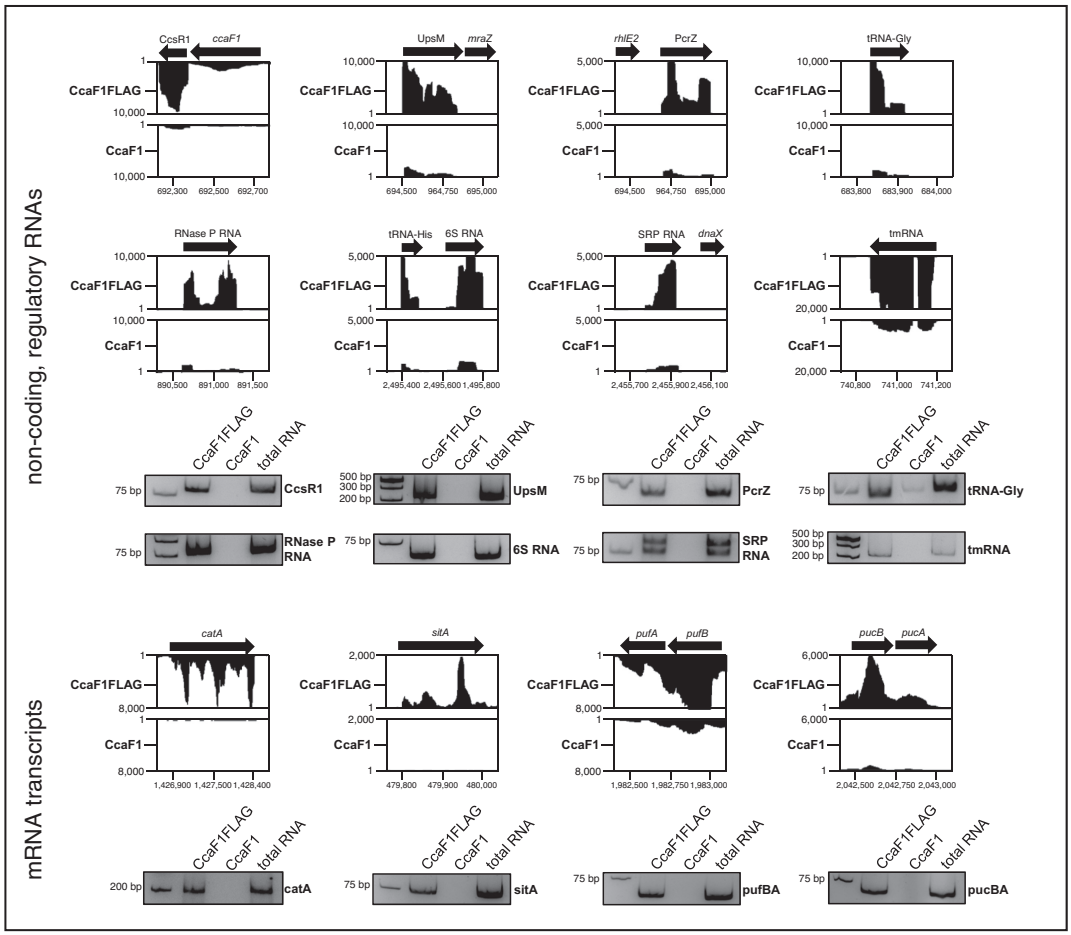


Figure 7. CoIP analysis identifies targets of CcaF1. Analysis of co-immunoprecipitated RNA by RNAseq (Rip-seq) using CcaF1 with 3xFLAG-tag (CcaF1FLAG) or without 3xFLAG-tag (CcaF1, control) in exponential growth phase at 32°C and microaerobic conditions. Read coverage plots from the Integrated Genome Browser display the sequencing reads for selected RNAs. The specific non-coding RNAs and mRNA transcripts of the co-immunoprecipitation were also analyzed by reverse transcription (RT) PCR. The RT-PCR products were separated on a 10% polyacrylamide gel and analyzed by ethidium-bromide staining and are shown below the corresponding read coverage plots.

Of the RNAs with expression changes in the DESeq2 analysis upon over-expression of CcaF1, only 5 RNAs were enriched in the CoIP (S3 Table): *expE1* for a hemolysin-type calcium-binding region, *sitA*, *sitB*, *sitC* encoding subunits of an ABC Mn^{2+} transporter and RSP_0850 for a hypothetical protein. This supports the view that many changes in the transcriptome are not due to a direct interaction with CcaF1. Vice versa, one could expect that a direct interaction to CcaF1 as suggested by the CoIP would result in changed levels of those RNAs in the DESeq2 analysis. As seen in Supplementary Table S5, this is only true for the five RNAs mentioned above. All other RNAs enriched in the CoIP did not reach the *P*-value we used for the cut-off in the DESeq2.

To further support the view that CcaF1 specifically binds to a set of RNAs we also performed CoIP analysis with the FLAG-tagged RSP_0557 protein. RSP_0557 encodes a 70 amino acid long DUF1127 protein which is not found in a CIN1 locus. The gene is under control of a RpoHI/HII-dependent promoter and is controlled by the sRNA Pos19 (17). Previous studies revealed strong increase of *ccaF1* and RSP_0557 transcript levels in transition from exponential to stationary phase (46). To date no function could be assigned to the RSP_0557 protein. Supplementary Figure S13 demonstrates that some RNAs like CcsR1, UpsM and tmRNA are bound by both, CcaF1 and RSP_0557 proteins, with similar efficiency. However, other RNAs are preferentially bound by only one of the two DUF1127 proteins. These data also support the view that not only DUF1127 proteins of CIN1 loci function as RNA-binding proteins.

CcaF1 affects stability of some of its targets

We assumed that CcaF1 controls the amounts of its targets by either affecting their maturation, their stability, or both. To test the effect of CcaF1 on stability, we compared half-lives of selected RNAs in the wild type and the CcaF1 overexpression strain by quantifying the RNA levels after addition of rifampicin (inhibits initiation of transcription) by Northern blot analyses (Figure 8). These experiments confirmed a destabilizing effect of CcaF1 on the sRNAs CcsR1 and UpsM, the RNaseP-RNA and the *pufBA* mRNA. Such an effect was not observed for the sRNA PcrZ, tRNA-Gly. These RNAs turned out to be very stable and half-life determination after long time periods in the presence of rifampicin is not reliable. We also did not observe changed half-lives for *pucBA*, or *catA*, although the half-lives were in the same range as for CcsR1, which was less stable in the overexpression strain. *PucBA* is part of a polycistronic transcript, while *catA* mRNA is monocistronic. The effect of CcaF1 on the amounts of these RNAs needs further investigation.

We conclude that CcaF1 can control RNA levels by affecting the stability of the mature transcript, but that also other mechanisms, like maturation from precursor transcripts are involved.

DISCUSSION

Although numerous small open reading frames are found in bacterial genomes, the importance of small proteins was realized only about a decade ago (1,2). Numerous small proteins, mostly found in alpha- and g–ammaproteobacteria, harbor the domain of unknown function DUF1127. First evidence for the involvement of a DUF1127 protein in bacterial physiology was provided for RSP_6037 (CcaF1) that has a role in stress responses in *R. sphaeroides* (18). In *Brucella abortus* deletion of the gene for a DUF1127 protein caused a defect in fucose metabolism (50). Recently a role of DUF1127 proteins in phosphate and carbon metabolism in *Agrobacterium tumefaciens* was demonstrated (51), as well as a role of the *Salmonella* protein YjiS in virulence (52). YjiS is a DUF1127 protein with 20% identity to CcaF1. The mechanisms by which these DUF1127 proteins affect physiology remain elusive. This study identifies CcaF1 of *R. sphaeroides* as a new type of RNA-binding protein.

The *ccaF1* gene of *R. sphaeroides* is co-transcribed with 4 homologous sRNAs. Our bioinformatic analysis revealed that the arrangement of sRNAs and genes for DUF1127 proteins (sRNAs in the 5′ or 3′ UTR) in CIN1 loci correlates with phylogenetic assignments by 16S rRNA. If duplicate loci occur in a family, these loci fall into two distinct clusters. Remarkably one single genus (*Sulfitobacter*) showed 'Cuckoo'-RNAs adjacent to both sides of a gene for a DUF1127 protein. The high correlation between the 16S rRNA phylogenetic tree and the phylogenetic tree based on CcaF1 amino acid sequences indicates, that the DUF1127 protein coding ORFs have likely been acquired by a common ancestor and sequences co-evolved. This hypothesis is underlined by the observation, that two similar CIN1 loci, which might be a result of a locus duplication, occur in many *Brucellaceae* and form distinct clusters based on the CcaF1 amino acid sequence. Furthermore, occurrence of multiple highly similar 'Cuckoo'-RNAs in *Sinorhizobium* on the chromosome and on plasmids, with only weak association with genes for DUF1127 proteins, shows that CIN in general and CIN1 loci in particular are subject to duplication and genomic reorganization and may result in different locus configurations.

Our data revealed a strong effect of the CcaF1 protein on CcsR levels, which did not require the N-terminal 21 amino acids. CcaF1 promotes RNase E-dependent cleavage of the *ccaF1*-CcsR1–4 precursor transcript. The cleavages occur adjacent to U residues, a preference that was also revealed by a previous global study mapping RNase E cleavage sites in *R. sphaeroides* (41). Overexpression of CcaF1 also resulted in the accumulation of precursor transcripts, that showed a pattern differing from the precursors accumulated in the rne^{ts} mutant. This suggests that CcaF1 is not solely reducing RNase E activity. It is conceivable that structural changes upon binding of CcaF1 affect cleavages of the precursor transcripts. This needs to be tested in the future. Gel retardation assays with purified CcaF1 proved that the protein can specifically bind to CcsR1 and to the CcsR1–4 precursor.

RNAseq analyses demonstrated that not only the amount of CcsR RNAs is influenced by CcaF1 but that this protein affects the levels of many RNAs including sRNAs, tRNAs, rRNAs, other non-coding RNAs like 6S RNAs or signal recognition RNA, and mRNAs. While the non-coding RNAs are mostly small and highly structured, also the amount of several large mRNAs, including polycistronic transcripts was affected. Considering the mRNAs mostly affected in the RNAseq data, there was no clear preference for certain functional groups, indicating that CcaF1 may influence multiple biological functions. Our experiments verified an effect of CcaF1 on resistance to oxidative stress, $CdCl_2$ and heat stress.

Results from co-immunoprecipitation with tagged CcaF1 followed by RNAseq support a direct interaction between CcaF1 and many RNAs, including sRNAs, tRNAs, rRNAs, other non-coding RNAs and mRNAs. *In vitro* gel retardation experiments with CcaF1 and CcsR1, CcsR1–4 or UpsM confirmed that CcaF1 is a *bona fide* RNA binding protein. How can a small protein affect the levels of many different RNAs? It is unlikely that the small DUF1127 domain exhibits a catalytic function that degrades RNA. But binding of proteins to RNA can restrict or enhance the action of ribonucleases (53). Such a function of CcaF1 is supported by our data: the amount of CcsR RNAs is dependent on RNase E activity and CcaF1 abundance. In *Drosophila,* Smaug can recruit the Argonaute 1 protein to an mRNA to trigger translational repression and/or decay (54). Argonaute proteins are also encoded by many bacterial and archaeal genomes (55) and a plasmid-encoded Argonaute from *Rhodobacter sphaeroides* ATCC17025 was investigated in more detail (56–58). The genome of *R. sphaeroides* 2.4.1 does not encode such a protein excluding the involvement of Argonaute in CcaF1-dependent RNA destabilization.

Several RNA-binding proteins, some of them small proteins, have established functions as RNA chaperones in prokaryotes (59,60). The Hfq protein (77 aa in *R. sphaeroides*) is considered as a global regulator of sRNA-based networks. It acts as an RNA chaperone in gram-negative bacteria by stabilizing the imperfect base-pairing

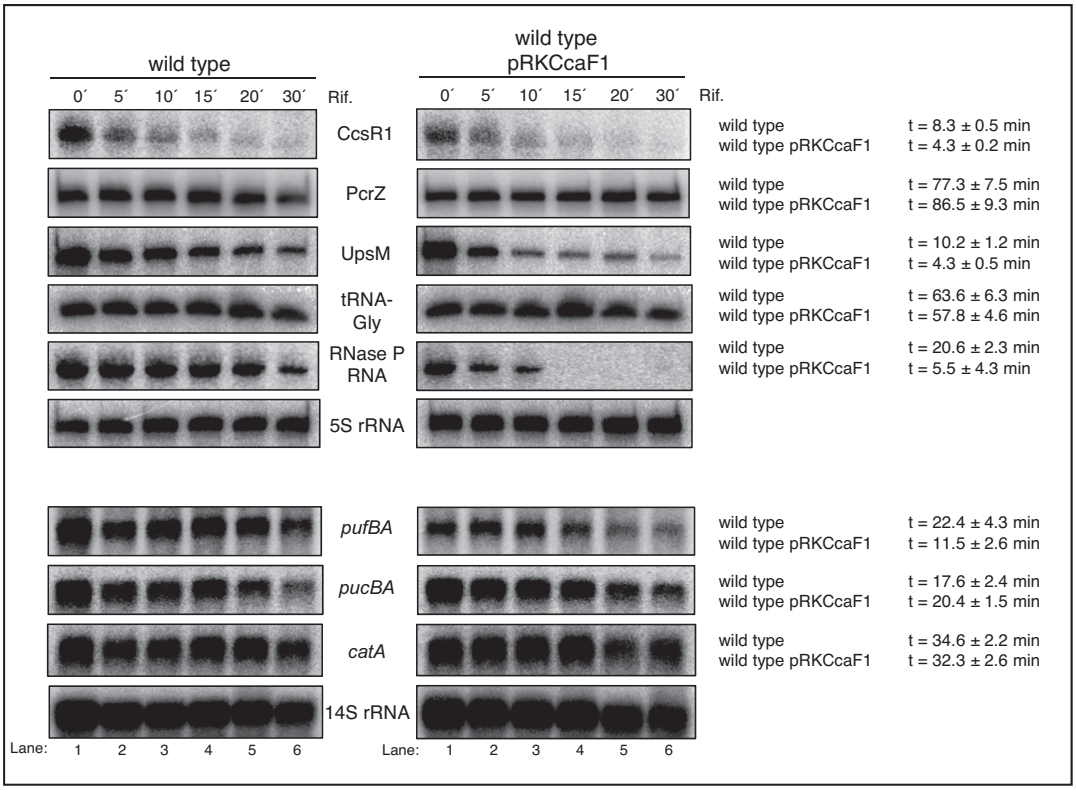


Figure 8. CcaF1 affects stability of some coding and non-coding RNA transcripts. Determination of RNA half-life of *R. sphaeroides* wild type and wild type with pRKCcaF1 for selected coding and non-coding RNA transcripts. Cells were harvested in exponential growth phase under microaerobic conditions. Samples were taken after adding rifampicin at different time points. Total RNA was isolated and either run on a 10% denaturing polyacryamide gel (upper six panels) or on a 1% formaldehyde agarose gel (lower four panels) and blotted. 5S or 14S rRNA serve as loading controls. For quantification RNA signal intensities were normalized to 5S or 14S rRNA signals. The average half-life was calculated from three independent experiments and the standard deviation is indicated.

between trans-encoded sRNAs and their mRNA targets. The 72 aa CsrA protein of *E. coli* (member of the CsrA/RsmA family) recognizes an AUGGA motif in RNA loop regions. It preferentially binds to the ribosome binding site or to the start codon of mRNAs and either represses translation or regulates transcript stability. CsrA was also shown to act as a chaperone that can promote complex formation between an sRNA and its mRNA target in *Bacillus subtilis* (61). ProQ (about 220 aa) is an RNA chaperone of the FinO family that is commonly found in Proteobacteria (62). It binds double-stranded RNAs and prefers highly structured RNAs, mostly promoting binding of sense and anti-sense RNAs, but can also regulate trans-acting sRNAs (63). CspA is another small (68 aa in *R. sphaeroides*) RNA chaperone and can passively remodel RNA structures by preferentially binding to pyrimidine-rich RNA sequences (64). In *Staphylococcus aureus* a RIP-CHIP assay also identified sRNAs as CspA targets (65). CspA binding destabilizes secondary structures to promote translation or alter mRNA turnover. CcaF1 has only low sequence similarity (maximal 30%) to these well studied RNA-binding proteins and no obvious structural homology (based on Phyre 2).

The CcaF1 coding region which is part of the *ccaF1-CcsR* precursor transcript is also enriched in the CoIP. Higher amounts of CcaF1 lead to impeded maturation of the CcsR RNAs and to stabilization of the precursor RNA, strongly suggesting that processing by RNase E is negatively affected. Nevertheless, further degradation of CcsR1 is accelerated by CcaF1. Thus, binding of CcaF1 can have different effects on the stability of individual transcripts. To understand the exact mechanisms, how CcaF1 acts on the stability of its targets, future work needs to identify all RNases involved in the maturation and degradation of CcaF1 targets and to follow changes in RNA structure upon binding by CcaF1. Such structural changes may lead to sequestration as well as to exposure of RNase cleavage sites, which may cause stabilization or destabilization of transcripts.

A binding motif for the RNA-binding SAM domain of the eukaryotic Smaug protein was identified in the past (66,67). The Smaug-recognition element (SRE) consists of a stem–loop structure with the sequence CUGGC in the loop. A translational control element (TCE) of 184 nucleotides is required for translational control of *nos* mRNA in *Drosophila*. The TCE contains a pair of redundant

SREs (68,69). Interestingly, the two stem-loops of the CcsR RNAs that were highly enriched in the CcaF1 CoIP, also mostly contain the CUGGC sequence. For many of the CcaF1 binding partners, pairs of stem–loop structures are predicted (not shown), but identification of CcaF1 bindings sites will require further investigation.

The CoIP data strongly suggest that another small DUF1127 protein from *R. sphaeroides*, the RSP_0557 protein binds to RNA. RSP_0557 is not associated with *ccsR* genes on the chromosome. Despite their small size and the strong similarity of the DUF1127 domains, the CcaF1 and RSP_0557 proteins show differences in their preference for RNA binding partners. It will be interesting to further elucidate the molecular basis for this binding specificity.

DATA AVAILABILITY

The sequencing data are available at NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo) under the accession number GSE144523 and GSE145045.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Kerstin Haberzettl and Andrea Weisert (Klug lab) for excellent technical assistance, Elnaz Jajin (Förstner lab) for help with mapping of the CoIP reads, Günter Lochnit (University of Giessen) for the MS analysis, Jeong Kock Lee (Sogang University, South Korea) for providing strain *R. sphaeroides* KD131, and Oliver Roßbach, Patrick Barth, and Kai Thormann (University of Giessen) for support of our project. We are grateful to Matthew McIntosh for corrections on the manuscript. We thank the Core Unit SysMed at the University of Würzburg for RNA-seq data generation.

FUNDING

Deutsche Forschungsgemeinschaft [Kl 561/37-1, RTG 2355]; IZKF at the University Würzburg [project Z-6]. Funding for open access charge: Deutsche Forschungsgemeinschaft/University of Giessen.
Conflict of interest statement. None declared.

REFERENCES

1. Hobbs,E.C., Fontaine,F., Yin,X. and Storz,G. (2011) An expanding universe of small proteins. *Curr. Opin. Microbiol.*, **14**, 167–173.
2. Storz,G., Wolf,Y.I. and Ramamurthi,K.S. (2014) Small proteins can no longer be ignored. *Annu. Rev. Biochem.*, **83**, 753–777.
3. Braatsch,S., Gomelsky,M., Kuphal,S. and Klug,G. (2002) A single flavoprotein, AppA, integrates both redox and light signals in *Rhodobacter sphaeroides*. *Mol. Microbiol.*, **3**, 827–569
4. Zeilstra-Ryalls,J.H. and Kaplan,S. (2004) Oxygen intervention in the regulation of gene expression: the photosynthetic bacterial paradigm. *Cell. Mol. Life Sci.: CMLS*, **61**, 417–436.
5. Mackenzie,C., Eraso,J.M., Choudhary,M., Roh,J.H., Zeng,X., Bruscella,P., Puskás,A. and Kaplan,S. (2007) Postgenomic adventures with *Rhodobacter sphaeroides*. *Annu. Rev. Microbiol.*, **61**, 283–307.
6. Yin,L. and Bauer,C.E. (2013) Controlling the delicate balance of tetrapyrrole biosynthesis. *Philos. Trans. Roy. Soc. London. B, Biol. Sci.*, **368**, 20120262.
7. Berghoff,B.A., Glaeser,J., Sharma,C.M., Vogel,J. and Klug,G. (2009) Photooxidative stress-induced and abundant small RNAs in *Rhodobacter sphaeroides*. *Mol. Microbiol.*, **74**, 1497–1512.
8. Nuss,A.M., Glaeser,J., Berghoff,B.A. and Klug,G. (2010) Overlapping alternative sigma factor regulons in the response to singlet oxygen in *Rhodobacter sphaeroides*. *J. Bacteriol.*, **192**, 2613–2623.
9. Berghoff,B.A., Konzer,A., Mank,N.N., Looso,M., Rische,T., Förstner,K.U., Krüger,M. and Klug,G. (2013) Integrative "omics"-approach discovers dynamic and regulatory features of bacterial stress responses. *PLos Genet.*, **9**, e1003576
10. Glaeser,J., Nuss,A.M., Berghoff,B.A. and Klug,G. (2011) Singlet oxygen stress in microorganisms. *Adv. Microb. Physiol.*, **58**, 141–173.
11. Hess,W.R., Berghoff,B.A., Wilde,A., Steglich,C. and Klug,G. (2014) Riboregulators and the role of Hfq in photosynthetic bacteria. *RNA Biology*, **11**, 413–426.
12. Adnan,F., Weber,L. and Klug,G. (2015) The sRNA SorY confers resistance during photooxidative stress by affecting a metabolite transporter in *Rhodobacter sphaeroides*. *RNA Biol.*, **12**, 569–577.
13. Rui,B., Shen,T., Zhou,H., Liu,J., Chen,J., Pan,X., Liu,H., Wu,J., Zheng,H. and Shi,Y. (2010) A systematic investigation of *Escherichia coli* central carbon metabolism in response to superoxide stress. *BMC Syst. Biol.*, **4**, 122.
14. Valdivia-González,M., Pérez-Donoso,J.M. and Vásquez,C.C. (2012) Effect of tellurite-mediated oxidative stress on the *Escherichia coli* glycolytic pathway. *Biometals*, **25**, 451–458.
15. Bignucolo,A., Appanna,V.P., Thomas,S.C., Auger,C., Han,S., Omri,A. and Appanna,V.D. (2013) Hydrogen peroxide stress provokes a metabolic reprogramming in *Pseudomonas* fluorescens: enhanced production of pyruvate. *J. Biotechnol.*, **167**, 309–315.
16. Peng,T., Berghoff,B.A., Oh,J.-I., Weber,L., Schirmer,J., Schwarz,J., Glaeser,J. and Klug,G. (2016) Regulation of a polyamine transporter by the conserved 3′ UTR-derived sRNA SorX confers resistance to singlet oxygen and organic hydroperoxides in *Rhodobacter sphaeroides*. *RNA Biol.*, **13**, 988–999.
17. Müller,K.M.H., Berghoff,B.A., Eisenhardt,B.D., Remes,B. and Klug,G. (2016) Characteristics of Pos19 - a small coding RNA in the oxidative stress response of *Rhodobacter sphaeroides*. *PLoS One*, **11**, e0163425.
18. Billenkamp,F., Peng,T., Berghoff,B.A. and Klug,G. (2015) A cluster of four homologous small RNAs modulates C1 metabolism and the pyruvate dehydrogenase complex in *Rhodobacter sphaeroides* under various stress conditions. *J. Bacteriol.*, **197**, 1839–1852.
19. Nuss,A.M., Glaeser,J. and Klug,G. (2009) RpoH(II) activates oxidative-stress defense systems and is controlled by RpoE in the singlet oxygen-dependent response in *Rhodobacter sphaeroides*. *J. Bacteriol.*, **191**, 220–230.
20. Dufour,Y.S., Landick,R. and Donohue,T.J. (2008) Organization and evolution of the biological response to singlet oxygen stress. *J. Mol. Biol.*, **383**, 713–730.
21. Remes,B., Rische-Grahl,T., Müller,K.M.H., Förstner,K.U., Yu,S.-H., Weber,L., Jäger,A., Peuser,V. and Klug,G. (2017) An RpoHI-dependent response promotes outgrowth after extended stationary phase in the alphaproteobacterium *Rhodobacter sphaeroides*. *J. Bacteriol.*, **199**, e00249-17.
22. Reinkensmeier,J. and Giegerich,R. (2015) Thermodynamic matchers for the construction of the cuckoo RNA family. *RNA Biology*, **12**, 197–207.
23. Mitchell,A.L., Attwood,T.K., Babbitt,P.C., Blum,M., Bork,P., Bridge,A., Brown,S.D., Chang,H.-Y., El-Gebali,S., Fraser,M.I. *et al.* (2018) InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res.*, **47**, D351–D360.
24. Barnett,M.J., Bittner,A.N., Toman,C.J., Oke,V. and Long,S.R. (2012) Dual RpoH sigma factors and transcriptional plasticity in a symbiotic bacterium. *J. Bacteriol.*, **194**, 4983–4994.
25. Robledo,M., Peregrina,A., Millán,V., García-Tomsig,N.I., Torres-Quesada,O., Mateos,P.F., Becker,A. and Jiménez-Zurdo,J.I. (2017) A conserved α-proteobacterial small RNA contributes to osmoadaptation and symbiotic efficiency of rhizobia on legume roots. *Environ. Microbiol.*, **19**, 2661–2680.
26. Remes,B., Berghoff,B.A., Förstner,K.U. and Klug,G. (2014) Role of oxygen and the OxyR protein in the response to iron limitation in *Rhodobacter sphaeroides*. *BMC Genomics* **15**, 794.
27. Baumgardt,K., Šmídová,K., Rahn,H., Lochnit,G., Robledo,M. and Evguenieva-Hackenberg,E. (2016) The stress-related, rhizobial small

RNA RcsR1 destabilizes the autoinducer synthase encoding mRNA *sinI* in *Sinorhizobium meliloti*. *RNA Biol.*, **13**, 486–499.
28. Mank,N.N., Berghoff,B.A., Hermanns,Y.N. and Klug,G. (2012) Regulation of bacterial photosynthesis genes by the small noncoding RNA PcrZ. *PNAS*, **109**, 16306–16311.
29. Klug,G. and Drews,G. (1984) Construction of a gene bank of *Rhodopseudomonas* capsulata using a broad host range DNA cloning system. *Arch. Microbiol.*, **139**, 319–325.
30. Gibson,D.G., Young,L., Chuang,R.-Y., Venter,J.C., Hutchison,C.A. and Smith,H.O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods*, **6**, 343–345.
31. Li,K., Härtig,E. and Klug,G. (2003) Thioredoxin 2 is involved in oxidative stress defence and redox-dependent expression of photosynthesis genes in *Rhodobacter capsulatus*. *Microbiology*, **149**, 419–430.
32. Damm,K., Bach,S., Müller,K.M.H., Klug,G., Burenina,O.Y., Kubareva,E.A., Grünweller,A. and Hartmann,R.K. (2015) Impact of RNA isolation protocols on RNA detection by Northern blotting. *Methods Mol. Biol.*, **1296**, 29–38.
33. Pfeiffer,V., Sittka,A., Tomer,R., Tedin,K., Brinkmann,V. and Vogel,J. (2007) A small non-coding RNA of the invasion gene island (SPI-1) represses outer membrane protein synthesis from the *Salmonella* core genome. *Mol. Microbiol.*, **66**, 1174–1191.
34. Förstner,K.U., Vogel,J. and Sharma,C.M. (2014) READemption-a tool for the computational analysis of deep-sequencing-based transcriptome data. *Bioinformatics (Oxford, England)*, **30**, 3421–3423.
35. Hoffmann,S., Otto,C., Kurtz,S., Sharma,C.M., Khaitovich,P., Vogel,J., Stadler,P.F. and Hackermüller,J. (2009) Fast mapping of short sequences with mismatches, insertions and deletions using index structures. *PLoS Comput. Biol.*, **5**, e1000502.
36. Hoffmann,S., Otto,C., Doose,G., Tanzer,A., Langenberger,D., Christ,S., Kunz,M., Holdt,L.M., Teupser,D., Hackermüller,J. and Stadler,P.F. (2014) A multi-split mapping algorithm for circularRNA, splicing, trans-splicing and fusion detection. *Genome Biol.*, **15**, R34.
37. Anders,S. and Huber,W. (2010) Differential expression analysis for sequence count data. *Genome Biol.*, **11**, R106.
38. Kumar,S., Stecher,G., Li,M., Knyaz,C. and Tamura,K. (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.*, **35**, 1547–1549.
39. Sneath,& Sokal 1973, Sneath,P.H.A and Sokal,R.R. (1973) In: *Numerical Taxonomy: The Principles and Practice of Numerical Classification*. Freeman, San Francisco, p. 573.
40. Felsenstein,J. (1985) Confidence limits on phylogenies: An approach using the bootstrap. *Evolution.* **39**, 783–791.
41. Förstner,K.U., Reuscher,C.M., Haberzettl,K., Weber,L. and Klug,G. (2018) RNase E cleavage shapes the transcriptome of *Rhodobacter sphaeroides* and strongly impacts phototrophic growth, *Life Sci. Alliance*, **1**, e201800080.
42. Green,J.B., Edwards,T.A., Trincao,J., Escalante,C.R., Wharton,R.P. and Aggarwal,A.K. (2002) Crystallization and characterization of Smaug: a novel RNA-binding motif. *Biochem. Biophys. Res. Commun.* **297**, 1085–1088
43. Tadros,W., Goldman,A.L., Babak,T., Menzies,F., Vardy,L., Orr-Weaver,T., Hughes,T.R., Westwood,J.T., Smibert,C.A. and Lipshitz,H.D. (2007) SMAUG is a major regulator of maternal mRNA destabilization in *Drosophila* and its translation is activated by the PAN GU kinase. *Dev. Cell*, **12**, 143–155.
44. Green,J.B., Gardner,C.D., Wharton,R.P. and Aggarwal,A.K. (2003) RNA recognition via the SAM domain of Smaug. *Mol. Cell*, **11**, 1537–1548.
45. Chao,Y., Li,L., Girodat,D., Förstner,K.U., Said,N., Corcoran,C., Śmiga,M., Papenfort,K., Reinhardt,R., Wieden,H.-J. *et al.* (2017) *In vivo* cleavage map Illuminates the central role of RNase E in coding and non-coding RNA pathways. *Mol. Cell*, **65**, 39–51.
46. McIntosh,M., Eisenhardt,K., Remes,B., Konzer,A. and Klug,G. (2019) Adaptation of the Alphaproteobacterium *Rhodobacter sphaeroides* to stationary phase. *Environ. Microbiol.*, **21**, 4425–4445.
47. Zhao,Z., Peng,T., Oh,J.-I., Glaeser,J., Weber,L., Li,Q. and Klug,G. (2019) A response regulator of the OmpR family is part of the regulatory network controlling the oxidative stress response of *Rhodobacter sphaeroides*. *Environ. Microbiol. Rep.*, **11**, 118–128.
48. Weber,L., Thoelken,C., Volk,M., Remes,B., Lechner,M. and Klug,G. (2016) The conserved *Dcw* gene cluster of *R. sphaeroides* is preceded by an uncommonly extended 5′ leader featuring the sRNA UpsM. *PLoS One*, **11**, e0165694.
49. Mank,N., Berghoff,B.A. and Klug,G. (2013) A mixed incoherent feed-forward loop contributes to the regulation of bacterial photosynthesis genes. *PLoS One* **10**, 347–352.
50. Budnick,J.A., Sheehan,L.M., Kang,L., Michalak,P. and Caswell,C.C. (2018) Characterization of three small proteins in *Brucella abortus* linked to fucose utilization. *J. Bacteriol.*, **200**, e00127-18.
51. Kraus,A., Weskamp,M., Zierles,J., Balzer,M., Busch,R., Eisfeld,J., Lambertz,J., Nowaczyk,M.M. and Narberhaus,F. (2020) Arginine-rich small proteins with a domain of unknown function, DUF1127, play a role in phosphate and carbon metabolism of *Agrobacterium tumefaciens*. *J. Bacteriol.*, **202**.
52. Venturini,E., Svensson,S.L., Maaß,S., Gelhausen,R., Eggenhofer,F., Li,L., Cain,A.K., Parkhill,J., Becher,D., Backofen,R., Barquist,L., Sharma,C.M., Westermann,A.J. and Vogel,J. (2020) A global data-driven census of Salmonella small proteins and their potential function in bacteria virulence. *microLife*, **1**, uqaa002.
53. Bandyra,K.J. and Luisi,B.F. (2013) Licensing and due process in the turnover of bacterial RNA. *RNA Biology*, **10**, 627–635.
54. Pinder,B.D. and Smibert,C.A. (2013) microRNA-independent recruitment of Argonaute 1 to nanos mRNA through the Smaug RNA-binding protein. *EMBO Rep.*, **14**, 80–86.
55. Makarova,K.S., Wolf,Y.I., van der Oost,J. and Koonin,E.V. (2009) Prokaryotic homologs of Argonaute proteins are predicted to function as key components of a novel system of defense against mobile genetic elements. *Biol. Direct*, **4**, 29.
56. Olovnikov,I., Chan,K., Sachidanandam,R., Newman,D.K. and Aravin,A.A. (2013) Bacterial argonaute samples the transcriptome to identify foreign DNA. *Mol. Cell*, **51**, 594–605.
57. Miyoshi,T., Ito,K., Murakami,R. and Uchiumi,T. (2016) Structural basis for the recognition of guide RNA and target DNA heteroduplex by Argonaute. *Nat. Commun.*, **7**, 11846.
58. Liu,Y., Esyunina,D., Olovnikov,I., Teplova,M., Kulbachinskiy,A., Aravin,A.A. and Patel,D.J. (2018) Accommodation of helical imperfections in *Rhodobacter sphaeroides* argonaute ternary complexes with guide RNA and target DNA. *Cell Rep.*, **24**, 453–462.
59. Holmqvist,E. and Vogel,J. (2018) RNA-binding proteins in bacteria. *Nat. Rev. Microbiol.*, **16**, 601–615.
60. Quendera,A.P., Seixas,A.F., dos Santos,R.F., Santos,I., Silva,J.P.N., Arraiano,C.M. and Andrade,J.M. (2020) RNA-binding proteins driving the regulatory activity of small non-coding RNAs in bacteria. *Front. Mol. Biosci.*, **7**, 78.
61. Müller,P., Gimpel,M., Wildenhain,T. and Brantl,S. (2019) A new role for CsrA: promotion of complex formation between an sRNA and its mRNA target in *Bacillus subtilis*. *RNA Biol.*, **16**, 972–987.
62. Olejniczak,M. and Storz,G. (2017) ProQ/FinO-domain proteins: another ubiquitous family of RNA matchmakers? *Mol. Microbiol.*, **104**, 905–915.
63. Smirnov,A., Schneider,C., Hör,J. and Vogel,J. (2017) Discovery of new RNA classes and global RNA-binding proteins. *Curr. Opin. Microbiol.*, **39**, 152–160.
64. Wang,N., Yamanaka,K. and Inouye,M. (1999) CspI, the ninth member of the CspA family of *Escherichia coli*, is induced upon cold shock. *J. Bacteriol.*, **181**, 1603–1609.
65. Caballero,C.J., Menendez-Gil,P., Catalan-Moreno,A., Vergara-Irigaray,M., García,B., Segura,V., Irurzun,N., Villanueva,M., Ruiz de los Mozos,I., Solano,C. *et al.* (2018) The regulon of the RNA chaperone CspA and its auto-regulation in *Staphylococcus aureus*. *Nucleic Acids Res.*, **46**, 1345–1361.
66. Dahanukar,A., Walker,J.A. and Wharton,R.P. (1999) Smaug, a novel RNA-binding protein that operates a translational switch in *Drosophila*. *Mol. Cell*, **4**, 209–218.
67. Smiber,C.A., Lie,Y.S., Shillinglaw,W., Henzel,W.J. and Macdonald,P.M. (1999) Smaug, a novel and conserved protein, contributes to repression of nanos mRNA translation *in vitro*. *RNA*, **5**, 1535–1547.
68. Smibert,C.A., Wilson,J.E., Kerr,K. and Macdonald,P.M. (1996) Smaug protein represses translation of unlocalized nanos mRNA in the *Drosophila* embryo. *Genes Dev.*, **10**, 2600–2609
69. Dahanukar,A. and Wharton,R.P. (1996) The Nanos gradient in *Drosophila* embryos is by translational regulation. *Genes Dev.*, **10**, 2610–2620.
70. Kordes,E., Jock,S., Fritsch,J., Bosch,F. and Klug,G. (1994) Cloning of a gene involved in rRNA precursor processing and 23S rRNA cleavage in *Rhodobacter capsulatus*. *J. Bacteriol.*, **176**, 1121–1127.

Chapter 5

A complex network of sigma factors and sRNA StsR regulates stress responses in *R. sphaeroides*

International Journal of
Molecular Sciences

MDPI

Article

A Complex Network of Sigma Factors and sRNA StsR Regulates Stress Responses in *R. sphaeroides*

Katrin M. H. Eisenhardt †, Bernhardt Remes †, Julian Grützner †, Daniel-Timon Spanka, Andreas Jäger and Gabriele Klug *

Institute of Microbiology and Molecular Biology, Justus-Liebig University of Giessen, IFZ, Heinrich-Buff-Ring 26-32, 35392 Giessen, Germany; kmeisenhardt@gmail.com (K.M.H.E.); Bernhard.Remes@gmx.de (B.R.); Julian.Gruetzner@mikro.bio.uni-giessen.de (J.G.); Daniel_Timon.Spanka@mikro.bio.uni-giessen.de (D.-T.S.); Andreas.Jaeger@mikro.bio.uni-giessen.de (A.J.)
* Correspondence: Gabriele.Klug@mikro.bio.uni-giessen.de
† Equal contribution.

Abstract: Adaptation of bacteria to a changing environment is often accompanied by remodeling of the transcriptome. In the facultative phototroph *Rhodobacter sphaeroides* the alternative sigma factors RpoE, RpoHI and RpoHII play an important role in a variety of stress responses, including heat, oxidative stress and nutrient limitation. Photooxidative stress caused by the simultaneous presence of chlorophylls, light and oxygen is a special challenge for phototrophic organisms. Like alternative sigma factors, several non-coding sRNAs have important roles in the defense against photooxidative stress. RNAseq-based transcriptome data pointed to an influence of the stationary phase-induced StsR sRNA on levels of mRNAs and sRNAs with a role in the photooxidative stress response. Furthermore, StsR also affects expression of photosynthesis genes and of genes for regulators of photosynthesis genes. In vivo and in vitro interaction studies revealed that StsR, that is under control of the RpoHI and RpoHII sigma factors, targets *rpoE* mRNA and affects its abundance by altering its stability. RpoE regulates expression of the *rpoHII* gene and, consequently, expression of *stsR*. These data provide new insights into a complex regulatory network of protein regulators and sRNAs involved in defense against photooxidative stress and the regulation of photosynthesis genes.

Keywords: anoxygenic photosynthesis; photooxidative stress; alternative sigma factor; sRNAs; transcriptome; regulatory networks

Citation: Eisenhardt, K.M.H.; Remes, B.; Grützner, J.; Spanka, D.-T.; Jäger, A.; Klug, G. A Complex Network of Sigma Factors and sRNA StsR Regulates Stress Responses in *R. sphaeroides*. *Int. J. Mol. Sci.* **2021**, *22*, 7557. https://doi.org/10.3390/ijms22147557

Academic Editor: Jill Zeilstra-Ryalls

Received: 11 June 2021
Accepted: 9 July 2021
Published: 14 July 2021

1. Introduction

In their natural environment, most bacteria are exposed to changing conditions that may limit survival and are considered stresses. Molecular mechanisms that allow bacteria to adapt to and to survive such stress situations have been known for decades. Nevertheless, new players in bacterial regulation, such as sRNAs, small proteins, or small signaling molecules, have been identified over the years and new models of regulation have emerged that are far more complex than anticipated.

It is widely accepted that adaptation of bacteria to stress conditions occurs mostly at transcriptional level, although in some cases a strong modulation of the proteome is observed that is not accompanied by similar strong changes at the transcriptome level (e.g., adaptation to the stationary phase of *R. sphaeroides* [1]). Transcriptional regulation is often controlled by sigma factors that recognize different promoter sequences and recruit the RNA polymerase to the promoters (reviewed in [2–4]). In addition, many DNA binding proteins other than sigma factors are known to influence promoter activities. Several pathways that activate sigma factors or other transcription factors in response to environmental signals have been unraveled in the past (e.g., [5–7]). sRNAs make important contributions to post-transcriptional regulation. Among several mechanisms of action, they often influence the translation and/or stability of their target mRNA (e.g., [8,9]).

Some sRNAs are controlled by alternative sigma factors and are consequently transcribed in response to external signals (e.g., [10,11]). Here we demonstrate that the sRNA StsR interacts with the mRNA of the RpoE sigma factor in *R. sphaeroides* and thus is part of a regulatory network affecting photooxidative stress defense and the formation of photosynthetic complexes (Figure 1).

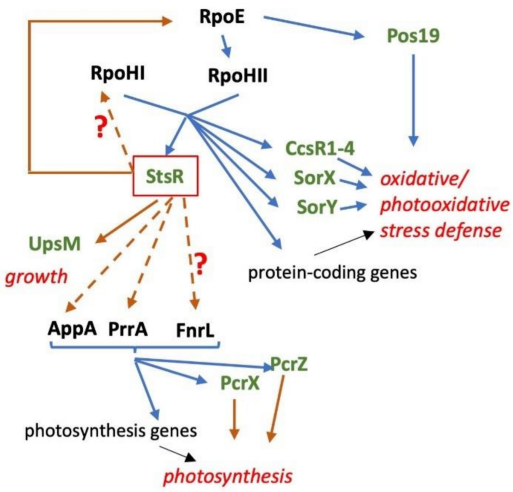


Figure 1. Schematic overview of the regulatory network of sRNAs (green) and proteins (black) affecting the photooxidative stress response and expression of photosynthesis genes. Blue arrows indicate activation, brown arrows repression.

R. sphaeroides is an Alphaproteobacterium that performs aerobic respiration as long as sufficient oxygen is present. If oxygen tension drops, photosynthetic complexes are assembled into intracytoplasmic membranes and allow the use of light for anoxygenic photosynthesis. If no light is present, anaerobic respiration or fermentation can generate ATP. To avoid photooxidative stress by the production of singlet oxygen, the formation of photosynthetic complexes is controlled by oxygen tension and light. Important factors in this regulation are the two component system proteins PrrA (response regulator) and PrrB (sensor kinase) that senses the electron flow through cbb3 cytochrome oxidase [12,13], the transcriptional repressor PpsR and the antirepressor proteins AppA, that senses oxygen through heme and light through the BLUF domain [14–18], and PpaA that uses cobalamine as a light sensor [19,20]. Furthermore, FnrL is an oxygen-responsive regulator of some photosynthesis genes [21–23]. In addition to transcription factors, sRNAs influence the expression of photosynthesis genes by having modulating effects as part of regulatory feed-forward loops [24–27]. The sRNAs, PcrX and asPcrL, interact with parts of the polycistronic *puf* mRNA that encodes the pigment-binding proteins of the reaction center (RC) and light harvesting (LH) complexes and the assembly factor PufX. They affect the stability of *puf* mRNA segments and, consequently, the stoichiometry of RC/LHI and LHII complexes [26,27]. LHII proteins are encoded by the *puc* mRNAs that are not affected by PrcX or asPcrL. PcrZ negatively affects its targets *puc2A* and *bchN* and thereby counteracts and balances the strong induction of photosynthesis genes upon a drop in oxygen [24,25]. Transcription of these sRNAs is controlled by the same proteins (PrrA, PpsR-AppA, FnrL) as expression of their targets (Figure 1).

The control of photosynthesis gene expression in response to external factors should avoid photooxidative stress. However, sudden changes in oxygen tension and/or light intensity after formation of photosynthetic complexes can take place and lead to photooxidative stress, mostly through the production of the reactive singlet oxygen [28]. A main role in the photooxidative stress response in *R. sphaeroides* was attributed to the

alternative sigma factors RpoE, RpoHI and RpoHII [29–31]. Under nonstress conditions, RpoE is inactivated by its antisigma factor ChrR [32]. Under oxidative stress, the proteases DegS and RseP promote degradation of ChrR [33]. The proteins RSP_1090 and RSP_1091 promote this process in the presence of singlet oxygen but not in the response to organic peroxides [33]. RpoE targets a relatively small number of genes, including its own gene, the gene for a photolyase, the gene for the sRNA Pos19 and the gene for the RpoHII sigma factor (Figure 1). RpoHII controls a rather large regulon including genes with functions in singlet oxygen quenching, methylglyoxal detoxification, methionine sulfoxide reduction, the GSH-dependent defense and quinone pool retention [30,31]. The RpoHII regulon has considerable overlap with the RpoHI regulon [30,31]. While *rpoHII* mRNA levels show a much stronger increase upon singlet oxygen exposure than after heat shock, *rpoHI* mRNA shows a stronger increase after heat shock [29]. Both RpoH sigma factors also play an important role in the stationary phase and are required for fast outgrowth from the stationary phase [34]. Consensus binding sequences for RpoHI, and RpoHII have been identified [27,28].

RpoHI and RpoHII do not only regulate expression of protein-coding genes but also regulate expression of sRNAs with a role in the stress defense in *R. sphaeroides* [15,32,35–37] (Figure 1). StsR (formerly RSs0827) was first described as an sRNA induced upon iron starvation [36], and was later found to be the most highly induced RNA in late stationary phase [1,34]. StsR is under control of RpoHI/RpoHII [37] and, therefore, induced by multiple stress factors such as heat and oxidative stress. This sRNA was named StsR (sRNA targeting sRNA) due to its binding to the sRNA UpsM. [38]. UpsM is derived from the 5′ UTR of the *dcw* (cell division and cell wall) genes [38], and binding of StsR to UpsM and to the *dcw* 5′ UTR affects *dcw* gene expression (and consequently cell division) in trans and in cis [37]. Here, we demonstrate that StsR also affects several mRNAs for regulators of photosynthesis genes and photosynthesis gene expression, as well as expression of sRNAs with a role in photooxidative stress in the stationary phase; we identify the *rpoE* mRNA as one target of StsR.

2. Results

2.1. Overview on the Effect of StsR on Expression of Protein-Coding Genes

The *R. sphaeroides* sRNA StsR shows very low abundance in the exponential phase but is highly abundant in the stationary phase [37]. To evaluate the effect of StsR on the transcriptome, we performed RNAseq analysis with wild type cells of *R. sphaeroides* and with a mutant (ΔStsR) lacking the *stsR* gene. For each strain and condition, triplicates were sequenced, each stemming from a mixture of three independent cultures.

Principal component analysis (PCA) revealed very good reproducibility within the replicates of every group (Figure S1A). Using DESeq2 [39], the transposed count matrix was used to compute the Euclidean sample-to-sample distances and to perform a hierarchical clustering (Figure S1B). A heatmap revealed strong similarities between the samples from both strains, which were taken during the exponential growth phase. In contrast, the transcriptomes of the wild type and the StsR mutant strain differed greatly during the stationary phase. Within these clusters, the three samples belonging to one strain formed distinct subclusters. The growth phase-dependent differences in the cellular RNA composition were visible in volcano plots: Only few transcripts varied between the strains during the exponential phase, but during the stationary phase more than a third of all transcripts were classified as differentially expressed (Figure S1C,D).

We considered all genes as differentially expressed when the $\log_2$-fold change between the two strains was ≥ 1.0 or ≤ -1.0 and the adjusted p-value was ≤ 0.05 (Supplementary Table S1). Although StsR showed very low abundance in the exponential phase, 26 protein-coding genes showed higher expression in the mutant. Among those were several genes for flagellar synthesis and chemotaxis. Seventeen protein-coding genes showed lower expression in the mutant, including *znuB* and *znuC* (*znuA* missed the cut-off for p-value) for a zinc transporter and *pufK*, which is part of the photosynthesis gene cluster. Seventy two

hours after inoculation to an OD_{660nm} of 0.2, *Rhodobacter* cells were in the late stationary phase and StsR was highly expressed [37]. Under these conditions 618 protein-coding genes showed higher expression in the mutant strain (Supplementary Table S1). Among those were *bchI* (*bch* genes are required for bacteriochlorophyll synthesis) and *tspO* from the photosynthesis gene cluster. TspO is an outer membrane protein that negatively regulates expression of photosynthesis genes in response to oxygen by controlling the efflux of porphyrin intermediates [40]. The *fnrL* gene for a regulator of photosynthesis genes also showed higher expression in the mutant. Furthermore, the *rpoHI*, *rpoHII*, *rpoE* genes, and RSP_3095 for another sigma factor, all showed 4.6-5.7 times greater expression in the *stsR* mutant (Table 1).

Table 1. log_2-fold changes in read counts determined by RNAseq within a strain between different growth phases or between wild type (WT) and the *stsR* mutant in the same growth phase, as calculated by DEseq2 [39]. Brackets indicate that the adjusted *p*-value is >0.05. Growth curves for the two strains are shown in Figure S2 and the time points of sampling are indicated.

Gene		Log_2-Fold WT Station./ Expon. Phase	Log_2-Fold Mutant Station./ Expon. Phase	Log_2-Fold Mutant/WT Expon. Phase	Log_2-Fold Mutant/WT Station. Phase
Photosynth. genes					
pufX	RSP_0255	−2.54	−3.79	(−0.14)	−1.26
pufM	RSP_0256	−2.65	−3.23	(−0.43)	−0.58
bchY	RSP_0261	−2.02	−3.57	(0.04)	−1.56
bchX	RSP_0262	−2.24	−3.14	0.25	−0.90
tspO	RSP_0269	−1.67	(−0.34)	(−0.03)	1.33
bchI	RSP_0273	−2.35	−1.26	0.28	1.10
bchJ	RSP_0280	−2.40	−4.00	−0.69	−1.67
bchE	RSP_0281	−1.66	−2.51	(−0.01)	−0.85
bchH	RSP_0287	−1.41	−2.86	(−0.07)	−1.00
bchL	RSP_0288	−1.18	−2.03	(0.15)	−0.85
hemN	RSP_0317	1.09	−2.70	(0.29)	−1.90
hemZ	RSP_0699	−1.12	−3.02	(0.05)	−1.95
Genes for regulatory proteins					
fnrL	RSP_0698	−0.82	1.00	(0.03)	1.80
prrA	RSP_1518	−1.32	−0.52	0.49	0.86
appA	RSP_1565	−0.80	−0.06	(−0.04)	0.70
Genes for alternative sigma factors/anti-sigma factors					
rpoHII	RSP_0601	1.10	3.33	(−0.04)	2.22
rpoE	RSP_1092	(−0.15)	2.89	(−0.15)	2.42
chrR	RSP_1093	−1.85	2.23	(−0.23)	(0.15)
rpoHI	RSP_2410	3.69	6.20	(0.06)	2.52
	RSP_3095	5.04	7.23	(0.12)	2.19
	RSP_3094	5.32	6.61	(0.32)	1.50

In the late stationary phase, 762 protein-coding genes showed lower expression in the mutant (Supplementary Table S1). Among them were several *bch* genes, *pufX* that is required for the assembly of the reaction center (RC) and light-harvesting I (LHI) complexes, and *hemZ* and *hemN* (for oxygen-independent coproporphyrinogen III oxidases that are required for synthesis of protoporphyrin IX).

2.2. Effects of StsR on Expression of Photosynthesis Genes and of Genes for Regulators of Photosynthesis Genes

Table 1 shows the expression changes for photosynthesis genes, for genes encoding regulators of photosynthesis, and for genes encoding sigma factors involved in stress responses that show $\log_2$-fold change of $\geq$0.5 or ≤ -0.5 between the two strains in the stationary phase (adj. *p*-value $\leq$0.05, otherwise numbers are in brackets, and genes with read counts <20 in both strains were excluded). In agreement with the low levels of StsR in the exponential phase, all these genes showed similar expression in the exponential phase in the wild type and mutant. Most photosynthesis genes showed a strong decrease in expression in the stationary phase in the wild type, and an even stronger decrease in the mutant strain. As a result, the mutant showed lower expression in the stationary phase, but transcript levels in both strains were very low compared to exponential growth phase, as shown for *bchJE* in Figure 2A.

A different effect of StsR was observed for *tspO,* and *bchI* (Table 1), which are not part of the same operon (results for *bchI* shown in Figure 2B). Expression levels were more decreased in the stationary phase in the wild type than in the *stsR* mutant, resulting in higher levels of *bchI* and *tspO* mRNAs in the mutant in the stationary phase. The expression pattern of *bchI* differed from that of *crtA*, *bchD* and *bchO*, although all these genes are in the same operon (these genes are not listed in Table 1 due to less than 20 reads in the stationary phase). Higher expression levels in the stationary phase in the mutant were only observed for *bchI.* This strongly suggests that StsR does not affect transcription of the operon but acts at the post-transcriptional level, like most sRNAs.

Transcriptional start sites in the *R. sphaeroides* transcriptome have previously been identified by differential RNAseq [34] (2017; GEO accession number GSE71844). The *tspO* gene is transcribed from a RpoHII-dependent promoter [31]. Since StsR also influences expression of this sigma factor (see below), its effect on *tspO* mRNA levels is likely indirect through altered levels of RpoHII in the mutant.

Altered expression of photosynthesis genes should also affect formation of the photosynthetic apparatus in the *stsR* mutant. Spectral analysis confirmed this assumption: the *stsR* mutant accumulated less photosynthetic complexes than the wild type under phototrophic conditions in the stationary phase (Figure S3).

Our data revealed that StsR also affects expression of some genes for important regulators of photosynthesis genes in the stationary phase (Table 1). This is the case for the *fnrL*, *prrA*, and *appA* genes, which all showed higher expression in the *stsR* mutant in the stationary phase. As observed for photosynthesis genes, expression was similar in the mutant and wild type in the exponential phase and dropped in the stationary phase in the wild type. In contrast to the results for most photosynthesis genes, expression in the stationary phase was higher in the mutant and reached similar levels as in the exponential phase (results for *prrA* shown in Figure 2C). Since the action of StsR on regulatory proteins impacts many other genes and, therefore, is of special importance, we confirmed the RNAseq data by qRT PCR for some selected regulator genes (Figure 3). These data confirmed the higher expression levels of *appA* and *prrA* in the mutant compared to the wild type in the stationary phase. The factors for expression changes are often higher in real time data than in RNAseq due to the high sensitivity of the PCR-based approach. Especially for *prrA,* the change observed by real time RT PCR was much higher. While the DEseq [39] analysis calculates the expression levels based on the read counts for the whole gene, only a small part of the mRNA is amplified in the real time analysis, which can account for such big differences.

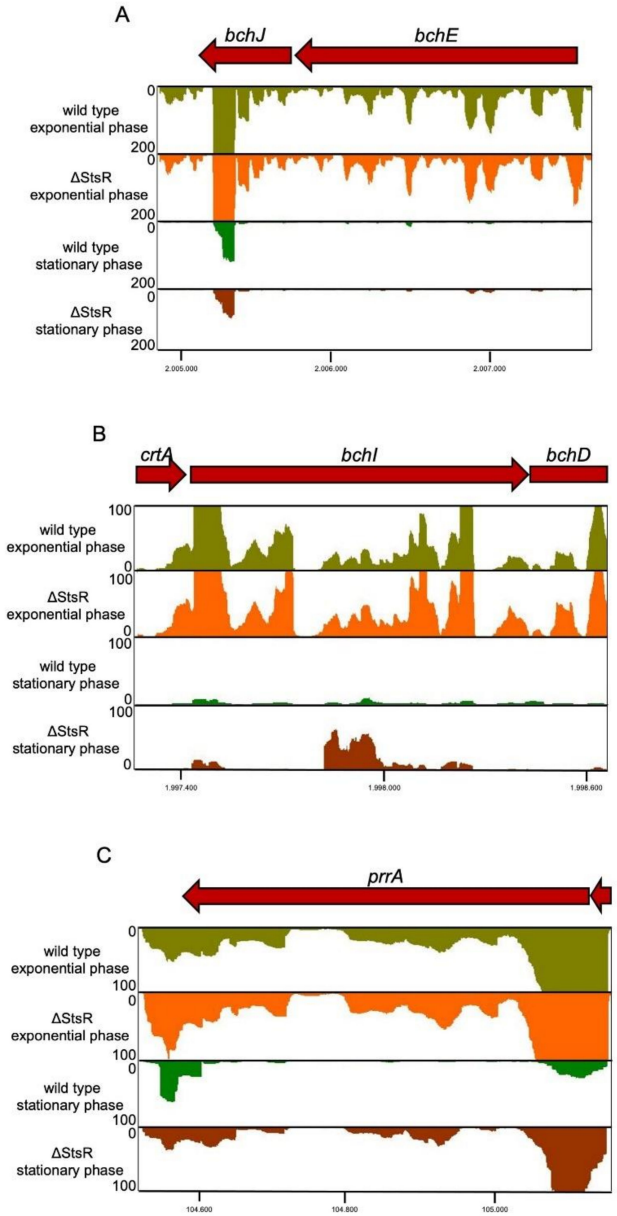


Figure 2. Effect of StsR on expression of selected photosynthesis genes and of the gene for a regulator of photosynthesis genes. Read numbers from RNAseq visualized by the Integrated Genome browser are shown for (**A**) *bchJ* and *bchE* genes, (**B**) the *bchI* gene required for bacteriochlorophyll synthesis and (**C**) the *prrA* gene encoding the response regulator of the PrrB/PrrA two component system. Reads are shown for RNA isolated from the wild type or a mutant lacking StsR (ΔStsR) in the exponential or stationary phase 72 h after inoculation. The read counts within one panel were all normalized to the same scale, as indicated.

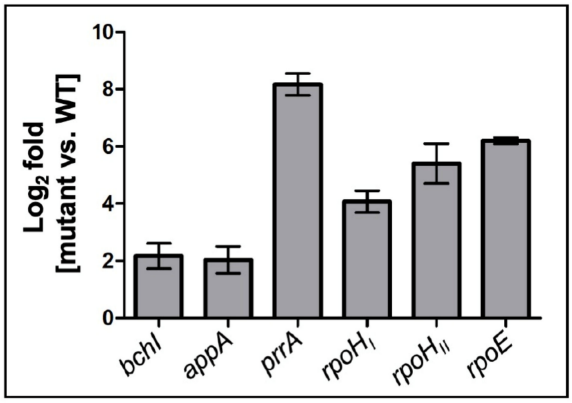


Figure 3. Ratio of expression (log_2-fold change) of selected genes as determined by real time RT PCR in the ΔStsR mutant compared to the wild type. An in vitro transcript of *sinI* RNA, an external spike-in RNA of known sequence and quantity, was used for normalization.

2.3. Effect of StsR on mRNAs for Alternative Sigma Factors

StsR is strongly expressed in the stationary phase in conditions known to also induce expression of the genes for the alternative sigma factors RpoHI, RpoHII and RSP_3095 in *R. sphaeroides* [1,34,41]. The mRNA levels for all these sigma factors were increased in the stationary phase in the wild type, and even more so in the mutant (e.g., a 150-fold higher level in the mutant in stationary phase than in the exponential phase for RSP_3095) (Table 1). This was not the case for *rpoE* mRNA in the wild type, but in the *stsR* mutant (Figure 4A). Figure 4B,C also shows expression levels for the *rpoHI* and *rpoHII* genes. For all these sigma factor mRNAs, the highest expression was observed in the stationary phase in the mutant, implicating that StsR counteracts high expression in the stationary phase. Real time PCR quantification of *rpoE*, *rpoHI*, and *rpoHII* mRNAs (Figure 3) confirmed their higher levels in the mutant in the stationary phase.

Table 1 also includes the data for RSP_1093, although it did not fulfil the criteria of fold-change and p-value for the difference between the two strains in the stationary phase. RSP_1093 encodes ChrR, the antisigma factor to RpoE [32]. It is noteworthy that the ratio of *chrR* mRNA levels between the two strains did not change to the same extent for *rpoE*, although both genes are transcribed from the same promoter. This strongly indicates additional regulation at the post-transcriptional level. RSP_3095 is cotranscribed with RSP_3094, most likely encoding the antisigma factor to the RSP_3095 protein. The expression pattern for both genes was very similar (Table 1). Expression levels of other alternative sigma factors (4 RpoN sigma factors with a role in nitrogen metabolism [42] and a second RpoE with unknown function) were similar for both strains in the stationary phase. Interestingly, the mRNA level for the house-keeping sigma factor (RpoD) was decreased in the stationary phase in the mutant compared to the wild type (log_2-fold change: −0.9) (data not shown).

These data demonstrate that StsR affects expression of many genes, especially in the stationary phase. However, these data cannot discriminate between direct effects by binding to target RNAs or indirect effects. The impact of StsR on expression of alternative sigma factors and regulators of photosynthesis genes suggests that many effects on the transcriptome may be indirect.

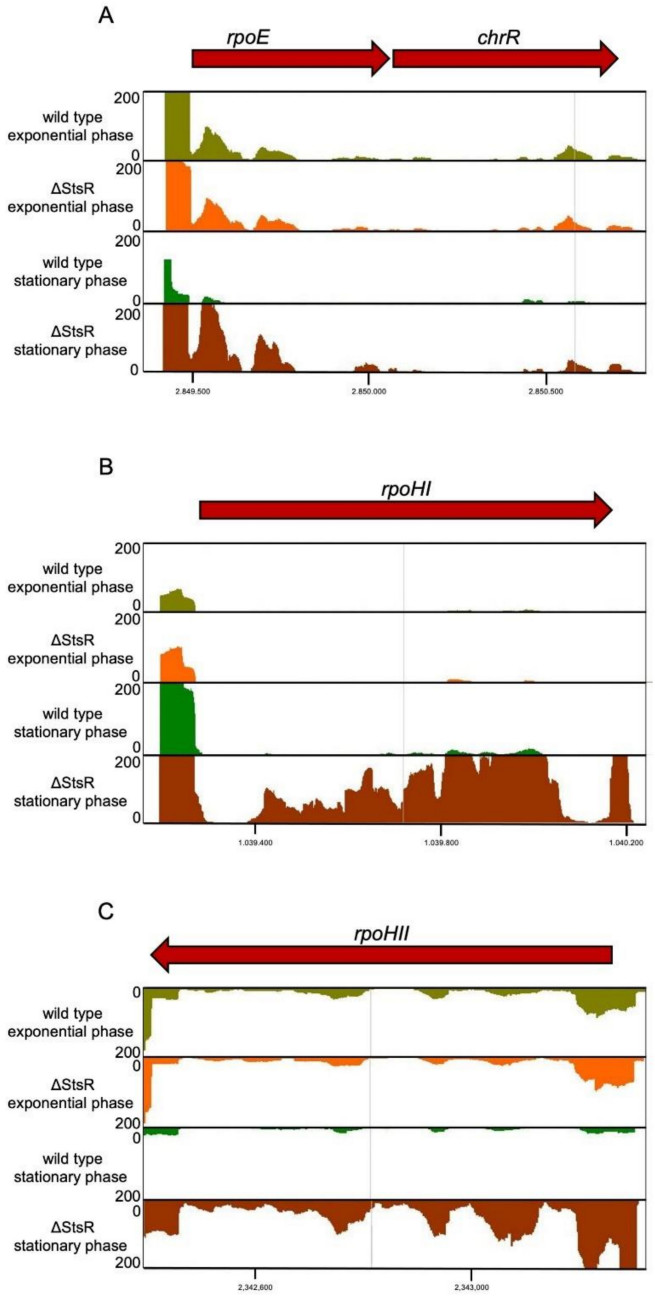


Figure 4. Effect of StsR on expression of selected genes for alternative sigma factors. Normalized read numbers from RNAseq visualized by the Integrated Genome browser are shown for (**A**) *rpoE* and *chrR* genes encoding a sigma factor and its antisigma factor, (**B**) the *rpoHI* gene and (**C**) the *rpoHII* gene encoding alternative sigma factors with an important role in stress responses. Reads are shown for RNA isolated from the wild type or a mutant lacking StsR (ΔStsR) in the exponential or stationary phase 72 h after inoculation. The read counts within one panel were all normalized to the same scale, as indicated.

2.4. Effect of StsR on Expression of sRNAs with a Role in Stress Responses or Photosynthesis Gene Expression

Two trans-acting sRNAs (PcrX and PcrZ) are known to affect photosynthesis gene expression [24,26]. As part of incoherent feed-forward loops, they balance the induction of photosynthesis genes upon reduction of oxygen tension. Several sRNAs are induced in response to various stress conditions and were identified as important regulators in the photooxidative stress response of *R. sphaeroides* [43]. The four homologous sRNAs, CcsR1-4, target the mRNA for the FlhR regulator and affect the glutathione pool and the pyruvate dehydrogenase complex [35]. They are cotranscribed with the gene for the small RNA-binding protein CcaF1, that influences maturation and stability of several sRNAs and/or mRNAs [41]. Another sRNA that influences the glutathione pool, Pos19, affects the abundance of numerous mRNAs involved in sulfur-metabolism [44]. SorX targets the mRNA for the subunit of a spermidine transporter [45]. SorY reduces the metabolic flux through the tricarboxylic acid cycle by targeting the mRNA for a malate transporter [46].

These sRNAs are not included in Table 1, since quantification of these short and mostly highly abundant sRNAs by DEseq is often problematic and generates high p-values. We therefore performed northern blots to examine the effect of StsR on the expression levels of these important RNA regulators (Figure 5). For all tested sRNAs, levels in the wild type and mutant were similar in the exponential phase.

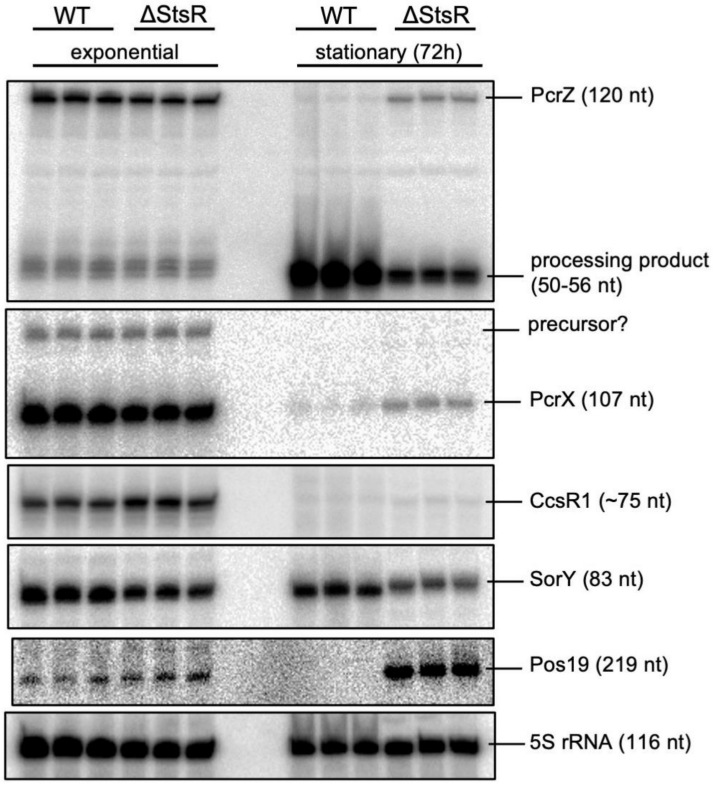


Figure 5. Northern blot of the sRNAs with a role in regulation of photosynthesis genes or in the oxidative stress response in the wild type (WT) and *stsR* mutant in the exponential or stationary phase. For each strain, RNA from three independent cultures was loaded. 8 µg of total RNA were applied to each lane, and 5S rRNA served as loading control. The identical membrane was subsequently hybridized to the specific probes.

PcrZ was previously shown to undergo growth phase-dependent processing: a shorter, stable segment derived from the 5´ end of the PcrZ transcript accumulates in stationary phase [25]. Northern blot analysis revealed that processing of PcrZ was impeded in absence of StsR (Figure 5). It is known that StsR interacts with the UpsM sRNA and promotes its RNase E-dependent cleavage [37]. An interaction between StsR and PcrZ was, however, not predicted by IntaRNA, suggesting a different effect of StsR on PcrZ processing, which may also be indirect.

PcrX and CcsR1-4 are derived from the 3′ UTRs of genes by processing of a precursor transcript. In both cases RNase E has an important role in maturation of these sRNAs [23,33,42]. The amount of PcrX was clearly decreased in the stationary phase, and a stronger decrease was observed for the wild type. StsR had no strong effect on CcsR1 levels, which were significantly lower in both strains in the stationary phase. No processing events are involved in the generation of Pos19 and SorY that are directly transcribed from their genes, and do not undergo further processing [23,34]. Pos19 levels were strongly increased in the mutant but were not detected in the wild type in the stationary phase. SorY had slightly lower levels in the mutant than in the wild type in both growth phases.

Our data demonstrate that StsR can have very different effects on the abundance of individual sRNAs. This effect may also be indirect, mediated by regulatory proteins or other sRNAs. Considering the important functions of the tested sRNAs in regulation, StsR indirectly affects the targets of PcrZ, PcrX, and Pos19 (Figure 1).

2.5. StsR Targets rpoE mRNA and Affects Its Stability

In order to get an idea on putative targets of StsR, we applied IntaRNA [47], a bioinformatic tool for the prediction of RNA-RNA interaction. RpoE mRNA was suggested to be a target of StsR, and an energy value of −18 kJ was calculated for the interaction (Figure 6A; the numbering for *rpoE* mRNA gives the position of nucleotides in relation to the translational start (GUG), numbering for StsR refers to the nucleotide position within the 72 nt long StsR). To verify this interaction in vivo, we compared activity of a *rpoE-lacZ* fusion in the wild type and in the mutant strain. A fragment from position −101 to +111 in relation to the start codon of *rpoE* (not including the promoter of *rpoE*) was cloned into pPHU4352 [24]. In the resulting plasmid (pPHU_1092) the *rpoE* sequence was transcribed from the 16S promoter and translationally fused to *lacZ*. As seen in Figure 6B, introduction of a second plasmid that overexpresses StsR (pBBR_StsR) led to reduced ß-galactosidase activity, while monitoring reporter gene activity in ΔStsR resulted in increased ß-galactosidase activity. This strongly supports the view that StsR reduces expression of RpoE, which is in agreement with the RNAseq data shown in Figure 4A.

To further validate these results, we tested in vitro interaction of the radio-labelled 72 nt StsR and a 153 nt in vitro transcript spanning positions −19 to +134 relative to the *rpoE* translational start (the transcriptional start site for RpoE is at −96 relative to the translational start). As shown in Figure 6C, addition of increasing amounts of the *rpoE* transcript (150–15,000 fmol) resulted in retardation of the radiolabeled StsR (150 fmol), providing further support for direct interaction between these two RNAs.

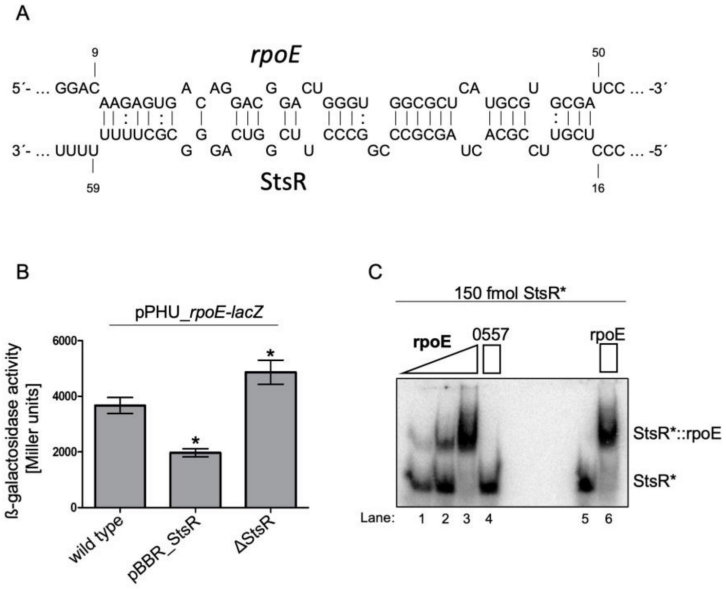


Figure 6. StsR interacts with the *rpoE* mRNA. (**A**) Seed region for the interaction between *rpoE* and StsR as predicted by the IntaRNA tool. (**B**) *lacZ*-based in vivo reporter assay. All strains contain a plasmid with a *rpoE-lacZ* fusion. While over-expression of StsR reduces *rpoE-lacZ* activity, lack of StsR leads to higher expression. (**C**) Gel retardation assay showing the interaction of *rpoE* and StsR in vitro. Radio-labelled StsR (150 fmol) in vitro transcript was incubated with increasing amounts of a 150 nt *rpoE* in vitro transcript (150–15,000 fmol lanes 1–3). A negative control StsR was incubated together with a 100-fold molar excess of an RSP_0557 [23] in vitro transcript (lane 4). As further controls, the StsR transcript was loaded alone (lane 5) or after de and renaturation together with a 100-fold molar excess of the *rpoE* transcript. * $p < 0.5$.

Most sRNAs affect translation of their target RNAs by binding close to the translational start, or influence the stability of the target mRNA, or both (reviewed in [8]). The reporter assay shown in Figure 6B cannot discriminate between these mechanisms, since both lead to reduced ß-galactosidase activity. Binding of the sRNA can either stabilize the target by protecting single-stranded regions from cleavage by RNases attacking single stranded regions (e.g., RNase E) or can promote degradation by generating targets for double strand-specific RNases (e.g., RNase III) [8]. Many sRNAs, among them StsR, are associated with the RNA chaperon Hfq [48] that can stabilize the sRNA-mRNA interaction, but can also recruit RNase E and promote destabilization of the target mRNA [49]. As shown in Figure 6A, the seed region for the interaction between StsR and *rpoE* mRNA (purple) is close to the translational start site, starting 10 nt downstream of the GTG. In this region, and just downstream of the interaction site, several RNase E cleavage sites were mapped [50]. To test whether StsR pairing influences *rpoE* mRNA stability, we determined the half-life of *rpoE* mRNA in the wild type and in the *stsR* mutant. Cultures were grown to late exponential phase and rifampicin was added to stop further initiation of transcription. At short intervals after addition of rifampicin, samples were collected for RNA isolation and *rpoE* mRNA was quantified by real time RT PCR. Figure 7 shows that in the strain lacking StsR *rpoE*, the half-life was about doubled compared to the wild type (1.0 min versus 0.55 min). These data strongly suggest that StsR reduces the half-life of *rpoE* mRNA and subsequently its level.

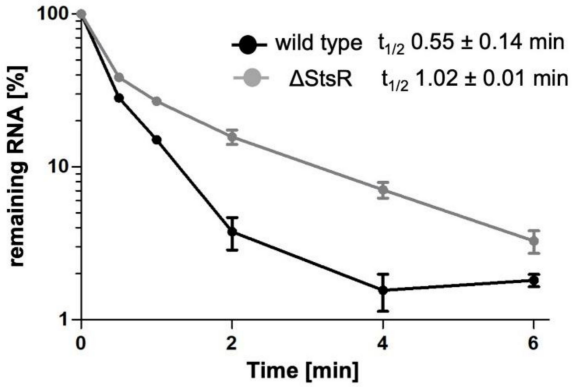


Figure 7. StsR decreases the half-life of *rpoE* mRNA. Rifampicin was added to cultures in the exponential growth phase to inhibit initiation of transcription. RNA was isolated at several time points and *rpoE* mRNA levels were quantified by real time RT-PCR and plotted against time. The values represent the average from three independent cultures, and the standard deviation is indicated. Lack of StsR increased the *rpoE* half-life about two-fold.

3. Discussion

Initiation of transcription is a major check point of regulation of prokaryotes in adapting to their environment, and mechanisms of transcriptional regulation have been studied for decades. Many important protein regulators and regulatory DNA elements have been identified and characterized. Today, the involvement of RNA regulators in adaptation is well recognized, and different mechanisms of this regulation, mostly acting on post-transcriptional levels, have been unraveled [8,9]. Special challenges for facultative phototrophic bacteria are to regulate the formation of the photosynthetic apparatus in order to avoid photooxidative stress and, if this is not possible, to defend against photooxidative stress. This is achieved by a complex network consisting of proteins that regulate transcription, and of sRNAs acting at the post-transcriptional level (Figure 1). The alternative sigma factors RpoE, RpoHI and RpoHII not only control transcription of genes for proteins with a function in stress responses, but also transcription of the sRNAs Pos19, CcsR1-4, SorX and SorY with an important contribution to these responses.

This work attributes a central role to StsR in this network. RpoHI and RpoHII increase transcription of the *stsR* gene in response to stress, while StsR destabilizes *rpoE* mRNA. Thus, RpoE, RpoHII, and StsR form a negative feed-back loop consisting of protein regulators and sRNA. Such regulatory loops have been reported for other bacteria (e.g., [51–54]). Regulation in the RpoE-RpoHII-StsR loop is based on different mechanisms. RpoE increases transcription of the *rpoHII* gene, while RpoHII increases transcription of *stsR*. The negative effect of StsR on *rpoE* mRNA levels is due, at least in part, to faster turn-over in the presence of StsR. We cannot exclude an additional effect on translation and, consequently, on RpoE protein levels. Since lack of StsR stabilizes the *rpoE* transcript, protection against RNase E cleavage by base pairing of StsR can be excluded. Instead, StsR promotes decay of *rpoE* mRNA. We have recently shown that base pairing of StsR to the sRNA UpsM promotes its cleavage by RNase E. This is due to a structural change of UpsM upon binding of StsR that gives access for RNase E to a previously double-stranded region [37]. A similar mechanism may apply to the effect of StsR on *rpoE* stability.

Our data reveal that *rpoHI* mRNA levels are also increased in the *stsR* mutant (Figures 3 and 4B). The *rpoHI* promoter is not under control of RpoE [31] and no information is available with regard to its activation by stress. Direct interaction of StsR and *rpoHI* was not supported by IntaRNA prediction.

By affecting levels of *rpoE* and *rpoHII* mRNAs, StsR indirectly affects levels of genes that are part of the RpoE and RpoHII regulons, such as the sRNAs Pos19, CcsR1-4, SorX,

and SorY (Figure 1). Considering the different expression patterns of these sRNAs, it is obvious that the effect of StsR cannot only be mediated via RpoHII. Lack of StsR increased *rpoE* and *rpoHII* mRNA levels, as seen in Figures 3 and 4. Higher levels of RpoE and RpoHII should increase expression of genes that are controlled by these proteins, which is in agreement with increased levels of Pos19, SorX and SorY in the *stsR* mutant (Figure 5). However, levels of CcsR, PcrZ and PcrX were decreased in the *stsR* mutant (Figure 5). The CcsR1-4 RNAs are transcribed together with an upstream gene, *ccaF1* (RSP_6037). CcaF1 was recently identified as a small RNA-binding protein involved in RNA maturation and turn-over. Increased amounts of CcaF1 interfere with the maturation of the CcsR RNAs from the *ccaF1*-CcsR1-4 precursor transcript and reduce the CcsR1 half-life [41]. These effects can account for reduced CcsR levels even when the *ccaF1*-CcsR promoter is activated.

PcrZ and PcrX are not under control of RpoE or RpoHI/HII but are regulated by PrrA, AppA and FnrL [24,26]. Lack of StsR results in increased levels of mRNAs for these regulators in the stationary phase, and in reduced levels of PcrZ and PcrX. PrrA and FnrL are activators of gene expression, AppA is an antirepressor of PpsR and, consequently, indirectly activates gene expression. We cannot exclude that StsR affects PcrZ and PcrX levels through other mechanisms.

StsR also affects photosynthesis gene expression. Most photosynthesis genes showed lower expression in the mutant in the stationary phase than in the wild type, while the opposite effect of StsR was observed for *bchI* and *tspO*. Regardless of the different expression levels of most photosynthesis genes in the two strains (up to four-fold), these differences may not be of physiological relevance, since expression in the stationary phase was very low in both strains. The effect of StsR on *bchI*, *prrA*, *appA*, and *fnrL* are likely to have larger impact, since only in the absence of StsR were significant amounts of the mRNAs observed. Thus, StsR has an important role in reducing expression of these genes in the stationary phase.

At present, it is not possible to explain the effect of StsR on expression of the individual genes. Although the function of AppA/PpsR, PrrB/PrrA and FnrL have been addressed in numerous studies (rev. [55]), we are far from understanding this complex regulatory network for photosynthesis gene expression. Elucidation of the underlying mechanisms is not straightforward: mutation or overexpression of one gene will at the same time affect other regulators, and the regulatory loop may compensate for the effects caused by the altered level of a single component of the loop. In vitro experiments with only two components may give some more insights into the mechanisms of regulation. For example, identifying direct targets of StsR, and investigating its effects on its targets as shown here for RpoE, may provide helpful information. We do not know, at present, whether some mRNAs for photosynthesis genes, for protein regulators, or for PcrZ and PcrX, are directly targeted by StsR. Nevertheless, in vitro experiments cannot completely mimic the in vivo situation and, in most cases, cannot include the effect of changing environments. Indeed, the complexity of the regulatory network for regulation of the photooxidative stress response, and for photosynthesis gene expression, is even greater, as outlined in Figure 1. StsR can bind the RNA chaperone Hfq [37] that affects many cellular processes. Deletion of Hfq in *R. sphaeroides* has pleiotropic effects, including reduced pigmentation and altered photooxidative stress response. More than 70% of the Hfq-bound sRNAs are affected by singlet oxygen [56]. In the exponential phase, about 60% of the RNA-bound Hfq protein is bound to UpsM (formerly RSs0682) [56]. UpsM is highly abundant in the exponential phase but strongly decreases in stationary phase due to the action of StsR that promotes degradation of UpsM by RNase E. StsR reaches high levels in the stationary phase similar to UpsM in the exponential phase, and also binds Hfq [37]. The sRNAs CcsR, SorY, SorY and Pos19 are known to interact with Hfq, and the function of PcrX is affected by Hfq [26,47,57]. Competition among targets over Hfq binding plays an important factor in regulation (e.g., [57–63]).

StsR has strong effects on the expression of other genes, mostly in the stationary phase, which is rarely included in studies analyzing bacterial gene expression. In natural habitats, however, bacteria are in the stationary phase for most of the time, so that regulation at this state should not be ignored.

Taken together, our study demonstrates that regulation of photosynthesis genes and of the oxidative stress response in *R. sphaeroides* is far more complex than was anticipated in the past. Complex regulatory loops complicate the elucidation of the role of individual components in regulation. Most likely, it will take a lot more studies to know all components of these complex networks, to understand their interaction and the process of adaptation to different growth conditions.

4. Materials and Methods

4.1. Bacterial Strains, Plasmids and Growth Conditions

The wild type *R. sphaeroides* 2.4.1 [64] was used for this study. Construction of the mutant strain lacking StsR (2.4.1 Δ*S*tsR), and of plasmid pBBR_StsR for overexpression of StsR is described in [37]. For cultivation of *R. sphaeroides* strains at 32 °C, malate minimal-salt medium was used [65]. Cultures were grown under microaerobic growth conditions, with a dissolved oxygen concentration of about 25–30 μM within the exponential phase. Erlenmeyer flasks containing 80% culture by volume were shaken at 140 rpm. For phototrophic cultivation, the strains were incubated in sealed Metplat flasks filled to the top and illuminated with 60 Wm^{-2} of white light. When necessary, kanamycin (25 μg mL^{-1}), tetracycline (2 μg mL^{-1}) or spectinomycin (10 μg mL^{-1}) was added to liquid and solid growth media (1.6% agar).

4.2. Construction of the rpoE-lacZ Fusion

For the *rpoE-lacZ* translational fusion, a 218 nt fragment of the *rpoE* gene was amplified with the primer pair rpoE_f and rpoE_r (Table S2). The fragment was subcloned into the pDrive cloning vector (Qiagen, Hilden, Germany) and the *rpoE* sequence was excised by XbaI and HindII and ligated into the corresponding sites of the pPHU4352 [24]. The resulting reporter plasmid pPHU_1092 (Tc^r) carried the translational *rpoE-lacZ* fusion under control of the 16S rRNA promoter (RSP_4352 promoter) and was transferred into *R. sphaeroides* strains by conjugation as described in [66].

4.3. β-Galactosidase Activity Measurements

For measuring ß-galactosidase activity, strains carrying the plasmid with the translational *rpoE-lacZ* fusion under control of the 16S rRNA promoter were incubated in biological triplicates at 32 °C under microaerobic conditions. ß-galactosidase activity was measured by the hydrolysis of O-nitrophenyl-β-D-galactopyranoside (ONPG) (Serva, Heidelberg, Germany) and expressed as Miller Units. Strains were grown until they reached an OD660 of 0.6. Cells were harvested, and the assay was performed as described in Klug et al. [67].

4.4. RNA Isolation

R. sphaeroides cultures from three independent starter cultures were inoculated separately and grown in triplicate to OD_{660nm} 0.5. For northern blot analysis, quantitative real-time RT-PCR and RNAseq analysis, RNA was isolated using the hot phenol method [68]. Afterwards, the RNA was precipitated with 1/10× vol. 3 M sodium acetate pH 4.5 and 2.5× vol. 96% ethanol.

4.5. Northern Blot Analysis

For Northern Blot analysis 10% polyacrylamide/urea gels were used to fractionate 8 μg total RNA, as described earlier [69]. Oligodeoxynucleotides were used for end-labelling with [γ-^{32}P]-ATP (SRP-30; Hartmann Analytic, Braunschweig, Germany) by T4 polynucleotide kinase (#EK0031, Fermentas, Ontario, Canada). A low stringency Church buffer was used for hybridization. Membranes were washed in 5× SCC buffer + 0.1%

SDS. After exposure on phosphoimaging screens (Bio-Rad), images were analyzed by 1D-Quantity One software (Bio-Rad, Feldkirchen, Germany). Oligonucleotides used for hybridization are listed in Table S2.

4.6. Quantitative Real-Time RT PCR

For qRT-PCR, total RNA was isolated using peqGOLD TriFast™ (VWR) as described by the manufacturer. Afterwards the RNA was treated with TURBO DNA-free™ Kit (Ambion/ ThermoFisher Scientific, Waltham, MA, USA) to remove DNA contaminations. For qRT-PCR, the Brilliant III Ultra-Fast SYBR® Green QPCR Master Mix was used for reverse transcription and PCR, as described in the manufacturer's manual. Each 10 µL reaction mixture contained 5 µL Master Mix (supplied), 0.1 µL DTT (100 mM, supplied), 0.5 µL RiboBlock solution (supplied), 0.4 µL water, 1 µl of each primer (10 pmol/L) and 2 µL DNA-free RNA (20 ng/µL). The reactions were performed in a spectrofluorometric thermal cycler (Biorad, Feldkirchen, Germany) and were visualized with BioRad CFX Manager 3.0. For all qRT-PCR experiments, means and standard deviations of biological triplicates were calculated, each performed in technical duplicates. For all primers, a no template-control was included. The expression of the target mRNAs in the strain of interest was calculated relative to the respective control strain and an in vitro transcript of *sinI* RNA, an external spike-in RNA of known sequence and quantity, was used for normalization [70]. Primers are listed in Table S1.

4.7. Gel Retardation Assay

For gel retardation assays, RNA was transcribed in vitro using T7 Polymerase (NEB, Massachusetts, USA) and PCR products with a T7 promoter region at the 5′ ends as the template. The assays were carried out with 150 fmol radio-labelled in vitro transcript and various molar ratios of nonlabelled in vitro transcripts in a final volume of 8 µL. RNAs were denatured separately for 1 min at 95 °C and renatured by cooling for 2 min on ice and for 5 min at 32 °C. After these de and renaturing steps, the radio-labelled and nonlabelled RNAs were mixed and 4 µL of 5× structure buffer (25 mM $MgCl_2$ and 300 mM KCl) were added for a final volume of 20 µL. For complex formation, the samples were incubated for 30 min at 32 °C. Afterwards, the reactions were mixed with 5 µL of loading dye (50% glycerol, 0.5× TBE, 0.2% bromophenol blue) and loaded onto a 6% nondenaturing polyacrylamide gel containing 0.5× TBE. Gels were pre-run at 100 V for 60 min at 4 °C before loading. Electrophoresis was performed at 4 °C by applying 200 V for 4 h. Gels were dried, exposed on phosphoimaging screens (Bio-Rad, Feldkirchen, Germany) and analyzed using 1D-Quantity One software (Bio-Rad, Feldkirchen, Germany).

4.8. RNAseq Analysis and Evaluation

RNA isolation, library preparation and bioinformatic analysis were performed as previously described [34,50]. DESeq2 (version 1.16.1; [39]) was applied for quantitative comparison of the data, and the p-value (Benjamini–Hochberg correction) was calculated. The data are deposited in GEO under the accession number GSE175997. Coverage plots in wiggle format representing the number of aligned reads per nucleotide were generated based on the aligned reads and visualized in the Integrated Genome Browser [71]. The raw coverage values of the graphs were normalized to the total number of reads that could be aligned for the respective library and multiplied by the minimum number of mapped reads of all libraries.

4.9. Half-Life Determination of the mRNA

RNA samples were prepared at different time points after addition of rifampicin to the cultures as described in RNA isolation and quantification. Half-lives were calculated based on real time RT-PCR with 20 ng of total RNA for both the target gene and for the standard gene *rpoZ*.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/ijms22147557/s1, Figures S1–S3 and Tables S1 and S2.

Author Contributions: Conceptualization, B.R., K.M.H.E., J.G., G.K.; Validation, B.R., K.M.H.E., J.G., D.-T.S., A.J., G.K.; investigation: B.R., K.M.H.E., J.G., D.-T.S., A.J.; data curation, B.R., K.M.H.E., J.G., A.J., D.-T.S.; formal analysis, D.-T.S., writing, G.K.; writing—review and editing, J.G., G.K.; visualization, B.R., K.M.H.E., J.G., A.J., G.K.; supervision, B.R., K.M.H.E., J.G., G.K.; project administration, G.K.; funding acquisition., G.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Deutsche Forschungsgemeinschaft (Kl563/28-1, Kl563/41-1).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The RNAseq data are available in the NCBI gene expression omnibus repository (GEO accession number GSE71844).

Acknowledgments: We thank Sabine Martini, Janis Kiebel, and Markus Späth for contributing experiments concerning the role of StsR, under supervision of the authors.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bathke, J.; Konzer, A.; Remes, B.; McIntosh, M.; Klug, G. Comparative analyses of the variation of the transcriptome and proteome of Rhodobacter sphaeroides throughout growth. *BMC Genom.* **2019**, *20*, 1–13. [CrossRef]
2. Wösten, M.M. Eubacterial sigma-factors. *FEMS Microbiol. Rev.* **1998**, *22*, 127–150. [CrossRef]
3. Paget, M.S.; Helmann, J.D. The σ^{70} family of sigma factors. *Genome Biol.* **2003**, *4*, 203. [CrossRef] [PubMed]
4. Feklístov, A.; Sharon, B.D.; Darst, S.A.; Gross, C.A. Bacterial Sigma Factors: A Historical, Structural, and Genomic Perspective. *Annu. Rev. Microbiol.* **2014**, *68*, 357–376. [CrossRef]
5. De Bruijn, F.J. (Ed.) *Stress and Environmental Regulation of Gene Expression and Adaptation in Bacteria*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2016; ISBN 9781119004813.
6. *Advances in Microbial Physiology*; Academic: London, UK, 2011; Volume 59, ISBN 9780123876614.
7. Yoon, S.H.; Waters, C.M. The ever-expanding world of bacterial cyclic oligonucleotide second messengers. *Curr. Opin. Microbiol.* **2021**, *60*, 96–103. [CrossRef] [PubMed]
8. Waters, L.S.; Storz, G. Regulatory RNAs in Bacteria. *Cell* **2009**, *136*, 615–628. [CrossRef] [PubMed]
9. Jørgensen, M.G.; Pettersen, J.S.; Kallipolitis, B.H. sRNA-mediated control in bacteria: An increasing diversity of regulatory mechanisms. *Biochim. Biophys. Acta Gene Regul. Mech.* **2020**, *1863*, 194504. [CrossRef] [PubMed]
10. Wagner, E.G.H.; Romby, P. Small RNAs in bacteria and archaea: Who they are, what they do, and how they do it. *Fungal Phylogenet. Phylogenomics* **2015**, *90*, 133–208. [CrossRef]
11. Michaux, C.; Verneuil, N.; Hartke, A.; Giard, J.-C. Physiological roles of small RNA molecules. *Microbiology* **2014**, *160*, 1007–1019. [CrossRef]
12. O'Gara, J.P.; Eraso, J.M.; Kaplan, S. A Redox-Responsive Pathway for Aerobic Regulation of Photosynthesis Gene Expression in Rhodobacter sphaeroides 2.4.1. *J. Bacteriol.* **1998**, *180*, 4044–4050. [CrossRef]
13. Oh, J.-I.; Ko, I.-J.; Kaplan, S. Reconstitution of the Rhodobacter sphaeroides cbb3-PrrBA Signal Transduction Pathway in Vitro. *Biochemistry* **2004**, *43*, 7915–7923. [CrossRef]
14. Han, Y.; Meyer, M.H.F.; Keusgen, M.; Klug, G. A haem cofactor is required for redox and light signalling by the AppA protein of Rhodobacter sphaeroides. *Mol. Microbiol.* **2007**, *64*, 1090–1104. [CrossRef]
15. Braatsch, S.; Gomelsky, M.; Kuphal, S.; Klug, G. A single flavoprotein, AppA, integrates both redox and light signals in Rhodobacter sphaeroides. *Mol. Microbiol.* **2002**, *45*, 827–836. [CrossRef] [PubMed]
16. Masuda, S.; Bauer, C.E. AppA Is a Blue Light Photoreceptor that Antirepresses Photosynthesis Gene Expression in Rhodobacter sphaeroides. *Cell* **2002**, *110*, 613–623. [CrossRef]
17. Moskvin, O.V.; Kaplan, S.; Gilles-Gonzalez, M.-A.; Gomelsky, M. Novel Heme-based Oxygen Sensor with a Revealing Evolutionary History. *J. Biol. Chem.* **2007**, *282*, 28740–28748. [CrossRef] [PubMed]
18. Gomelsky, M.; Klug, G. BLUF: A novel FAD-binding domain involved in sensory transduction in microorganisms. *Trends Biochem. Sci.* **2002**, *27*, 497–500. [CrossRef]
19. Gomelsky, L.; Sram, J.; Moskvin, O.V.; Horne, I.M.; Dodd, H.N.; Pemberton, J.M.; McEwan, A.G.; Kaplan, S.; Gomelsky, M. Identification and in vivo characterization of PpaA, a regulator of photosystem formation in Rhodobacter sphaeroides. *Microbiology* **2003**, *149*, 377–388. [CrossRef] [PubMed]
20. Vermeulen, A.J.; Bauer, C.E. Members of the PpaA/AerR Antirepressor Family Bind Cobalamin. *J. Bacteriol.* **2015**, *197*, 2694–2703. [CrossRef]

21. Zeilstra-Ryalls, J.H.; Kaplan, S. Role of the fnrL Gene in Photosystem Gene Expression and Photosynthetic Growth of Rhodobacter sphaeroides 2.4.1. *J. Bacteriol.* **1998**, *180*, 1496–1503. [CrossRef]
22. Zeilstra-Ryalls, J.H.; Kaplan, S. Aerobic and anaerobic regulation in Rhodobacter sphaeroides 2.4.1: The role of the fnrL gene. *J. Bacteriol.* **1995**, *177*, 6422–6431. [CrossRef]
23. Imam, S.; Noguera, D.R.; Donohue, T.J. Global Analysis of Photosynthesis Transcriptional Regulatory Networks. *PLoS Genet.* **2014**, *10*, e1004837. [CrossRef] [PubMed]
24. Mank, N.N.; Berghoff, B.A.; Hermanns, Y.N.; Klug, G. Regulation of bacterial photosynthesis genes by the small noncoding RNA PcrZ. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 16306–16311. [CrossRef]
25. Mank, N.N.; Berghoff, B.A.; Klug, G. A mixed incoherent feed-forward loop contributes to the regulation of bacterial photosynthesis genes. *RNA Biol.* **2013**, *10*, 347–352. [CrossRef]
26. Eisenhardt, K.M.H.; Reuscher, C.M.; Klug, G. PcrX, an sRNA derived from the 3′- UTR of the Rhodobacter sphaeroides puf operon modulates expression of puf genes encoding proteins of the bacterial photosynthetic apparatus. *Mol. Microbiol.* **2018**, *110*, 325–334. [CrossRef] [PubMed]
27. Reuscher, C.M.; Klug, G. Antisense RNA asPcrL regulates expression of photosynthesis genes in Rhodobacter sphaeroides by promoting RNase III-dependent turn-over of puf mRNA. *RNA Biol.* **2021**, 1–13. [CrossRef]
28. Glaeser, J.; Nuss, A.; Berghoff, B.; Klug, G. Singlet Oxygen Stress in Microorganisms. *Adv. Bact. Pathog. Biol.* **2011**, *58*, 141–173. [CrossRef]
29. Nuss, A.M.; Glaeser, J.; Klug, G. $RpoH_{II}$ Activates Oxidative-Stress Defense Systems and Is Controlled by RpoE in the Singlet Oxygen-Dependent Response in Rhodobacter sphaeroides. *J. Bacteriol.* **2009**, *191*, 220–230. [CrossRef]
30. Nuss, A.M.; Glaeser, J.; Berghoff, B.A.; Klug, G. Overlapping Alternative Sigma Factor Regulons in the Response to Singlet Oxygen in Rhodobacter sphaeroides. *J. Bacteriol.* **2010**, *192*, 2613–2623. [CrossRef]
31. Dufour, Y.S.; Imam, S.; Koo, B.-M.; Green, H.A.; Donohue, T.J. Convergence of the Transcriptional Responses to Heat Shock and Singlet Oxygen Stresses. *PLoS Genet.* **2012**, *8*, e1002929. [CrossRef] [PubMed]
32. Anthony, J.R.; Newman, J.D.; Donohue, T.J. Interactions Between the Rhodobacter sphaeroides ECF Sigma Factor, σE, and its Anti-sigma Factor, ChrR. *J. Mol. Biol.* **2004**, *341*, 345–360. [CrossRef]
33. Nuss, A.M.; Adnan, F.; Weber, L.; Berghoff, B.A.; Glaeser, J.; Klug, G. DegS and RseP Homologous Proteases Are Involved in Singlet Oxygen Dependent Activation of RpoE in Rhodobacter sphaeroides. *PLoS ONE* **2013**, *8*, e79520. [CrossRef] [PubMed]
34. Remes, B.; Rische-Grahl, T.; Müller, K.M.H.; Förstner, K.U.; Yu, S.-H.; Weber, L.; Jäger, A.; Peuser, V.; Klug, G. An RpoHI-Dependent Response Promotes Outgrowth after Extended Stationary Phase in the Alphaproteobacterium Rhodobacter sphaeroides. *J. Bacteriol.* **2017**, *199*, e00249-17. [CrossRef] [PubMed]
35. Billenkamp, F.; Peng, T.; Berghoff, B.A.; Klug, G. A Cluster of Four Homologous Small RNAs Modulates C_1 Metabolism and the Pyruvate Dehydrogenase Complex in Rhodobacter sphaeroides under Various Stress Conditions. *J. Bacteriol.* **2015**, *197*, 1839–1852. [CrossRef]
36. Peuser, V.; Remes, B.; Klug, G. Role of the Irr Protein in the Regulation of Iron Metabolism in Rhodobacter sphaeroides. *PLoS ONE* **2012**, *7*, e42231. [CrossRef]
37. Grützner, J.; Remes, B.; Eisenhardt, K.M.H.; Scheller, D.; Kretz, J.; Madhugiri, R.; McIntosh, M.; Klug, G. sRNA-mediated RNA processing regulates bacterial cell division. *Nucleic Acids Res.* **2021**. [CrossRef] [PubMed]
38. Weber, L.; Thoelken, C.; Volk, M.; Remes, B.; Lechner, M.; Klug, G. The Conserved Dcw Gene Cluster of R. sphaeroides Is Preceded by an Uncommonly Extended 5′ Leader Featuring the sRNA UpsM. *PLoS ONE* **2016**, *11*, e0165694. [CrossRef] [PubMed]
39. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 1–21. [CrossRef]
40. Zeng, X.; Kaplan, S. TspO as a Modulator of the Repressor/Antirepressor (PpsR/AppA) Regulatory System in Rhodobacter sphaeroides 2.4.1. *J. Bacteriol.* **2001**, *183*, 6355–6364. [CrossRef] [PubMed]
41. Grützner, J.; Billenkamp, F.; Spanka, D.-T.; Rick, T.; Monzon, V.; Förstner, K.U.; Klug, G. The small DUF1127 protein CcaF1 from Rhodobacter sphaeroides is an RNA-binding protein involved in sRNA maturation and RNA turnover. *Nucleic Acids Res.* **2021**, *49*, 3003–3019. [CrossRef]
42. Domenzain, C.; Camarena, L.; Osorio, A.; Dreyfus, G.; Poggio, S. Evolutionary origin of the Rhodobacter sphaeroides specialized RpoN sigma factors. *FEMS Microbiol. Lett.* **2012**, *327*, 93–102. [CrossRef] [PubMed]
43. Berghoff, B.A.; Klug, G. An Omics View on the Response to Singlet Oxygen. In *Stress and Environmental Regulation of Gene Expression and Adaptation in Bacteria*; De Bruijn, F.J., Ed.; John Wiley & Sons, Inc: Hoboken, NJ, USA, 2016; pp. 619–631. ISBN 9781119004813.
44. Müller, K.M.H.; Berghoff, B.A.; Eisenhardt, B.D.; Remes, B.; Klug, G. Characteristics of Pos19–A Small Coding RNA in the Oxidative Stress Response of Rhodobacter sphaeroides. *PLoS ONE* **2016**, *11*, e0163425. [CrossRef] [PubMed]
45. Peng, T.; Berghoff, B.A.; Oh, J.-I.; Weber, L.; Schirmer, J.; Schwarz, J.; Glaeser, J.; Klug, G. Regulation of a polyamine transporter by the conserved 3′ UTR-derived sRNA SorX confers resistance to singlet oxygen and organic hydroperoxides in Rhodobacter sphaeroides. *RNA Biol.* **2016**, *13*, 988–999. [CrossRef] [PubMed]
46. Adnan, F.; Weber, L.; Klug, G. The sRNA SorY confers resistance during photooxidative stress by affecting a metabolite transporter inRhodobacter sphaeroides. *RNA Biol.* **2015**, *12*, 569–577. [CrossRef]

47. Mann, M.; Wright, P.R.; Backofen, R. IntaRNA 2.0: Enhanced and customizable prediction of RNA–RNA interactions. *Nucleic Acids Res.* **2017**, *45*, W435–W439. [CrossRef]
48. Kavita, K.; de Mets, F.; Gottesman, S. New aspects of RNA-based regulation by Hfq and its partner sRNAs. *Curr. Opin. Microbiol.* **2018**, *42*, 53–61. [CrossRef]
49. Morita, T.; Maki, K.; Aiba, H. RNase E-based ribonucleoprotein complexes: Mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs. *Genes Dev.* **2005**, *19*, 2176–2186. [CrossRef] [PubMed]
50. Förstner, K.U.; Reuscher, C.M.; Haberzettl, K.; Weber, L.; Klug, G. RNase E cleavage shapes the transcriptome ofRhodobacter sphaeroidesand strongly impacts phototrophic growth. *Life Sci. Alliance* **2018**, *1*, e201800080. [CrossRef]
51. Saoud, J.; Carrier, M.-C.; Massé, É.; Faucher, S.P. The small regulatory RNA Lpr10 regulates the expression of RpoS in Legionella pneumophila. *Mol. Microbiol.* **2021**, *115*, 789–806. [CrossRef]
52. Engel, F.; Ossipova, E.; Jakobsson, P.-J.; Vockenhuber, M.-P.; Suess, B. sRNA scr5239 Involved in Feedback Loop Regulation of Streptomyces coelicolor Central Metabolism. *Front. Microbiol.* **2020**, *10*, 3121. [CrossRef]
53. Khan, M.A.; Durica-Mitic, S.; Göpel, Y.; Heermann, R.; Görke, B. Small RNA-binding protein RapZ mediates cell envelope precursor sensing and signaling in Escherichia coli. *EMBO J.* **2020**, *39*, e103848. [CrossRef] [PubMed]
54. Butz, H.A.; Mey, A.R.; Ciosek, A.L.; Payne, S.M. Vibrio cholerae CsrA Directly Regulates varA to Increase Expression of the Three Nonredundant Csr Small RNAs. *MBio* **2019**, *10*, e01042-19. [CrossRef]
55. Zeilstra-Ryalls, J.H.; Kaplan, S. Oxygen intervention in the regulation of gene expression: The photosynthetic bacterial paradigm. *Cell. Mol. Life Sci.* **2004**, *61*, 417–436. [CrossRef] [PubMed]
56. Berghoff, B.A.; Glaeser, J.; Sharma, C.M.; Zobawa, M.; Lottspeich, F.; Vogel, J.; Klug, G. Contribution of Hfq to photooxidative stress resistance and global regulation in Rhodobacter sphaeroides. *Mol. Microbiol.* **2011**, *80*, 1479–1495. [CrossRef]
57. Hussein, R.; Lim, H.N. Disruption of small RNA signaling caused by competition for Hfq. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 1110–1115. [CrossRef]
58. Olejniczak, M. Despite Similar Binding to the Hfq Protein Regulatory RNAs Widely Differ in Their Competition Performance. *Biochemistry* **2011**, *50*, 4427–4440. [CrossRef]
59. Moon, K.; Gottesman, S. Competition among Hfq-binding small RNAs in Escherichia coli. *Mol. Microbiol.* **2011**, *82*, 1545–1562. [CrossRef]
60. Santiago-Frangos, A.; Woodson, S.A. Hfq chaperone brings speed dating to bacterial sRNA. *Wiley Interdiscip. Rev. RNA* **2018**, *9*, e1475. [CrossRef] [PubMed]
61. Sonnleitner, E.; Prindl, K.; Bläsi, U. The Pseudomonas aeruginosa CrcZ RNA interferes with Hfq-mediated riboregulation. *PLoS ONE* **2017**, *12*, e0180887. [CrossRef]
62. Faigenbaum-Romm, R.; Reich, A.; Gatt, Y.E.; Barsheshet, M.; Argaman, L.; Margalit, H. Hierarchy in Hfq Chaperon Occupancy of Small RNA Targets Plays a Major Role in Their Regulation. *Cell Rep.* **2020**, *30*, 3127–3138.e6. [CrossRef] [PubMed]
63. Małecka, E.; Stróżecka, J.; Sobańska, D.; Olejniczak, M. Structure of Bacterial Regulatory RNAs Determines Their Performance in Competition for the Chaperone Protein Hfq. *Biochemistry* **2015**, *54*, 1157–1170. [CrossRef]
64. Van Niel, C.B. The culture, general physiology, morphology, and classification of the non-sulfur purple and brown bacteria. *Bacteriol. Rev.* **1944**, *8*, 1–118. [CrossRef]
65. Remes, B.; A Berghoff, B.A.; Förstner, K.U.; Klug, G. Role of oxygen and the OxyR protein in the response to iron limitation in Rhodobacter sphaeroides. *BMC Genom.* **2014**, *15*, 1–11. [CrossRef]
66. Klug, G.; Drews, G. Construction of a gene bank of Rhodopseudomonas capsulata using a broad host range DNA cloning system. *Arch. Microbiol.* **1984**, *139*, 319–325. [CrossRef] [PubMed]
67. Klug, G.; Jäger, A.; Heck, C.; Rauhut, R. Identification, sequence analysis, and expression of the lepB gene for a leader peptidase in Rhodobacter capsulatus. *Mol. Genet. Genom.* **1997**, *253*, 666–673. [CrossRef] [PubMed]
68. Damm, K.; Bach, S.; Müller, K.M.H.; Klug, G.; Burenina, O.Y.; Kubareva, E.A.; Grünweller, A.; Hartmann, R.K. Impact of RNA Isolation Protocols on RNA Detection by Northern Blotting. *Methods Mol. Biol.* **2015**, *1296*, 29–38. [CrossRef] [PubMed]
69. Berghoff, B.A.; Glaeser, J.; Sharma, C.M.; Vogel, J.; Klug, G. Photooxidative stress-induced and abundant small RNAs in Rhodobacter sphaeroides. *Mol. Microbiol.* **2009**, *74*, 1497–1512. [CrossRef]
70. Pfaffl, M.W. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* **2001**, *29*, e45. [CrossRef]
71. Nicol, J.W.; Helt, G.A.; Blanchard, S.G.; Raja, A.; Loraine, A.E. The Integrated Genome Browser: Free software for distribution and exploration of genome-scale datasets. *Bioinformatics* **2009**, *25*, 2730–2731. [CrossRef] [PubMed]

Chapter 6

High-throughput proteomics identifies proteins with importance to postantibiotic recovery in depolarized persister cells

frontiers in Microbiology

ORIGINAL RESEARCH
published: 06 March 2019
doi: 10.3389/fmicb.2019.00378

High-Throughput Proteomics Identifies Proteins With Importance to Postantibiotic Recovery in Depolarized Persister Cells

Daniel-Timon Spanka[1†], *Anne Konzer*[2†], *Daniel Edelmann*[1] *and Bork A. Berghoff*[1*]

[1] *Institute for Microbiology and Molecular Biology, Justus Liebig University Giessen, Giessen, Germany,* [2] *Biomolecular Mass Spectrometry, Max-Planck-Institute for Heart and Lung Research, Bad Nauheim, Germany*

Bacterial populations produce phenotypic variants called persisters to survive harmful conditions. Persisters are highly tolerant to antibiotics and repopulate environments after the stress has vanished. In order to resume growth, persisters have to recover from the persistent state, but the processes behind recovery remain mostly elusive. Deciphering these processes is an essential step toward understanding the persister phenomenon in its entirety. High-throughput proteomics by mass spectrometry is a valuable tool to assess persister physiology during any stage of the persister life cycle, and is expected to considerably contribute to our understanding of the recovery process. In the present study, an *Escherichia coli* strain, that overproduces the membrane-depolarizing toxin TisB, was established as a model for persistence by the use of high-throughput proteomics. Labeling of TisB persisters with stable isotope-containing amino acids (pulsed-SILAC) revealed an active translational response to ampicillin, including several RpoS-dependent proteins. Subsequent investigation of the persister proteome during postantibiotic recovery by label-free quantitative proteomics identified proteins with importance to the recovery process. Among them, AhpF, a component of alkyl hydroperoxide reductase, and the outer membrane porin OmpF were found to affect the persistence time of TisB persisters. Assessing the role of AhpF and OmpF in TisB-independent persisters demonstrated that the importance of a particular protein for the recovery process strongly depends on the physiological condition of a persister cell. Our study provides important insights into persister physiology and the processes behind recovery of depolarized cells.

Keywords: persister cells, antibiotic tolerance, depolarization, TisB toxin, recovery, SILAC, proteomics

OPEN ACCESS

Edited by:
Satoshi Tsuneda,
Waseda University, Japan

Reviewed by:
Anna D. Tischler,
University of Minnesota Twin Cities, United States
Hanne Ingmer,
University of Copenhagen, Denmark

****Correspondence:***
Bork A. Berghoff
bork.a.berghoff@mikro.bio.uni-giessen.de

[†]These authors have contributed equally to this work

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

Received: *12 September 2018*
Accepted: *13 February 2019*
Published: *06 March 2019*

Citation:
Spanka D-T, Konzer A, Edelmann D and Berghoff BA (2019) High-Throughput Proteomics Identifies Proteins With Importance to Postantibiotic Recovery in Depolarized Persister Cells. Front. Microbiol. 10:378. doi: 10.3389/fmicb.2019.00378

INTRODUCTION

The rise of antibiotic resistance among pathogens is a major threat to the human health care system (Lewis, 2013), and gains ever-expanding attention. Bacteria have, however, developed alternative strategies to survive an antibiotic challenge. For instance, Joseph W. Bigger realized very early after the introduction of penicillin that a small subpopulation of otherwise susceptible *Staphylococcus aureus* cultures survived a penicillin treatment for several days. He termed these surviving cells "persisters" (Bigger, 1944). Persisters are transiently drug-tolerant phenotypic variants within

isogenic populations and, in contrast to resistant bacteria, do not proliferate in the presence of antibiotics. Furthermore, the minimum inhibitory concentration (MIC) for a persistent strain is not altered in comparison to a strain that is susceptible to a particular antibiotic (Brauner et al., 2016). Even though persisters have been extensively studied during the last years, there is still a lack of knowledge regarding the physiological state of persisters. The general perception of persisters is that they are non-growing or slowly growing cells (Balaban et al., 2004), and that reduced activity of antibiotic targets renders them tolerant to antibiotics (Lewis, 2010). Clearly, different types of persisters exist (Balaban et al., 2004), and it is feasible to assume that physiological states are different as well. For example, extremely dormant persisters might be similar to viable but non-culturable (VBNC) cells (Ayrapetyan et al., 2015; Kim et al., 2018). By contrast, persisters induced by carbon source shifts retain metabolic activity, exhibit slow growth, and are marked by a distinct proteome pattern (Radzikowski et al., 2016). Hypothetically, the physiology of persisters is dependent on the particular mechanism that has triggered entry into the persistent state.

Endogenous factors, that reduce cellular activity and potentially favor persistence, are toxins from toxin-antitoxin (TA) systems. TA systems were discovered on plasmids, where they are implicated in plasmid maintenance during cell proliferation, but were later also identified on bacterial chromosomes in surprisingly high numbers (Hayes, 2003; Gerdes et al., 2005). They are classified according to the specific mechanism by which the antitoxin inhibits expression or activity of its toxin counterpart (Page and Peti, 2016). In type I TA systems, the antitoxin is an antisense RNA that specifically blocks translation of the toxin mRNA (Fozo et al., 2008a; Brantl and Jahn, 2015). In type II TA systems, the antitoxin inhibits activity of the toxin via protein-protein interaction (Gerdes and Maisonneuve, 2012). Several lines of evidence indicate that chromosomal TA systems play a role in persister formation. In fact, the first "persister gene" to be discovered was *hipA* from the type II TA system *hipAB* in *E. coli* (Moyed and Bertrand, 1983; Black et al., 1991, 1994). Toxin HipA inactivates glutamyl-tRNA-synthetase (GltX) by phosphorylation, which leads to disturbed aminoacylation (Germain et al., 2013; Kaspy et al., 2013). As a consequence, the stringent response alarmone (p)ppGpp is produced and transition into the persistent state is favored (Korch et al., 2003). The possible role of chromosomal TA systems in bacterial persistence was further underscored in the mid 2000's, when it was observed that several toxin genes from type II TA systems were upregulated in persister cells, which for example applies to *mazF* and *relE* (Keren et al., 2004; Shah et al., 2006). MazF and RelE are mRNA endonucleases that impede translation and cause growth stasis. Intriguingly, both activation of MazF and overexpression of RelE mediated persister formation in *E. coli* (Keren et al., 2004; Tripathi et al., 2014). Further evidence for toxins as "persistence factors" comes from work with the mRNA endonucleases YafQ and MqsR. Single gene deletions of *yafQ* and *mqsR* caused a reduction in persistence (Harrison et al., 2009; Kim and Wood, 2010). Finally, toxins from type I TA systems have been directly linked to persister formation. These toxins are often small hydrophobic proteins that preferentially localize to the inner membrane to cause break-down of the proton motive force or ATP leakage (Fozo et al., 2008b; Unoson and Wagner, 2008; Gurnev et al., 2012; Weel-Sneve et al., 2013; Wilmaerts et al., 2018). Membrane depolarization and depletion of intracellular ATP potentially trigger entry into a persistent state, as shown for toxins TisB and HokB (Dörr et al., 2010; Verstraeten et al., 2015; Berghoff et al., 2017; Wilmaerts et al., 2018). In *E. coli* and *S. aureus*, artificial ATP depletion by the addition of arsenate is sufficient to induce persister formation (Conlon et al., 2016; Shan et al., 2017). However, membrane depolarization alone, rather than ATP depletion, might be a determinant of persistence in *S. aureus* as well (Wang et al., 2018).

Another central question in the persister field concerns the mechanisms that affect persister awakening. The "microbial scout" model suggests that dormant cells awake stochastically to sample their environment for suitable conditions (Buerger et al., 2012). In case of *Bacillus* spores, the model seems reasonable, and some spores were indeed shown to awaken spontaneously without sensing of outside signals (Sturm and Dworkin, 2015). Work with the model bacterium *E. coli* showed that the outgrowth medium has an influence on wake-up kinetics of persister cells (Jõers et al., 2010), indicating that persister cells may sense their environment. It is, however, not known which cellular proteins are at the forefront of awakening. Do persisters express specific proteins to neutralize the effect of a particular toxin? The answer is possibly yes; it was demonstrated that acetylation of tRNAs by toxin TacT is reversed by a peptidyl-tRNA hydrolase (Cheverton et al., 2016). However, this specific enzyme cannot explain awakening of persisters that have formed through other mechanisms. Moreover, even though neutralization of toxins might initiate awakening, it can be expected that further proteins serve specific functions during the subsequent recovery process. For instance, DNA repair during recovery seems to be key to persistence of non-growing *E. coli* cells after ofloxacin treatment (Völzing and Brynildsen, 2015).

In the present study, we have chosen TisB as a model system for "persistence by depolarization" during exponential growth phase (Dörr et al., 2010; Berghoff et al., 2017). TisB is the toxin moiety of the TisB/IstR-1 TA system and is induced upon DNA damage due to activation of the SOS response via cleavage of the LexA repressor (Vogel et al., 2004). TisB targets the inner membrane and causes depolarization (Unoson and Wagner, 2008; Gurnev et al., 2012). Translation of the primary *tisB* mRNA is repressed by an inhibitory secondary structure in its 5′ untranslated region (UTR). After cleavage of the 5′ UTR structure, translation is blocked by the antitoxin IstR-1 (Darfeuille et al., 2007; Berghoff and Wagner, 2017). We recently deleted both regulatory RNA elements in *E. coli* K-12 wild type MG1655. The resulting double deletion strain Δ1-41 Δ*istR* exhibits unchecked expression of TisB, is highly persistent when treated with different antibiotics, and represents, therefore, a suitable system to study TisB-dependent persisters (Berghoff et al., 2017). Here, state-of-the-art mass spectrometry (MS) methods were applied to assess the persister proteome, both during antibiotic challenge and during postantibiotic recovery. A pulsed-SILAC (stable isotope labeling by amino

acids in cell culture) approach was applied during ampicillin treatment, highlighting 43 proteins with significantly increased protein synthesis. Since many of these proteins are stress-related and serve protective functions, we conclude that TisB-dependent persisters mount an active response to ampicillin. Furthermore, protein samples from the recovery phase were analyzed by label-free quantitative MS to identify proteins with differential abundance. Among the 24 proteins with increased abundance during recovery, we identified a component of the alkyl hydroperoxide reductase (AhpF) and an outer membrane porin (OmpF). Deletions of *ahpF* and *ompF* in the Δ1-41 Δ*istR* background caused an extended period of persistence after antibiotic treatment, indicating that both proteins play important roles during recovery from the persistent state. However, subsequent experiments showed that these functions are specific to TisB-dependent persister cells. We conclude that functions needed for the recovery process have to match the specific physiological state of a persister cell.

MATERIALS AND METHODS

Growth Conditions

For physiological experiments, *E. coli* strains were grown in lysogeny broth (LB) or M9 minimal medium at 37°C with continuous shaking at 180 rpm (aerobic growth). *E. coli* strains containing temperature-sensitive plasmids were grown at 30°C. If applicable, antibiotics were added at the following concentrations: 50 μg ml^{-1} kanamycin, 15 μg ml^{-1} chloramphenicol, 6 μg ml^{-1} tetracycline, 50 μg ml^{-1} ampicillin. Over-night cultures were diluted 100-fold into fresh medium and incubated until an optical density at 600 nm (OD_{600}) of 0.3–0.6 (exponential phase) was reached. For stationary phase experiments, liquid cultures were inoculated with single colonies and grown for 20 h. For growth curves, the initial OD_{600} was adjusted to 0.02. Growth curves were monitored using a Cell density meter model 40 (Fisher Scientific). Doubling times were calculated from exponential growth phase and *P*-values were assessed using Student's *t*-test.

Construction of Bacterial Strains

E. coli strains used in this study are derivatives of K-12 wild type MG1655 and are listed in **Supplementary Table 1**. Chromosomal deletions of candidate genes were constructed by homologous recombination using the λ red genes (Yu et al., 2000). A chloramphenicol acetyltransferase (*cat*) or kanamycin resistance (*kan*) gene was PCR-amplified together with specific overhangs (40 bp) on each side to enable recombination within the desired gene locus. The corresponding oligodeoxyribonucleotides are listed in **Supplementary Table 2**. The linear amplification products were transformed into *E. coli* strains containing temperature-sensitive plasmid pSIM5 for heat-inducible expression of λ red genes (Datta et al., 2006). After recombination, clones were selected on LB agar plates containing chloramphenicol (12.5 μg ml^{-1}) or kanamycin (25 μg ml^{-1}), respectively. Insertion of the *cat* or *kan* gene was verified by PCR with oligodeoxyribonucleotides listed in **Supplementary Table 2**. Deletion constructs containing selectable markers were transferred to the wild-type background by P1 transduction.

For simultaneous deletion of *ahpF* and *ompF* in strain B133 (Δ1-41 Δ*istR::frt-kan-frt*), the *kan* gene of strain B133 was removed by FLP-mediated site-directed mutagenesis using plasmid 709-FLPe (Gene Bridges) according to the manufacturer's protocol. Clones were tested for kanamycin sensitivity, and *ahpF* and *ompF* were subsequently deleted by λ red-mediated recombination using *cat* and *kan* genes, respectively, for selection. The corresponding oligodeoxyribonucleotides are listed in **Supplementary Table 2**.

Pulse-Labeling With Stable Isotopes

E. coli strain B133 (Δ1-41 Δ*istR::frt-kan-frt*) was pulse-labeled with the stable isotope-containing amino acid L-lysine-$^{13}C_6$,$^{15}N_2$ (Lys8). The protocol is a combination of native and pulsed-SILAC approaches that have been successfully applied to prototrophic bacteria (Michalik et al., 2012; Berghoff et al., 2013; Fröhlich et al., 2013). *E. coli* was cultivated in M9 minimal medium (final concentrations: 47.7 mM Na_2HPO_4, 22 mM KH_2PO_4, 8.6 mM NaCl, 18.7 mM NH_4Cl, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 1 μg ml^{-1} thiamine) containing 0.4% glucose as carbon source (M9+glu). Regular L-lysine (Lys0) was added to over-night cultures at a final concentration of 30 μg ml^{-1}. For pulse-labeling experiments, Erlenmeyer flasks (50 ml) were filled with 10 ml M9+glu, supplemented with 30 μg ml^{-1} Lys0, and inoculated with cells from over-night cultures. When an OD_{600} of 0.3–0.4 was reached, cells were pelleted by centrifugation (10,000 g, 3 min), washed in 1 ml M9+glu to remove Lys0, resuspended in 10 ml fresh M9+glu, and transferred to Erlenmeyer flasks (50 ml). Lys8 (30 μg ml^{-1}) and ampicillin (200 μg ml^{-1}) were immediately added to start pulse-labeling and antibiotic challenge in parallel. A culture incubated in the presence of Lys8 but without ampicillin served as treatment control. Cultures were incubated for 4 h at 37°C and 180 rpm. Cells were harvested by cold centrifugation (10,000 g, 3 min, 4°C), washed with 1 ml ice-cold M9+glu, and pelleted by centrifugation (10,000 g, 3 min, 4°C). All cell pellets were stored at −20°C until preparation of protein samples for mass spectrometry analysis.

Recovery Experiments

Recovery experiments were performed in biological triplicates with *E. coli* strain B133 (Δ1-41 Δ*istR::frt-kan-frt*). Erlenmeyer flasks (100 ml) were filled with 20 ml LB medium and inoculated with cells from over-night cultures. When exponential phase (OD_{600} 0.3–0.6) was reached, cultures were treated with 200 μg ml^{-1} ampicillin to lyse non-persistent cells. Persister cells were harvested by centrifugation (10,000 g, 3 min) after 2 h of treatment. Cells were either washed with 1 ml ice-cold NaCl solution (0.9%) and pelleted by centrifugation (10,000 g, 3 min, 4°C) or prepared for recovery in fresh LB medium without antibiotics. For this purpose, cells were washed in 1 ml NaCl solution (0.9%), followed by centrifugation (10,000 g, 3 min), resuspension in 20 ml LB medium, and incubation in Erlenmeyer flasks (100 ml) at 37°C and 180 rpm. At different time points during and following recovery, cells were harvested by cold

centrifugation and washed with ice-cold NaCl solution (0.9%) as described before. All cell pellets were stored at −20°C until preparation of protein samples for SDS-PAGE or mass spectrometry analysis.

SDS-PAGE of Protein Samples

Cell pellets from recovery experiments were resuspended in 500 µl phosphate buffer (50 mM, pH 7.2) and disrupted by sonication on ice (3 × 30 s; cycle: 70%; power: 70%). Cell debris and intact cells were removed by centrifugation (13,000 rpm, 10 min, 4°C). Protein concentration of supernatants was determined using a NanoDrop ND-1000 spectrophotometer (PeqLab). Two hundred micrograms of protein were precipitated with acetone at −20°C for 3 h. After centrifugation (13,000 rpm, 10 min, 4°C), protein pellets were washed two times with ice-cold acetone. Pellets were resolved in 1 × SDS sample buffer at 95°C for 10 min, and 50 µg of protein were loaded on 12% polyacrylamide gels. Gels were stained with colloidal Coomassie (Roth) for visualization of protein bands.

Protein Sample Preparation for Mass Spectrometry

E. coli cell pellets were lysed in SDS-buffer (4% SDS in 0.1 M Tris/HCl, pH 7.6) by heating at 70°C for 10 min and sonication. Next, solubilized proteins were separated from cell debris by centrifugation at 16,000 g for 10 min and protein concentration in the supernatant was determined using the DC protein assay (BioRad). From each sample, an equal protein amount was precipitated by 4 volumes of 100% acetone at −20°C for 1 h, pelleted at 14,000 g for 10 min and washed with 90% acetone. Dried protein pellets were dissolved in urea buffer (6 M urea, 2 M thiourea, 10 mM HEPES, pH 8.0) and enzymatic protein digest was performed by in-solution digestion as previously described (Andersen et al., 2005). In brief, protein disulfide bonds were reduced with 10 mM dithiothreitol and alkylated with 55 mM iodoacetamide. Next, proteins were cleaved enzymatically in two steps at room temperature: pre-digestion was performed by Lys-C (100:1 protein-to-enzyme ratio) (Wako Chemicals GmbH) for 3 h followed by an overnight treatment with Trypsin (100:1 protein-to-enzyme ratio) (Serva). Finally, resulting peptides were desalted by stop and go extraction (STAGE) tips (Rappsilber et al., 2003) before LC-MS/MS analysis.

Mass Spectrometry

LC-MS/MS analysis was performed using an UHPLC system (EASY-nLC 1000, Thermo Fisher Scientific) and a QExactive HF Orbitrap mass spectrometer (Thermo Fisher Scientific) as already described (Worzfeld et al., 2018). The same parameters were used except for the gradient of the reverse-phase chromatography: peptides were separated using a linearly increasing concentration of solvent B (80% acetonitrile, 0.1% formic acid) over solvent A (0.1% formic acid) from 5 to 30% for 215 min and from 30 to 60% for 5 min, followed by washing with 95% of solvent B for 5 min and re-equilibration with 5% of solvent B.

Evaluation of Proteome Data

MS raw data were processed by MaxQuant (1.5.6.5) (Cox and Mann, 2008) and the implemented Andromeda search engine using a Uniprot database of *E. coli* (strain K12) containing 4,306 entries (release 2017-01). The following parameters were used for data processing: maximum of two miss cleavages and mass tolerance of 4.5 ppm for main search. Trypsin was set as digesting enzyme. As fixed modification we used carbamidomethylation of cysteins and variable modifications were defined as oxidation of methionine and acetylation of the protein N-terminus. Beside these default parameters of MaxQuant, "label-free quantification" (LFQ) was enabled for protein quantification. For downstream data analysis, only proteins with at least two peptides and at least one unique peptide were considered as identified and processed with Perseus (1.5.6.0) to calculate *P*-values based on Benjamini–Hochberg multiple testing correction with a FDR threshold of 0.05. Processed data were evaluated using R statistical language (http://www.r-project.org/) and DAVID bioinformatics database (Huang et al., 2009) (https://david.abcc.ncifcrf.gov/home.jsp). Principal component analysis (PCA) was applied to LFQ intensities ($\log_2$) using R function *prcomp*. Package "factoextra" was used for visualization of PCA plots. Significantly regulated proteins were clustered according to functional annotations using the DAVID bioinformatics database. Gene ontology (GO) terms BP (biological process), CC (cellular component), MF (molecular function), KEGG pathway information (https://www.genome.jp/kegg/), and InterPro protein domains (https://www.ebi.ac.uk/interpro/) were selected to identify enriched clusters. The proteome data can be found in **Supplementary Table 3**.

Persister Assays

Persister levels were calculated by plating serial dilutions of cultures before and after antibiotic treatment. Exponential phase cultures (OD_{600} 0.3–0.6) were treated with 200 µg ml^{-1} ampicillin or 10 µg ml^{-1} ciprofloxacin, and subsequently incubated for 3 h at 37°C and 180 rpm. Stationary phase cultures (20 h after inoculation) were treated with 10 µg ml^{-1} ciprofloxacin for 5 h. Serial dilutions were prepared with NaCl solution (0.9%) and plated on LB agar without antibiotics. Agar plates were incubated at 37°C and colonies either counted after ∼20 h (before treatment) or ∼40 h (after treatment). Colony counts were used to calculate colony forming units (CFU) per milliliter. The persister level was the ratio between CFU ml^{-1} from treated samples and CFU ml^{-1} from untreated samples. *P*-values were calculated using Student's *t*-test.

ScanLag Analysis

Colony growth on LB agar plates was monitored using the ScanLag method (Levin-Reisman et al., 2010, 2014). Agar plates were covered with black felt and placed on Epson Perfection V39 scanners to record time series of images using the *ScanningManager* application. The equipment was placed in a 37°C incubation room. Images (stored as **.tif* files) were recorded every 20 min for a total time period of 40 h. Images were processed in MatLab (MathWorks) using functions *PreparePictures*, *setMaskApp*, and *TimeLapse*. Function *ScanLagApp* was called to assess the quality of single colonies.

Afterwards, appearance and growth times were extracted from the data. The appearance time of a colony equals the time point at which the colony has a minimum size of 10 pixels. The growth time of a colony is defined as the time needed to increase in size from 80 to 160 pixels. Growth inhibition by neighboring colonies can produce late appearance of individual colonies on densely populated plates (>100 colonies), which would falsely skew distribution plots toward late appearance times. In most cases, these colonies exhibit extraordinarily long growth times as well. Therefore, raw data were corrected based on the interquartile range of growth times: the upper limit L was defined as $L = Q_3 + 1.5 * (Q_3 - Q_1)$, with Q_1 and Q_3 being the lower and upper quartile, respectively. All colonies with growth times above L were removed from the analysis. For statistical evaluation of growth parameters R statistical language (http://www.r-project.org/) was used. Shapiro–Wilk test (function *shapiro.test* from package "stats") was called to analyze the distribution of appearance and growth times, demonstrating that the data were not normally distributed. Mann–Whitney-Wilcoxon test (function *wilcox.test* from package "stats") was subsequently used to calculate *P*-values.

RESULTS

Assessing the Persister Proteome by a Pulsed-SILAC Approach

Ampicillin and other β-lactam antibiotics cause rapid lysis of actively growing cells due to inhibition of cell wall biosynthesis. A treatment with ampicillin will, therefore, efficiently eliminate the majority of cells in exponential cultures of ampicillin-sensitive bacteria, and is considered as a simple method to enrich persister fractions (Keren et al., 2004). Here, we treated the

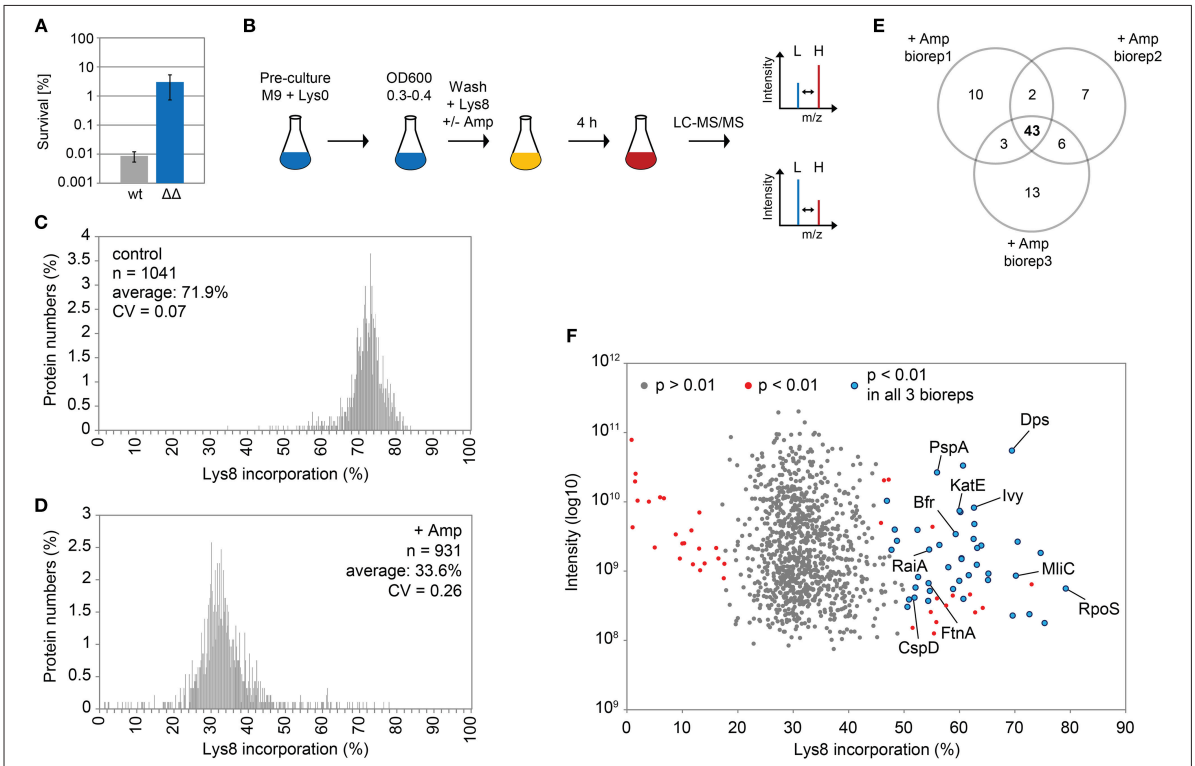


FIGURE 1 | Pulsed-SILAC during ampicillin treatment reveals newly synthesized proteins. **(A)** Persister levels after 4 h of ampicillin treatment (200 μg ml^{-1}) during exponential growth phase in M9+glu. CFU counts before and after ampicillin treatment were used to calculate survival. Data represent the mean of at least three independent biological replicates for wild type MG1655 (wt) and Δ1-41 Δ*istR* (ΔΔ). Error bars depict standard deviations. **(B)** Illustration of the pulsed-SILAC approach. Pre-cultures were grown in M9+glu in the presence of regular L-lysine (Lys0; 30 μg ml^{-1}). When exponential growth phase was reached, Lys0 was removed by washing, and cultures were incubated in the presence of L-lysine-$^{13}C_6$, $^{15}N_2$ (Lys8; 30 μg ml^{-1}) and ampicillin (200 μg ml^{-1}) for 4 h. A control culture was treated with Lys8, but allowed to grow in the absence of ampicillin. Protein samples were analyzed by LC-MS/MS to determine heavy (H) to light (L) protein ratios. H/L ratios were subsequently used to calculate Lys8 incorporation in percent. **(C)** Distribution plot of Lys8 incorporation (%) for the control culture. The number of quantified proteins (*n*), the average Lys8 incorporation and coefficient of variation (CV) are indicated. **(D)** Distribution plot of Lys8 incorporation (%) for ampicillin-treated cultures. Data represent the average of three biological replicates (**Supplementary Figure 1**). The number of quantified proteins (*n*), the overall average of Lys8 incorporation and coefficient of variation (CV) are indicated. **(E)** Venn diagram depicting proteins with significantly increased Lys8 incorporation (Significance B analysis, $P < 0.01$). Data were derived from ampicillin treatments in three biological replicates (biorep1-3). **(F)** Illustration of Significance B analysis for biorep1. Red dots represent proteins with a significantly altered ($P < 0.01$) Lys8 incorporation (%). Blue dots: $P < 0.01$ in all three biological replicates. Gray dots: $P > 0.01$.

TABLE 1 | Proteins with high Lys8 incorporation upon ampicillin treatment.

Protein name	Lys8 incorporation (%)					Significance B (*P* value)				Description
	Average	Biorep1	Biorep2	Biorep3	Control	Biorep1	Biorep2	Biorep3	Control	
RpoS	77.69	79.17	76.59	77.32		1.95E-60	4.69E-45	4.99E-55		RNA polymerase sigma factor RpoS
YmgG	76.30	75.36	77.23	76.30		3.51E-37	7.10E-49	2.67E-48		UPF0757 protein YmgG
YgdI	73.65	70.50	75.48	74.97		1.22E-38	3.35E-39	6.13E-41		Uncharacterized lipoprotein YgdI
Cfa	71.69	74.66	67.07	73.35	80.44	9.39E-63	5.72E-15	1.28E-53	0.0067	Cyclopropane-fatty-acyl-phospholipid synthase
MliC	70.68	70.19	70.06	71.79		1.86E-20	8.52E-21	5.40E-28		Membrane-bound lysozyme inhibitor of C-type lysozyme
Ugd	69.71	69.60	70.81	68.71		3.51E-19	1.27E-22	1.02E-19		UDP-glucose 6-dehydrogenase
YcfJ	69.37	72.63	69.00	66.47		6.13E-27	1.76E-18	2.06E-15		Uncharacterized protein YcfJ
Dps	67.93	69.47	67.76	66.57	80.83	2.04E-34	9.35E-26	1.49E-24	0.0038	DNA protection during starvation protein
BioB	67.18	65.15	66.53	69.85		5.80E-12	4.00E-14	2.02E-22		Biotin synthase
YiaD	66.55	65.23	66.76	67.67		4.68E-12	1.77E-14	1.41E-17		Probable lipoprotein YiaD
LpxC	64.36	63.94	64.55	64.58	81.65	7.07E-19	4.43E-18	1.46E-12	8.81E-04	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosamine deacetylase
Ivy	63.65	62.62	64.65	63.70	77.86	2.13E-16	2.84E-18	7.11E-18	0.1016	Inhibitor of vertebrate lysozyme
Gmd	63.45	63.15	63.79	63.43		9.64E-10	1.46E-10	2.34E-17		GDP-mannose 4,6-dehydratase
TauA	62.15	63.21	61.20	62.02		1.81E-17	1.37E-12	7.00E-15		Taurine-binding periplasmic protein
SsuA	61.92	60.37	62.96	62.44		2.23E-07	1.09E-09	1.41E-15		Putative aliphatic sulfonates-binding protein
SsuE	61.36	60.71	60.56	62.80		1.25E-07	1.42E-07	2.15E-10		FMN reductase (NADPH)
LoiP	61.19	60.37	61.78	61.42	70.02	2.22E-07	1.41E-08	6.35E-14	0.5673	Metalloprotease LoiP
PlaP	61.14	57.99	61.61	63.82		7.61E-06	1.98E-08	1.37E-11		Low-affinity putrescine importer PlaP
YghA	61.08	62.56	60.08	60.60	78.35	2.70E-16	3.59E-11	9.81E-13	0.0687	Uncharacterized oxidoreductase YghA
OsmY	60.83	60.64	61.30	60.55	76.34	2.57E-13	9.93E-13	1.13E-12	0.2640	Osmotically-inducible protein Y
Spy	60.83	61.67	59.96	60.86	79.07	2.15E-08	4.01E-07	1.77E-08	0.0363	Spheroplast protein Y
ClpA	60.00	62.68	59.81	57.52	80.69	1.63E-16	7.60E-11	4.15E-09	0.0047	ATP-dependent Clp protease ATP-binding subunit ClpA
Bfr	59.89	59.32	60.61	59.74	57.94	1.33E-11	8.17E-12	1.34E-11	0.0179	Bacterioferritin
Sbp	59.68	60.17	59.34	59.52	74.90	1.13E-12	2.57E-10	2.50E-11	0.5276	Sulfate-binding protein
YbgS	59.52	59.97	59.42	59.15	79.39	4.27E-07	9.57E-07	4.22E-07	0.0259	Uncharacterized protein YbgS
KatE	58.49	60.02	57.97	57.49	74.35	1.75E-12	6.55E-09	4.43E-09	0.5987	Catalase HPII
ZntA	57.92	58.84	56.42	58.49		2.38E-06	5.39E-05	1.24E-06		Lead, cadmium, zinc and mercury-transporting ATPase
LipA	56.42	54.37	57.10	57.79		3.93E-04	2.42E-05	3.62E-06		Lipoyl synthase
PspA	56.08	55.96	56.38	55.90	76.82	3.59E-08	1.63E-07	1.21E-07	0.2154	Phage shock protein A
YebE	55.69	56.39	56.07	54.61		1.53E-08	2.86E-07	1.22E-06		Inner membrane protein YebE
YciE	54.48	54.65	55.14	53.65	76.85	3.03E-04	2.11E-04	4.79E-04	0.2119	DUF892 domain-containing protein YciE
GatZ	54.19	48.73	56.51	57.31		7.58E-04	4.86E-05	7.11E-06		D-tagatose-1,6-bisphosphate aldolase subunit GatZ
RaiA	53.93	54.53	52.91	54.35	80.98	4.84E-07	3.43E-05	1.87E-06	0.0029	Ribosome-associated inhibitor A
HemA	53.90	50.61	54.58	56.51		0.0065	3.64E-04	2.08E-05		Glutamyl-tRNA reductase
CspD	53.74	51.89	55.83	53.49	78.89	0.0028	1.04E-04	5.55E-04	0.0437	Cold shock-like protein CspD
YfdC	52.46	52.53	53.33	51.50		0.0018	0.0011	0.0029		Inner membrane protein YfdC
Ahr	52.25	52.07	52.60	52.09	61.94	0.0025	0.0019	0.0018	0.0576	Aldehyde reductase Ahr
ArgF	51.86	47.73	53.12	54.72	57.32	0.0018	2.61E-05	1.02E-06	0.0151	Ornithine carbamoyltransferase chain F

(Continued)

TABLE 1 | Continued

Protein name	Lys8 incorporation (%)					Significance B (*P* value)				Description
	Average	Biorep1	Biorep2	Biorep3	Control	Biorep1	Biorep2	Biorep3	Control	
FtnA	51.69	54.44	49.24	51.41		3.69E-04	0.0017	0.0031		Bacterial non-heme ferritin
YibT	50.54	50.91	50.59	50.11	80.27	0.0054	0.0078	0.0076	0.0086	Uncharacterized protein YibT
PoxB	49.67	52.44	48.73	47.85	76.55	1.16E-05	0.0025	0.0037	0.2366	Pyruvate dehydrogenase [ubiquinone]; Alpha-peptide
OtsA	48.51	48.32	49.00	48.20	77.95	0.0011	0.0020	0.0028	0.1002	Alpha, alpha-trehalose-phosphate synthase [UDP-forming]
IlvB	48.01	46.91	47.53	49.58	60.54	0.0035	0.0064	8.03E-04	0.0381	Acetolactate synthase isozyme 1 large subunit

highly persistent *E. coli* strain Δ1-41 Δ*istR* with a high dose of ampicillin (200 μg ml^{-1}) during exponential growth phase (OD_{600} ∼0.3) in M9 minimal medium, a situation in which up to 3% persister cells are formed (**Figure 1A**). Since wild-type cultures produce >300-fold less persisters (**Figure 1A**), the vast majority of persister cells in strain Δ1-41 Δ*istR* can be expected to depend on TisB-induced depolarization for their generation (Berghoff et al., 2017). In parallel to the ampicillin challenge, pulse-labeling with the stable isotope-containing amino acid L-lysine-$^{13}C_6$,$^{15}N_2$ (Lys8) was performed to assess its incorporation into newly synthesized proteins during persister formation. Since our *E. coli* strains are prototrophic, pre-cultures were incubated in the presence of regular L-lysine (Lys0) to prime uptake and utilization of the externally added amino acid (Fröhlich et al., 2013). Lys0 was washed out the medium just before pulse-labeling with Lys8 was started (**Figure 1B**). After 4 h of ampicillin challenge and pulse-labeling, persister cells were harvested and analyzed by LC-MS/MS to determine heavy to light (H/L) protein ratios according to existing SILAC protocol (Schwanhäusser et al., 2009). H/L ratios were subsequently used to calculate Lys8 incorporation for the corresponding proteins. As a control, one culture was pulse-labeled with Lys8, but allowed to grow in the absence of ampicillin (**Figure 1B**). The control experiment demonstrated efficiency of the pulse-labeling protocol, as judged from the average Lys8 incorporation of 71.9% after 4 h of labeling (**Figure 1C**). By contrast, Lys8 incorporation in the ampicillin-treated persister cells was clearly reduced in all three biological replicates, with average values ranging from 32.8 to 34.4% (**Supplementary Figure 1**). Since reproducibility of Lys8 incorporation between replicates was high (Pearson's Rho ≥ 0.95; **Supplementary Figure 2**), data were combined in a single distribution plot displaying an overall average Lys8 incorporation of 33.6% for 931 proteins quantified in all biological replicates (**Figure 1D**). Notably, the variability of Lys8 incorporation, as measured by the coefficient of variation (CV), was considerably higher in persister cells (CV = 0.26) than in the control (CV = 0.07), indicating an increased heterogeneity in protein synthesis. The reduction of Lys8 incorporation of >2-fold compared to the control shows that ampicillin-treated persister cells are impaired in protein synthesis.

TisB-Dependent Persister Cells Mount an Active Response to Ampicillin

We were especially interested in proteins that displayed a Lys8 incorporation significantly higher than the average, which was indicative of enhanced synthesis and functional importance upon ampicillin stress. For this purpose, Significance B analysis was performed for each biological replicate (Cox and Mann, 2008). A common group of 43 proteins with $P < 0.01$ emerged from all ampicillin experiments (**Figures 1E,F**; for details see **Table 1**). Only six of these proteins showed a significantly higher Lys8 incorporation in the control experiment, which applied to Cfa, Dps, LpxC, ClpA, RaiA, and YibT. *P*-values were however lower by several orders of magnitude in ampicillin experiments, with YibT being the only exception (**Table 1**). Functional annotation clustering using gene ontology (GO) terms, KEGG pathway information, and InterPro protein domains (Huang et al., 2009) failed to highlight any major functional category, except for a small cluster formed by FtnA, Bfr, and Dps. Ferritin (FtnA) and bacterioferritin (Bfr) are iron storage proteins, that bind major proportions of iron under both normal and high-iron conditions (Sevcenco et al., 2011). Dps is a multi-function protein involved in iron sequestration, detoxification of reactive oxygen species (ROS), and mechanical protection of DNA during stationary phase (Zeth, 2012). Interestingly, the ROS detoxifying enzyme KatE (catalase II) showed increased Lys8 incorporation as well. Transcription of both *dps* and *katE* is triggered by RpoS, the master regulator of the general stress response in *E. coli*. Intriguingly, RpoS had the highest Lys8 incorporation (77.7%) of all identified proteins (**Table 1**), indicating that the general stress response is switched on in TisB-dependent persister cells. Further proteins that are related to stationary phase were RaiA and CspD. RaiA (or protein Y) is a modulator of ribosome activity and associates with 70S ribosomes to inhibit translation initiation during cold shock and stationary phase (Vila-Sanjurjo et al., 2004). CspD is an inhibitor of DNA replication that is implicated in persister formation (Kim and Wood, 2010). CspD levels are high in stationary phase, but decline when growth resumes due to degradation by Lon protease (Langklotz and Narberhaus, 2011). Furthermore, the bifunctional protein PspA was found to exhibit an increased Lys8 incorporation. PspA

inhibits its own transcriptional activator PspF under normal conditions. As soon as envelope stress occurs, PspA releases PspF and associates with PspBC to preserve functions of the inner membrane (Manganelli and Gennaro, 2017). Two other identified proteins, Ivy and MliC, are inhibitors of vertebrate C-type lysozyme and protect peptidoglycan against hydrolyzing attack. The above mentioned proteins are highlighted in a scatter plot that illustrates the Significance B analysis for one of the biological replicates (**Figure 1F**). The remaining proteins (**Table 1**) represent, e.g., periplasmic chaperones (OsmY, Spy), further RpoS-dependent proteins (cyclopropane-fatty-acyl-phospholipid synthase Cfa, trehalose-6-phosphate synthase OtsA, pyruvate dehydrogenase PoxB), proteins with functions in sulfate and sulfonate utilization [sulfate transporter subunit Sbp, sulfonate ABC transporter periplasmic binding protein SsuA, NAD(P)H-dependent FMN reductase SsuE], and proteins with a role in production of the capsular polysaccharide colanic acid (GDP-mannose 4,6-dehydratase Gmd, UDP-glucose 6-dehydrogenase Ugd). As a conclusion, our pulsed-SILAC approach identified 43 proteins with enhanced synthesis upon ampicillin stress, representing 4.6% of all quantified proteins. Many of these proteins serve stress-related and protective functions, and likely represent an active response of persister cells to ampicillin.

Proteome Analysis of Recovering Persister Cells

In a follow-up experiment, the proteome of strain Δ1-41 Δ*istR* was analyzed during the recovery phase. When liquid cultures were treated with a high dose of ampicillin (200 μg ml^{-1}) during exponential phase (OD_{600} ~0.3) in LB medium, the optical density clearly dropped within 2 h to values below 0.1 due to lysis of non-persistent cells (**Figure 2A**). After ampicillin treatment, persister cells were centrifuged and washed to remove cell debris, and finally transferred to fresh medium without antibiotics to enable recovery. After ~2 h of recovery, a slight increase of the optical density was visible, and from there cultures resumed exponential growth to complete a persister cycle (**Figure 2A**). Protein samples from persister cycle experiments were analyzed by SDS-PAGE, revealing a distinctive change in the protein pattern upon ampicillin treatment compared to exponential phase (**Figure 2B**). The protein pattern remained unchanged until 2 h of recovery, but reversed to the exponential phase pattern after 3 h of recovery (**Figure 2B**). Samples for proteome analysis by mass spectrometry were collected in biological triplicates after 2 h of ampicillin challenge ("amp") and at early time points during recovery ("rec1" and "rec1.5"; **Figures 2A,B**). Protein samples were analyzed by LC-MS/MS and quantified using a label-free approach (Cox et al., 2014). The retained "label-free quantification" (LFQ) values represent normalized protein intensities and can be used as a proxy for protein abundance. Reliability of the label-free normalization approach was high, as judged from LFQ intensity distributions (**Supplementary Figure 3**). LFQ intensities were subsequently used to identify proteins with changed abundances between

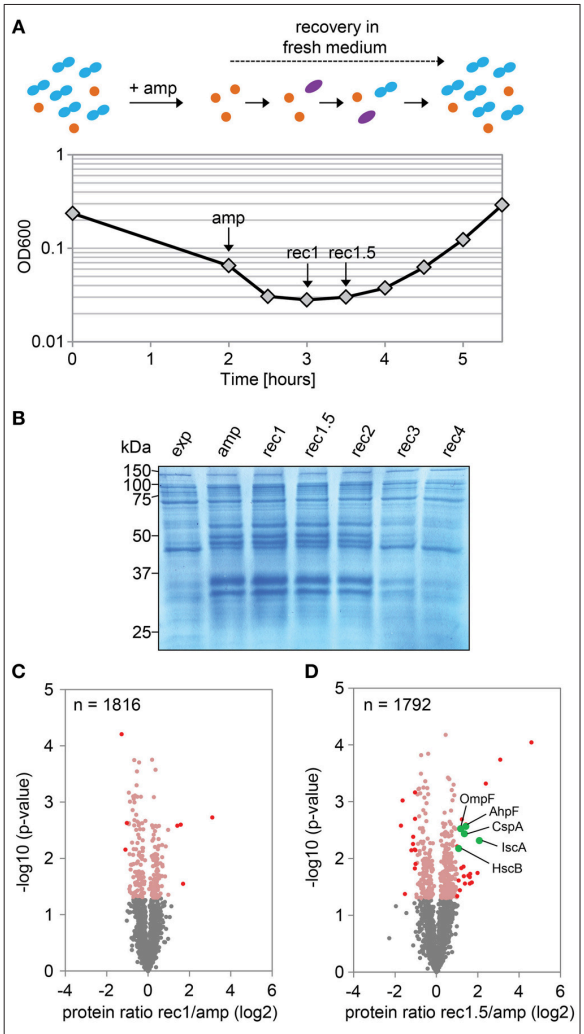


FIGURE 2 | Proteome analysis of persister cells during recovery from ampicillin treatment. **(A)** Graphical illustration of a persister cycle experiment with the highly persistent strain Δ1-41 Δ*istR*. Actively growing cells are blue, persisters are orange, and cells in the process of recovery are purple. Treatment with ampicillin (amp) during exponential growth phase leads to rapid lysis of non-persistent cells. After transfer of surviving persisters into recovery medium, cells eventually awake and resume growth. A representative growth curve is shown and time points for mass spectrometry analysis are indicated. **(B)** SDS-PAGE of protein samples from persister cycle experiments. Proteins (total amount of 50 μg per sample) were separated on 12% polyacrylamide gels and stained with colloidal Coomassie. A size marker (in kDa) is given on the left hand side. **(C,D)** Volcano plots for the comparison between recovery samples and ampicillin treatment. Data are derived from mass spectrometry analysis. Proteins with a significant change in abundance after 1 h **(B)** and 1.5 h **(C)** of recovery are in red (log_2 ratios >1 or < −1, $P < 0.05$). Pink dots: log_2 ratios <1 or >-1, $P < 0.05$. Gray dots: log_2 ratios <1 or >-1, $P > 0.05$. Exp: exponential phase; amp: 2 h ampicillin treatment; rec1: 1 h recovery; rec1.5: 1.5 h recovery; rec2: 2 h recovery; rec3: 3 h recovery; rec4: 4 h recovery.

TABLE 2 | Proteins with differential abundance during recovery.

Protein name	Protein ratio (log_2)		*P*-value		Description
	rec1/amp	rec1.5/amp	rec1/amp	rec1.5/amp	
RpmH	3.11	4.59	0.0019	9.07E−05	50S ribosomal protein L34
RpmF	1.58	3.08	0.0025	1.82E−04	50S ribosomal protein L32
RpmG	1.41	2.39	0.0026	4.79E−04	50S ribosomal protein L33
IscA	1.13	2.07	0.0707	0.0048	Iron-binding protein IscA
PrfA	1.70	1.99	0.0284	0.0180	Peptide chain release factor 1
YbjQ	0.81	1.71	0.0594	0.0264	UPF0145 protein YbjQ
SmrA	0.97	1.63	0.2685	0.0186	Probable DNA endonuclease SmrA
RlmJ	0.88	1.61	0.1092	0.0212	Ribosomal RNA large subunit methyltransferase J
QueD	1.05	1.61	0.1105	0.0276	6-carboxy-5,6,7,8-tetrahydropterin synthase
QueA	0.45	1.55	0.0302	0.0201	S-adenosylmethionine:tRNA ribosyltransferase-isomerase
AhpF	0.68	1.42	0.0058	0.0027	Alkyl hydroperoxide reductase subunit F
RpsU	0.70	1.42	0.0138	0.0025	30S ribosomal protein S21
SrlA	0.51	1.38	0.4402	0.0277	Glucitol/sorbitol permease IIC component
CspA	0.53	1.35	0.0227	0.0037	Cold shock protein CspA
YniC	0.27	1.33	0.3309	0.0204	2-deoxyglucose-6-phosphate phosphatase
Nth	0.62	1.31	0.1636	0.0139	Endonuclease III
YebC	0.54	1.23	0.0200	0.0020	Probable transcriptional regulatory protein YebC
RibB	−0.09	1.20	0.7793	0.0148	3,4-dihydroxy-2-butanone 4-phosphate synthase
OmpF	0.97	1.17	0.0031	0.0030	Outer membrane protein F
PurM	0.12	1.13	0.8329	0.0361	Phosphoribosylformylglycinamidine cyclo-ligase
MurB	0.16	1.07	0.6072	0.0244	UDP-N-acetylenolpyruvoylglucosamine reductase
HscB	0.62	1.07	0.0517	0.0067	Co-chaperone protein HscB
MetJ	0.31	1.00	0.3302	0.0061	Met repressor
PtsN	0.61	1.00	0.1086	0.0462	Nitrogen regulatory protein
GlnP	−0.78	−1.01	0.0130	0.0125	Glutamine transport system permease protein GlnP
YebF	−1.02	−1.01	0.0024	0.0071	Protein YebF
Ivy	−0.93	−1.03	6.78E−04	6.86E−04	Inhibitor of vertebrate lysozyme
SeqA	−1.28	−1.04	6.22E−05	0.0020	Negative modulator of initiation of replication
NuoK	−0.13	−1.04	0.6921	0.0150	NADH-quinone oxidoreductase subunit K
LivJ	−0.64	−1.05	0.0745	0.0069	Leu/Ile/Val-binding protein
RraA	−0.43	−1.12	0.1030	0.0042	Regulator of ribonuclease activity A
MalM	−0.90	−1.13	0.0084	0.0055	Maltose operon periplasmic protein
RcnB	−0.57	−1.21	0.0516	0.0072	Nickel/cobalt homeostasis protein RcnB
FxsA	−0.44	−1.52	0.0231	0.0423	UPF0716 protein FxsA
YmgD	–	−1.64	–	0.0010	Uncharacterized protein YmgD
CsrD	−0.88	−1.71	0.0392	0.0026	RNase E specificity factor CsrD

conditions (log_2 protein ratios >1 or <-1, $P < 0.05$). We reasoned that proteins, which exhibit differential expression during the early stages of recovery, are potentially aiding the recovery process itself. Therefore, samples from the early recovery phase ("rec1" and "rec1.5") were compared to samples taken during the ampicillin treatment ("amp"). The proteome pattern after 1 h of recovery was almost identical to the pattern of ampicillin-treated cells, and only seven proteins showed a significant change matching our criteria (**Table 2** and **Figure 2C**). However, after 1.5 h in recovery medium, 24 proteins were significantly increased and 12 proteins were significantly reduced in abundance (**Table 2** and **Figure 2D**). Importantly, all 36 proteins showed the same direction of regulation after 1 h of recovery, albeit changes in abundance were less pronounced (**Supplementary Figure 4**). LFQ intensities from all samples were applied to principal component analysis (PCA) to reduce complexity of the data for illustration in a two-dimensional plot. The first dimension reveals the progressive separation of recovery samples from ampicillin samples, while the second dimension most likely reflects variation between replicates (**Supplementary Figure 5**). We conclude that (i) the TisB-dependent and ampicillin-challenged "persister proteome" undergoes minor changes in a gradual manner within the first 1.5 h of postantibiotic recovery, and that (ii) the small subset

TABLE 3 | Doubling times of deletion strains in liquid medium.

Strain	Doubling time [min] Mean ± SD	*P*-value
ΔΔ	27.4 ± 0.7	
ΔΔΔ*hscB*	32.5 ± 0.1	1.86E-06
ΔΔΔ*iscA*	32.7 ± 0.9	1.35E-05
ΔΔΔ*cspA*	27.0 ± 0.2	0.2765
ΔΔΔ*ahpF*	27.6 ± 0.1	0.6727
ΔΔΔ*ompF*	27.4 ± 0.3	0.9441

*Doubling times were calculated from exponential growth phase in liquid LB medium (**Figure 3**). Values represent the mean and standard deviation (SD) of at least three independent biological replicates. Statistical testing using Student's t-test (P-value) refers to double deletion strain Δ1-41 ΔistR (ΔΔ).*

of 36 proteins might represent functions important to the recovery process.

Selection of Proteins From the Recovery Phase for Further Analysis

We selected five candidates from the 24 proteins, which were increased in abundance during recovery (**Figure 2D**). The pulsed-SILAC approach partly guided the selection process: increased synthesis of catalase KatE during ampicillin treatment (**Figure 1F**) was indicative of oxidative stress caused by hydrogen peroxide, and increased synthesis of iron storage proteins FtnA, Bfr, and Dps (**Figure 1F**) implied that excess intracellular iron might have originated from decomposition of iron-sulfur clusters (Fe/S). We therefore selected AhpF, a component of the alkyl hydroperoxide reductase involved in peroxide detoxification, and two proteins with functions in Fe/S assembly, HscB and IscA. The transcriptional activator CspA was selected because its mRNA levels are highest immediately before cell division starts (Brandi et al., 2016), implying an important function for growth resumption. Finally, the outer membrane porin OmpF was selected because it plays a major role for influx of nutrients, antibiotics, and other small compounds (Nikaido, 2003; Pagès et al., 2008).

AhpF and OmpF Specifically Affect the Colony Appearance Time After Antibiotic Treatment

We constructed gene deletions for *hscB*, *iscA*, *cspA*, *ahpF*, and *ompF* in our highly persistent strain Δ1-41 Δ*istR*. The resulting triple deletion strains were tested with regard to growth and persistence. The doubling time of strain Δ1-41 Δ*istR* was 27.4 min during exponential phase in liquid LB medium. Deletion of *hscB* and *iscA* resulted in significantly increased doubling times of 32.5 and 32.7 min, respectively (**Table 3** and **Figures 3A,B**). In addition, the optical density after 6 h of growth was lower in both strains (**Figures 3A,B**). We also determined the fraction of surviving cells after 3 h of ampicillin treatment (200 μg ml^{-1}) during exponential phase, to test whether persister levels were affected in the triple deletion strains. The persister level of strain Δ1-41 Δ*istR* was determined as 7.5%. Deletion of *hscB* and *iscA* resulted in significantly decreased persister levels of 1.5 and 0.9%, respectively (**Figure 4A**). By contrast, doubling times and persister levels for the remaining deletion strains (*cspA*, *ahpF*, and *ompF*) were largely unaffected (**Table 3** and **Figures 3C–E**, **4A**). We next asked the question whether the five genes have an effect on the time individual cells remain in the persistent state after an antibiotic challenge. Liquid cultures of the deletion strains were treated with ampicillin (200 μg ml^{-1}) for 3 h and subsequently plated on LB agar without antibiotics to enable recovery of the surviving persister cells and formation of colonies. Colony growth was analyzed by the ScanLag method, which has been developed to simultaneously measure the appearance and growth times of hundreds of colonies on agar plates (Levin-Reisman et al., 2010, 2014). The appearance time indicates the very first detection event of a colony, which is represented by a colony size of 10 pixels in the scanned images. The growth time reflects an increase in colony size from 80 to 160 pixels (see section Materials and Methods). If the growth time of a colony is largely unaffected, the appearance time mainly depends on the time the colony-forming cell needs to recover and, therefore, reflects the persistence time. The median appearance time of strain Δ1-41 Δ*istR* ranged between 880 and 940 min and the median growth time was mostly 140 min (**Table 4**). Deletion of *hscB* and *iscA* resulted in a significant shift to later appearance and growth times, with colonies appearing on average after more than 1,100 min with a growth time of 180 min (**Table 4** and **Figures 4B,C**). These findings validated a general growth defect for the *hscB* and *iscA* deletions, which was apparent both on agar plates and in liquid medium (**Tables 3**, **4**). By contrast, and consistent with measurements in liquid LB medium, deletions of *cspA*, *ahpF*, and *ompF* did not extend the growth time of colonies (**Figures 4D–F**). The *ahpF* deletion even had a reduced growth time of 120 min (**Table 4**). However, the appearance time was shifted to later time points in all three deletion strains, which was particularly evident for *ahpF* and *ompF* deletions. Both strains lost a major fraction of colonies with an early appearance time and gained colonies with an appearance time later than 1,400 min (**Figures 4E,F**). Since growth time was not affected, the later appearance time indicated a delay in outgrowth of persister cells.

We repeated the experiment with the fluoroquinolone antibiotic ciprofloxacin and treated exponentially growing cultures with a high dose (10 μg ml^{-1}) for 3 h. The *cspA* deletion was not further investigated, since it has only caused a small effect on the colony appearance time after treatment with ampicillin. The persister level of strain Δ1-41 Δ*istR* was determined as 3.2% for ciprofloxacin, and a significant decrease to 0.4% was only observed for the *hscB* deletion (**Figure 5A**). After ciprofloxacin treatment, all strains had a median growth time on LB agar which was comparable to growth times obtained after ampicillin treatment. However, the appearance time distributions after ciprofloxacin treatment were overall shifted to later time points, when individual strains were compared to the respective ampicillin experiment (**Table 4**; compare **Figures 4**, **5**). In case of ciprofloxacin, deletion of *hscB* and *iscA* caused a shift of ≥140 min for the median appearance time and ≥60 min for the median growth time in comparison to strain Δ1-41

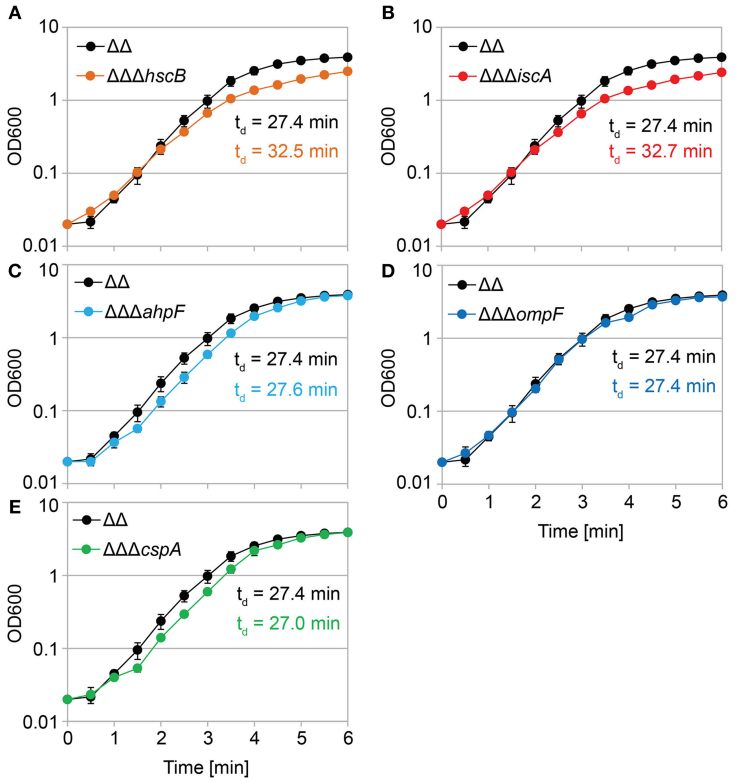


FIGURE 3 | Growth in liquid medium is impaired by *hscB* and *iscA* deletions. Growth curves of double deletion strain Δ1-41 Δ*istR* (ΔΔ) and corresponding triple deletion strains in liquid LB medium. The OD_{600} was initially adjusted to 0.02 from over-night cultures and measured over 6 h. Doubling times (t_d) during exponential phase are given for the corresponding strains in each graph. **(A)** ΔΔ vs. ΔΔΔ*hscB*, **(B)** ΔΔ vs. ΔΔΔ*iscA*, **(C)** ΔΔ vs. ΔΔΔahpF, **(D)** ΔΔ vs. ΔΔΔ*ompF*, and **(E)** ΔΔ vs. ΔΔΔ*cspA*. Data represent the mean of at least three independent biological replicates. Error bars depict standard deviations.

Δ*istR* (**Table 4** and **Figures 5B,C**). By contrast, *ahpF* and *ompF* deletions caused a delay in colony appearance (70-min shift of the median appearance time), without severely affecting the growth time (**Table 4** and **Figures 5D,E**).

As a conclusion, the growth phenotypes obtained for the *hscB*, *iscA*, *ahpF*, and *ompF* deletions were not specific to a distinct antibiotic, but instead applied to at least two different classes of antibiotics (β-lactams and fluoroquinolones). Furthermore, deletion of *ahpF* and *ompF* specifically affected the colony appearance time after antibiotic treatment in strain Δ1-41 Δ*istR*, possibly due to impairment of the recovery process.

Functions of AhpF and OmpF During Recovery Are Specific to TisB-Dependent Persister Cells

We were curious whether our findings were influenced by the genetic background of strain Δ1-41 Δ*istR*. To answer this question, *hscB*, *iscA*, *ahpF*, and *ompF* deletions were constructed in the MG1655 wild-type background. Strains were treated with ampicillin (200 μg ml^{-1}) for 3 h during exponential phase, as before, and analyzed by ScanLag. Interestingly, none of the deletions caused a reduction in persister levels (**Figure 6A**). While *hscB* and *iscA* deletions caused a shift to later appearance and growth times compared to wild type MG1655 (**Table 4** and **Figures 6B,C**), growth parameters of *ahpF* and *ompF* deletion strains were almost identical to those obtained for the wild type (**Table 4** and **Figures 6D,E**). The influence of *ahpF* and *ompF* on the colony appearance time (after ampicillin treatment) was therefore no general feature and appeared to be specific for TisB-dependent persister cells.

To further corroborate our findings, *ahpF* and *ompF* deletions were constructed in a Δ*tisB* background and analyzed as before. Deletion of *ahpF* did not affect the persistence time, while deletion of *ompF* even caused an earlier appearance of colonies (**Table 4** and **Figure 7A**). In a next experiment, wild-type *ahpF* and *ompF* deletion strains were treated with ciprofloxacin (10 μg ml^{-1}) during stationary phase for 5 h. ScanLag analysis revealed that the colony appearance time of stationary phase persisters was neither affected by *ahpF* nor *ompF* deletions (**Table 4** and **Figure 7B**). These experiments confirmed that AhpF and OmpF specifically affect recovery of

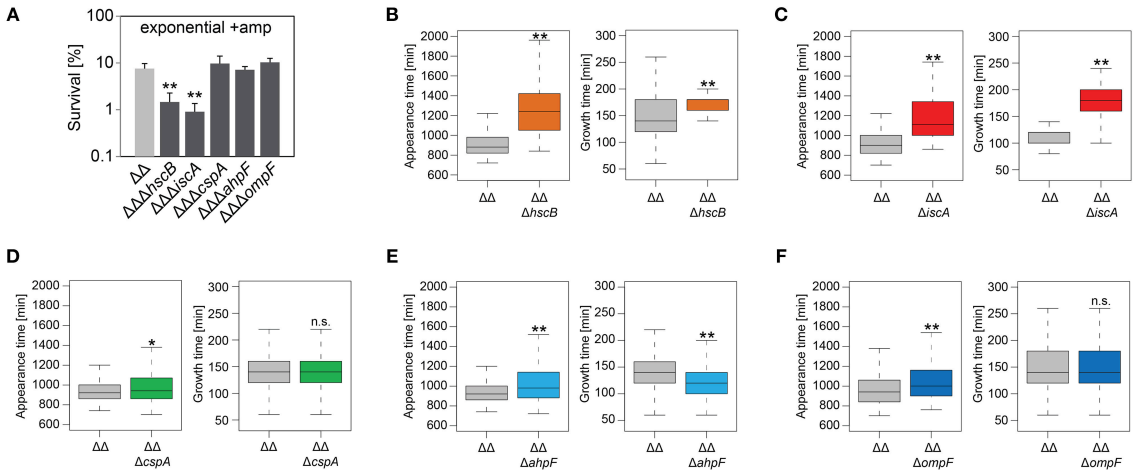


FIGURE 4 | Influence of candidate proteins on survival and recovery of TisB-dependent persister cells after ampicillin treatment. Exponential cultures of double deletion strain Δ1-41 Δ*istR* (ΔΔ) and corresponding triple deletion strains were treated for 3 h with ampicillin (+amp; 200 μg ml^{-1}) in liquid LB medium and subsequently plated on LB agar without antibiotics. **(A)** CFU counts before and after ampicillin treatment were used to calculate survival. Data represent the mean of at least three independent biological replicates. Error bars depict standard deviations. *P*-values were calculated using Student's *t*-test (**$P < 0.01$). **(B–F)** The ScanLag method (Levin-Reisman et al., 2010, 2014) was applied to monitor appearance and growth times of individual colonies after ampicillin treatment. Left and right panels show boxplots for appearance and growth times, respectively, for **(B)** ΔΔ ($n = 294$) vs. ΔΔΔ*hscB* ($n = 139$), **(C)** ΔΔ ($n = 191$) vs. ΔΔΔ*iscA* ($n = 102$), **(D)** ΔΔ ($n = 190$) vs. ΔΔΔ*cspA* ($n = 203$), **(E)** ΔΔ ($n = 190$) vs. ΔΔΔ*ahpF* ($n = 194$), and **(F)** ΔΔ ($n = 399$) vs. ΔΔΔ*ompF* ($n = 604$). *P*-values were calculated using Mann–Whitney-Wilcoxon test (**$P < 0.01$; *$P < 0.05$; n.s., not significant).

TABLE 4 | Growth parameters of deletion strains on solid medium after antibiotic treatment.

Condition	Appearance time [min]					Growth time [min]				
	Control strain	Median	Deletion strain	Median	*P*-value	Control strain	Median	Deletion strain	Median	*P*-value
Exponential +amp	ΔΔ	880	ΔΔΔ*hscB*	1,240	<2.2E-16	ΔΔ	140	ΔΔΔ*hscB*	180	3.01E-10
	ΔΔ	900	ΔΔΔ*iscA*	1,110	<2.2E-16	ΔΔ	120	ΔΔΔ*iscA*	180	<2.2E-16
	ΔΔ	920	ΔΔΔ*cspA*	940	0.0363	ΔΔ	140	ΔΔΔ*cspA*	140	0.4227
	ΔΔ	920	ΔΔΔ*ahpF*	980	2.37E-06	ΔΔ	140	ΔΔΔ*ahpF*	120	8.01E-05
	ΔΔ	940	ΔΔΔ*ompF*	1,000	5.75E-10	ΔΔ	140	ΔΔΔ*ompF*	140	0.3795
Exponential +CF	ΔΔ	1,480	ΔΔΔ*hscB*	1,650	4.77E-07	ΔΔ	140	ΔΔΔ*hscB*	200	<2.2E-16
	ΔΔ	1,480	ΔΔΔ*iscA*	1,620	0.0026	ΔΔ	140	ΔΔΔ*iscA*	220	<2.2E-16
	ΔΔ	1,290	ΔΔΔ*ahpF*	1,360	0.0294	ΔΔ	120	ΔΔΔ*ahpF*	120	0.2841
	ΔΔ	1,290	ΔΔΔ*ompF*	1,360	0.0355	ΔΔ	120	ΔΔΔ*ompF*	140	0.0039
Exponential +amp	wt	760	Δ*hscB*	900	<2.2E-16	wt	140	Δ*hscB*	200	<2.2E-16
	wt	760	Δ*iscA*	960	<2.2E-16	wt	140	Δ*iscA*	220	<2.2E-16
	wt	820	Δ*ahpF*	820	0.5294	wt	140	Δ*ahpF*	140	0.1106
	wt	760	Δ*ompF*	760	0.1024	wt	140	Δ*ompF*	140	0.6671
Exponential +amp	Δ*tisB*	820	Δ*tisB* Δ*ahpF*	820	0.2650	Δ*tisB*	160	Δ*tisB* Δ*ahpF*	120	<2.2E-16
	Δ*tisB*	820	Δ*tisB* Δ*ompF*	780	5.25E-13	Δ*tisB*	160	Δ*tisB* Δ*ompF*	140	2.89E-05
Stationary +CF	wt	1,080	Δ*ahpF*	1,080	0.6382	wt	140	Δ*ahpF*	140	0.7855
	wt	1,080	Δ*ompF*	1,030	0.1983	wt	140	Δ*ompF*	120	1.41E-05
Exponential +amp	ΔΔ	960	ΔΔΔ*ahpF*	1,140	8.24E-13	ΔΔ	140	ΔΔΔ*ahpF*	140	0.0017
	ΔΔ	960	ΔΔΔ*ahpF* Δ*ompF*	1,000	0.0004	ΔΔ	140	ΔΔΔ*ahpF* Δ*ompF*	120	<2.2E-16

*Appearance and growth times represent median values of ScanLag experiments (**Figures 4–7**). Statistical testing using Mann–Whitney-Wilcoxon test (P-value) refers to the corresponding control strain from the same experimental run.*

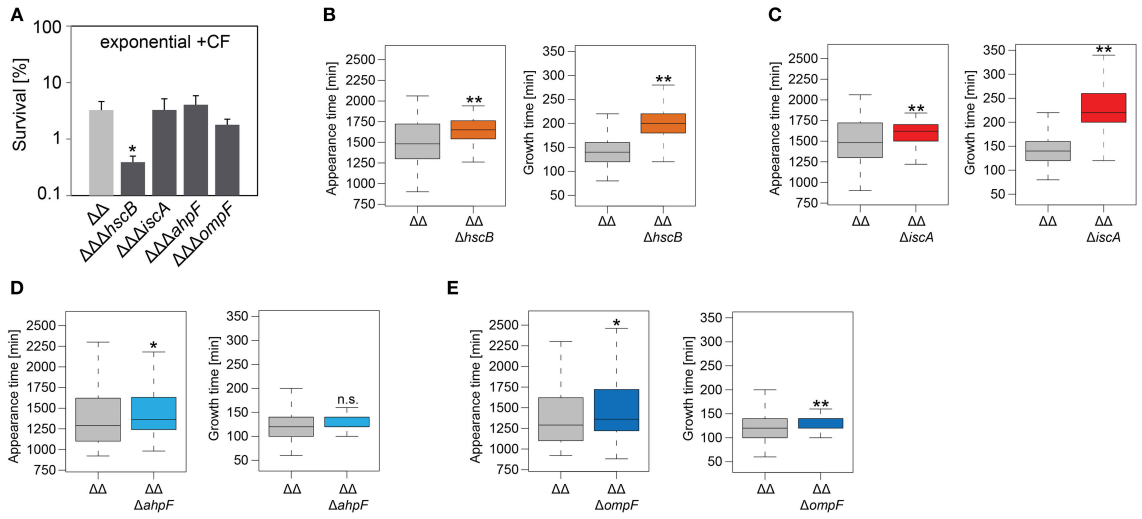


FIGURE 5 | Influence of candidate proteins on survival and recovery of TisB-dependent persister cells after ciprofloxacin treatment. Exponential cultures of double deletion strain Δ1-41 Δ*istR* (ΔΔ) and corresponding triple deletion strains were treated for 3 h with ciprofloxacin (+CF; 10 μg ml^{-1}) in liquid LB medium and subsequently plated on LB agar without antibiotics. **(A)** CFU counts before and after ciprofloxacin treatment were used to calculate survival. Data represent the mean of at least three independent biological replicates. Error bars depict standard deviations. *P*-values were calculated using Student's *t*-test (*$P < 0.05$). **(B–E)** The ScanLag method (Levin-Reisman et al., 2010, 2014) was applied to monitor appearance and growth times of individual colonies after ciprofloxacin treatment. Left and right panels show boxplots for appearance and growth times, respectively, for **(B)** ΔΔ ($n = 133$) vs. ΔΔΔ*hscB* ($n = 210$), **(C)** ΔΔ ($n = 133$) vs. ΔΔΔ*iscA* ($n = 215$), **(D)** ΔΔ ($n = 118$) vs. ΔΔΔ*ahpF* ($n = 104$), and **(E)** ΔΔ ($n = 118$) vs. ΔΔΔ*ompF* ($n = 75$). *P*-values were calculated using Mann–Whitney-Wilcoxon test (**$P < 0.01$; *$P < 0.05$; n.s., not significant).

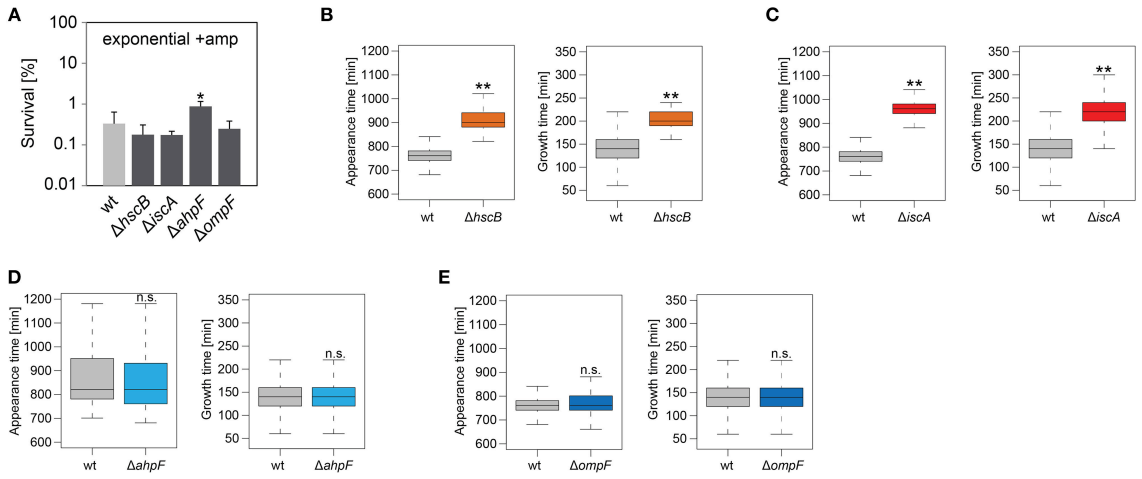


FIGURE 6 | Influence of candidate proteins on survival and recovery of wild-type persister cells after ampicillin treatment. Exponential cultures of wild type MG1655 (wt) and corresponding deletion strains were treated for 3 h with ampicillin (+amp; 200 μg ml^{-1}) in liquid LB medium and subsequently plated on LB agar without antibiotics. **(A)** CFU counts before and after ampicillin treatment were used to calculate survival. Data represent the mean of at least three independent biological replicates. Error bars depict standard deviations. *P*-values were calculated using Student's *t*-test (*$P < 0.05$). **(B–E)** The ScanLag method (Levin-Reisman et al., 2010, 2014) was applied to monitor appearance and growth times of individual colonies after ampicillin treatment. Left and right panels show boxplots for appearance and growth times, respectively, for **(B)** wt ($n = 80$) vs. Δ*hscB* ($n = 84$), **(C)** wt ($n = 80$) vs. Δ*iscA* ($n = 144$), **(D)** wt ($n = 224$) vs. Δ*ahpF* ($n = 219$), and **(E)** wt ($n = 276$) vs. Δ*ompF* ($n = 391$). *P*-values were calculated using Mann-Whitney-Wilcoxon test (**$P < 0.01$; n.s., not significant).

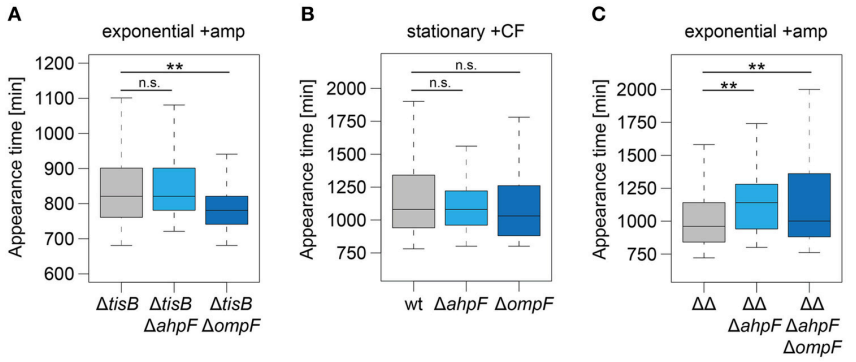


FIGURE 7 | The role of AhpF and OmpF in TisB-independent persisters and synergistic effects. Deletions of *ahpF* and *ompF* were constructed in a *tisB* deletion (Δ*tisB*), wild-type (wt), or Δ1-41 Δ*istR* (ΔΔ) background. Resulting strains were analyzed by ScanLag. **(A)** Cultures were treated with ampicillin (+amp; 200 μg ml^{-1}) for 3 h during exponential phase. Boxplots reflect appearance time distributions for Δ*tisB* ($n = 394$), Δ*tisB* Δ*ahpF* ($n = 204$), and Δ*tisB* Δ*ompF* ($n = 272$). **(B)** Cultures were treated with ciprofloxacin (+CF; 10 μg ml^{-1}) for 5 h during stationary phase. Boxplots reflect appearance time distributions for wt ($n = 113$), Δ*ahpF* ($n = 147$), and Δ*ompF* ($n = 58$). **(C)** Cultures were treated with ampicillin (+amp; 200 μg ml^{-1}) for 3 h during exponential phase. Boxplots reflect appearance time distributions for ΔΔ ($n = 325$), ΔΔΔ*ahpF* ($n = 233$), and ΔΔΔ*ahpF*Δ*ompF* ($n = 142$). *P*-values were calculated using Mann–Whitney-Wilcoxon test (**$P < 0.01$; n.s., not significant).

TisB-dependent persister cells, but excluded a more general role in recovery of persister cells that have formed through other mechanisms.

Simultaneous Deletion of *ahpF* and *ompF* Does not Cause a Synergistic Effect

Since ScanLag experiments with *ahpF* and *ompF* single deletions in strain Δ1-41 Δ*istR* indicated that both genes support the recovery process of TisB-dependent persister cells (**Figures 4**, **5**), we asked the question whether simultaneous deletion of both genes would cause a severe delay in colony appearance. To this end, a marker-less variant of strain Δ1-41 Δ*istR* was constructed by FLP-mediated recombination. Subsequent deletion of *ahpF*, resulting in a triple deletion strain, caused a prolonged appearance time of colonies after ampicillin treatment during exponential phase (**Figure 7C**), which confirmed our former findings (**Figure 4E**). Additional deletion of *ompF* produced a quadruple deletion strain, which was still highly persistent (**Supplementary Figure 6**). However, the colony appearance time in the quadruple deletion strain was not further increased in comparison to the triple deletion strain (**Figure 7C**). On the contrary, the median appearance time of the quadruple deletion strain was even intermediate between the initial strain (Δ1-41 Δ*istR*) and the triple deletion strain.

DISCUSSION

Antibiotic-tolerant persister cells increase the risk for relapsing infections, a threat especially prevalent in combination with bacterial biofilms (Lewis, 2007, 2010; Michiels et al., 2016). Assessing persister physiology is key to understanding the processes that drive bacterial persistence, which will ultimately guide development of therapeutic strategies. Several studies have assessed the persister transcriptome by microarray or RNA-seq analysis after enrichment of persister fractions by either lysis of non-persistent cells (Keren et al., 2004, 2011; Pu et al., 2016), or fluorescence-activated cell sorting (FACS) of reporter strains (Shah et al., 2006). While each of these enrichment methods has its limitations in terms of purity of the enriched persister fractions (Cañas-Duarte et al., 2014; Henry and Brynildsen, 2016), the transcriptome data provided conclusive insights into persister physiology. For example, pioneering work from the Lewis group highlighted the prevalence of toxin mRNAs in persister fractions and emphasized the importance of chromosomally encoded TA systems for bacterial persistence (Keren et al., 2004; Shah et al., 2006). However, alterations on transcript level do not necessarily affect the amount of a given protein due to posttranscriptional and posttranslational regulatory events. To complement the picture of persister physiology, high-throughput investigations of the proteome are needed, but have rarely been addressed.

Pulsed-SILAC Reveals Persister Physiology on the Translational Level

Pulsed-SILAC is a powerful tool to globally assess active protein translation (Schwanhäusser et al., 2009), but even though it has been successfully applied to study, e.g., protein synthesis in colistin-tolerant subpopulations of *Pseudomonas aeruginosa* biofilms (Chua et al., 2016), it has not yet been applied to planktonic persister cells during an antibiotic challenge. Here, a pulsed-SILAC approach was used to quantify protein synthesis in TisB-dependent persister cells after an ampicillin challenge during exponential phase (**Figure 1B**). The enriched persister fraction exhibited a >2-fold reduction in protein synthesis

compared to the exponentially growing control, as judged from the average Lys8 incorporation of 33.6 and 71.9%, respectively (**Figures 1C,D**). These findings are in line with the general notion that the likelihood of persister formation inversely correlates with the translational activity (Balaban et al., 2004; Shah et al., 2006; Orman and Brynildsen, 2013; Henry and Brynildsen, 2016). Additionally, Lys8 incorporation might be compromised in TisB-dependent persisters, since depolarization by TisB and subsequent ATP depletion (Unoson and Wagner, 2008; Gurnev et al., 2012) is expected to impede lysine uptake by ABC transporters. Even though protein synthesis is diminished on average, there is high variability in protein synthesis with 43 proteins exhibiting elevated levels of Lys8 incorporation (∼45–80%; **Figures 1E,F**). These proteins likely represent an active stress response in TisB-dependent persister cells. In an alternative and toxin-independent model of persistence, high levels of persister cells are formed after nutrient shifts, e.g., from glucose to fumarate. Quantitative proteomics revealed activation of a distinct stress response, mainly controlled by the sigma factor RpoS (Radzikowski et al., 2016). Interestingly, TisB-dependent persister cells strongly synthesize RpoS (Lys8 incorporation of 77.7%) and the RpoS-dependent proteins Dps, KatE, Cfa, OsmY, OtsA, and PoxB (**Figure 1F**), most of which serve protective functions. For some of them, enrichment in persister fractions has already been shown on the mRNA level by transcriptome analysis (Keren et al., 2004; Shah et al., 2006), which applies to KatE (catalase II), OsmY (periplasmic chaperone), OtsA (trehalose-6-phosphate synthase), and to the dual-function effector of the envelope stress response PspA. These proteins might represent a general hallmark of persister proteomes and serve as suitable biomarkers for persister cells in future studies. Other proteins found by our pulsed-SILAC approach might influence persistence directly through inhibition of replication (CspD) (Kim and Wood, 2010) or inhibition of translation (RaiA). In summary, we conclude that TisB-dependent persister cells mount an active response on the translational level, which has the potential to (i) actively protect persister cells from severe damage and (ii) induce or stabilize the persistent state. The question to which extent the proteins identified here contribute to TisB-dependent persistence—or persistence in general—needs to be addressed in future studies.

Assuming that the particular response to a stress factor (e.g., antibiotics or nutrient deprivation) depends on the molecular status (e.g., expression of TA systems or metabolic state) of a persister cell, it seems reasonable that the same stress would elicit different responses in different persister types. For example, stationary phase persisters might react in a different way to ciprofloxacin than depolarized persisters generated during exponential phase. Systematic investigations of the persister proteome from a set of defined persister populations might help to unravel shared and specialized features of stress responses in different persister types. Our pulsed-SILAC approach represents a first step toward this direction.

Postantibiotic Recovery of TisB-Dependent Persister Cells

Mechanisms that lead to persister formation are diverse but in many instances quite well-understood. Mechanisms that help bacteria to recover from the persistent state, however, are only beginning to be discovered, and can be classified as follows: (i) rescuing of targets, that have been corrupted by toxins, to enable awakening (De Jonge et al., 2009; Cheverton et al., 2016), (ii) intrinsic regulatory features of TA operons by a process called "conditional cooperativity" to regain inhibition of toxins by their cognate antitoxins (Page and Peti, 2016), and (iii) repair of antibiotic-induced damages during the postantibiotic recovery phase to maintain survival (Völzing and Brynildsen, 2015; Mok and Brynildsen, 2018). Here, we applied label-free quantitative MS to identify differentially expressed proteins in TisB-dependent persisters during recovery from ampicillin. Importantly, our sampling time points precede bulk growth resumption (**Figures 2A,B**), which is also reflected by the small proportion of proteins with decreased (12 proteins) or enhanced abundance (24 proteins) after 1.5 h of recovery (**Figure 2D**). At this time point, TisB persisters have obviously just started to remodel their proteome, and are—from a physiological point of view—still engaged in an intermediate state between persistence and growth resumption.

The proteomics snapshot identified proteins harboring interesting functions with respect to what we have learned from the pulsed-SILAC approach. Increased synthesis of catalase KatE during ampicillin treatment of strain Δ1-41 Δ*istR* (**Figure 1F**) indicates that TisB persisters need to detoxify hydrogen peroxide. Alternatively, increased KatG synthesis might be an inevitable consequence of activation of the general stress response by RpoS (**Figure 1F**) without increased ROS production. However, generation of ROS by antibiotics is a documented, albeit controversial, phenomenon (Kohanski et al., 2007; Dwyer et al., 2014), and we have reason to believe that TisB, and other depolarizing toxins, further enhance ROS production (our unpublished data). It is, therefore, feasible to assume that TisB persisters accumulate oxidative damage that needs to be repaired during postantibiotic recovery. In this regard, TisB-dependent persisters might be reminiscent of viable but non-culturable (VBNC) *E. coli* cells, that exhibit signatures of oxidative protein damage and have activated RpoS-dependent genes like *katE* (Desnues et al., 2003). But unlike VBNC cells, TisB persisters are able to produce colonies in a timely fashion, even though their appearance time is delayed due to high TisB levels (compare wt and ΔΔ in **Table 4**) (Berghoff et al., 2017). We observed increasing protein levels of AhpF, a component of alkyl hydroperoxide reductase Ahp, during recovery of strain Δ1-41 Δ*istR* (**Figure 2D**). Ahp is the primary scavenger of hydrogen peroxide under standard growth conditions (Imlay, 2013). Furthermore, Ahp has the potential to reduce a variety of alkyl hydroperoxides (Jacobson et al., 1989; Storz et al., 1989), and might be involved in repairing damaged molecules in TisB-dependent persister cells

during postantibiotic recovery. In line with this hypothesis, the colony appearance time after antibiotic treatment was prolonged in strain Δ1-41 Δ*istR* when *ahpF* was deleted (**Figures 4E**, **5D**). Since persister levels were not affected (**Figures 4A**, **5A**), the amount of oxidative damage in TisB persisters appears to be sublethal, maybe due to hydrogen peroxide detoxification by catalase KatE already during the persistent state. Interestingly, the effect of an *ahpF* deletion on the appearance time was not observed in a wild-type or Δ*tisB* background (**Figures 6D**, **7A**). Furthermore, an *ahpF* deletion did not extend the lag phase after diluting cells from stationary phase into fresh LB medium (**Figure 3C**). Together these data demonstrate that AhpF does not represent a crucial factor during outgrowth of *E. coli* in general. The specific importance of AhpF for recovery of TisB persisters rather supports the assumption that oxidative stress represents a particular threat for depolarized cells.

The pulsed-SILAC approach also demonstrated increased synthesis of iron storage proteins FtnA, Bfr, and Dps (**Figure 1F**), which was indicative of excess free iron within TisB persisters. Free iron might originate from decomposition of Fe/S by ROS (Imlay, 2006, 2008), and iron sequestration is, therefore, needed to avoid subsequent generation of genotoxic hydroxyl radicals through Fenton chemistry. In line with the proposed Fe/S decomposition, proteins with a function in Fe/S assembly were upregulated during recovery, which applies to the A-type Fe/S carrier protein IscA and chaperone HscB, which is involved in release of Fe/S from scaffold proteins (Roche et al., 2013). Growth defects have been reported for both *hscB* and *iscA* deletion strains (Tokumoto and Takahashi, 2001; Lu et al., 2008), which was also observed here in the Δ1-41 Δ*istR* background (**Table 3**) and for wild type MG1655 (data not shown). The general growth defect likely causes the strongly delayed appearance time of colonies in ScanLag experiments (**Figures 4**–**6**), and it is, therefore, difficult to draw conclusions about duration of the persistent state from these colony-based experiments. However, upon ampicillin treatment persister levels of strain Δ1-41 Δ*istR* were reduced by ∼5- and 8-fold due to *hscB* and *iscA* deletions, respectively (**Figure 4A**), which was not observed in a wild-type background (**Figure 6A**). We speculated that the reduced survival was caused by generation of hydroxyl radicals, which could not be confirmed in experiments with the hydroxyl radical scavenger thiourea (data not shown). We conclude that Fe/S assembly specifically supports the postantibiotic recovery process of TisB persisters to maintain survival, but cannot exclude an additional role in persister generation.

The Ability of OmpF to Influence Recovery Depends on the Physiological Condition

The intracellular concentration of an antibiotic is determined by two processes: influx and efflux. Efflux of antibiotics by bacterial pumps is mainly considered an important determinant for antibiotic resistance. It was only recently shown that *E. coli* persister cells extrude β-lactams by TolC-dependent pumps as an active defense and survival strategy (Pu et al., 2016). However, the parallel induction of *ompF*, as monitored on RNA level, was considered a paradox, since OmpF is a major entry gate for β-lactams and other antibiotics (Nikaido, 2003; Pagès et al., 2008). These results were interpreted as lack of cooperation between efflux and influx systems (Pu et al., 2016). Our data might provide a solution to this problem. One has to consider that OmpF and other porins do not only allow antibiotics to enter the periplasm, but also provide the nutrient supply that is needed for cell growth. If *ompF* is deleted in a Δ*tisB* background, cells recover more quickly after an ampicillin challenge (**Figure 7A**), likely because intracellular ampicillin accumulation is reduced. The situation reverses in TisB persisters: deleting *ompF* in the Δ1-41 Δ*istR* background causes a prolonged persistence time (**Figures 4F**, **5E**), indicating that OmpF is an important factor for recovery. Some nutrients, like sugars and other metabolites, have the potential to repolarize the inner membrane and reverse toxin-dependent depolarization (Allison et al., 2011; Verstraeten et al., 2015). Increased nutrient supply by OmpF upregulation might support this process. We conclude that high levels of OmpF are either detrimental or beneficial for persister cells, depending on the particular persistence mechanism and physiological condition. The surprising lack of synergistic effects in the quadruple deletion strain (Δ1-41 Δ*istR* Δ*ahpF* Δ*ompF*, **Figure 7C**) can be explained along the same lines. If AhpF is not present in TisB persisters, the physiological condition has changed. Now, an *ompF* deletion turns out to be beneficial for the recovery process. In summary, the functional importance of a particular protein during postantibiotic recovery of persister cells strongly depends on the physiological condition and is expected to show high variations on the single cell level among mixed persister populations.

CONCLUSIONS

Proteome analysis by state-of-the-art MS is a powerful tool to assess persister physiology, and was applied here to learn more about proteins with potential functions during postantibiotic recovery. TisB-dependent persisters were chosen as a model system for "persistence by depolarization." The investigated proteins with increased abundance during recovery fall into three classes: (i) proteins with no major impact, neither on persister level nor persistence time (CspA), (ii) proteins important for growth in general and persister survival in particular (HscB and IscA), and (iii) proteins with specific functions during recovery of TisB persisters (AhpF and OmpF). How expression of these proteins is regulated during the recovery phase remains an exciting question for future studies.

DATA AVAILABILITY

All datasets generated or analyzed for this study are included in the manuscript and/or the supplementary files.

AUTHOR CONTRIBUTIONS

AK and BB designed the study. D-TS performed most of the physiological experiments and established a workflow

for analysis of ScanLag data. DE and BB performed additional physiological experiments and analyzed the data. AK performed mass spectrometry, analyzed the data, and contributed to the material and methods section. BB wrote the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was supported by Fonds der Chemischen Industrie (material cost allowance to BB) and University of Giessen (Research Grant to BB).

ACKNOWLEDGMENTS

We are grateful to Gabriele Klug (University of Giessen) for critical reading of the manuscript. We thank Jochen Bathke and Timothy Bender (University of Giessen) for help with R statistical language and experimental support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2019.00378/full#supplementary-material

REFERENCES

Allison, K. R., Brynildsen, M. P., and Collins, J. J. (2011). Metabolite-enabled eradication of bacterial persisters by aminoglycosides. *Nature* 473, 216–220. doi: 10.1038/nature10069

Andersen, J. S., Lam, Y. W., Leung, A. K. L., Ong, S. E., Lyon, C. E., Lamond, A. I., et al. (2005). Nucleolar proteome dynamics. *Nature* 433, 77–83. doi: 10.1038/nature03207

Ayrapetyan, M., Williams, T. C., and Oliver, J. D. (2015). Bridging the gap between viable but non-culturable and antibiotic persistent bacteria. *Trends Microbiol.* 23, 7–13. doi: 10.1016/j.tim.2014.09.004

Balaban, N. Q., Merrin, J., Chait, R., Kowalik, L., and Leibler, S. (2004). Bacterial persistence as a phenotypic switch. *Science* 305, 1622–1625. doi: 10.1126/science.1099390

Berghoff, B. A., Hoekzema, M., Aulbach, L., and Wagner, E. G. H. (2017). Two regulatory RNA elements affect TisB-dependent depolarization and persister formation. *Mol. Microbiol.* 103, 1020–1033. doi: 10.1111/mmi.13607

Berghoff, B. A., Konzer, A., Mank, N. N., Looso, M., Rische, T., Förstner, K. U., et al. (2013). Integrative "omics"-approach discovers dynamic and regulatory features of bacterial stress responses. *PLoS Genet.* 9:e1003576. doi: 10.1371/journal.pgen.1003576

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

Bigger, J. W. (1944). The bactericidal action of penicillin on *Staphylococcus pyogenes*. *Ir. J. Med. Sci.* 19, 553–568. doi: 10.1007/BF02948386

Black, D. S., Irwin, B., and Moyed, H. S. (1994). Autoregulation of *hip*, an operon that affects lethality due to inhibition of peptidoglycan or DNA synthesis. *J. Bacteriol.* 176, 4081–4091. doi: 10.1128/jb.176.13.4081-4091.1994

Black, D. S., Kelly, A. J., Mardis, M. J., and Moyed, H. S. (1991). Structure and organization of *hip*, an operon that affects lethality due to inhibition of peptidoglycan or DNA synthesis. *J. Bacteriol.* 173, 5732–5739. doi: 10.1128/jb.173.18.5732-5739.1991

Brandi, A., Giangrossi, M., Giuliodori, A. M., and Falconi, M. (2016). An interplay among FIS, HNS, and guanosine tetraphosphate modulates transcription of the *Escherichia coli cspA* gene under physiological growth conditions. *Front. Mol. Biosci.* 3:19. doi: 10.3389/fmolb.2016.00019

Brantl, S., and Jahn, N. (2015). sRNAs in bacterial type I and type III toxin-antitoxin systems. *FEMS Microbiol. Rev.* 39, 413–427. doi: 10.1093/femsre/fuv003

Brauner, A., Fridman, O., Gefen, O., and Balaban, N. Q. (2016). Distinguishing between resistance, tolerance and persistence to antibiotic treatment. *Nat. Rev. Microbiol.* 14, 320–330. doi: 10.1038/nrmicro.2016.34

Buerger, S., Spoering, A., Gavrish, E., Leslin, C., Ling, L., and Epstein, S. S. (2012). Microbial scout hypothesis, stochastic exit from dormancy, and the nature of slow growers. *Appl. Environ. Microbiol.* 78, 3221–3228. doi: 10.1128/AEM.07307-11

Cañas-Duarte, S. J., Restrepo, S., and Pedraza, J. M. (2014). Novel protocol for persister cells isolation. *PLoS ONE* 9:e88660. doi: 10.1371/journal.pone.0088660

Cheverton, A. M., Gollan, B., Przydacz, M., Wong, C. T., Mylona, A., Hare, S. A., et al. (2016). A *Salmonella* toxin promotes persister formation through acetylation of tRNA. *Mol. Cell* 63, 86–96. doi: 10.1016/j.molcel.2016.05.002

Chua, S. L., Yam, J. K. H., Hao, P., Adav, S. S., Salido, M. M., Liu, Y., et al. (2016). Selective labelling and eradication of antibiotic-tolerant bacterial populations in *Pseudomonas aeruginosa* biofilms. *Nat. Commun.* 7:10750. doi: 10.1038/ncomms10750

Conlon, B. P., Rowe, S. E., Gandt, A. B., Nuxoll, A. S., Donegan, N. P., Zalis, E. A., et al. (2016). Persister formation in *Staphylococcus aureus* is associated with ATP depletion. *Nat. Microbiol.* 1:16051. doi: 10.1038/nmicrobiol.2016.51

Cox, J., Hein, M. Y., Luber, C. A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteomics* 13, 2513–2526. doi: 10.1074/mcp.M113.031591

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26, 1367–1372. doi: 10.1038/nbt.1511

Darfeuille, F., Unoson, C., Vogel, J., and Wagner, E. G. H. (2007). An antisense RNA inhibits translation by competing with standby ribosomes. *Mol. Cell* 26, 381–392. doi: 10.1016/j.molcel.2007.04.003

Datta, S., Costantino, N., and Court, D. L. (2006). A set of recombineering plasmids for gram-negative bacteria. *Gene* 379, 109–115. doi: 10.1016/j.gene.2006.04.018

De Jonge, N., Garcia-Pino, A., Buts, L., Haesaerts, S., Charlier, D., Zangger, K., et al. (2009). Rejuvenation of CcdB-poisoned gyrase by an intrinsically disordered protein domain. *Mol. Cell* 35, 154–163. doi: 10.1016/j.molcel.2009.05.025

Desnues, B., Cuny, C., Grégori, G., Dukan, S., Aguilaniu, H., and Nyström, T. (2003). Differential oxidative damage and expression of stress defence regulons in culturable and non-culturable *Escherichia coli* cells. *EMBO Rep.* 4, 400–404. doi: 10.1038/sj.embor.embor799

Dörr, T., Vulic, M., and Lewis, K. (2010). Ciprofloxacin causes persister formation by inducing the TisB toxin in *Escherichia coli*. *PLoS Biol.* 8:e1000317. doi: 10.1371/journal.pbio.1000317

Dwyer, D. J., Belenky, P. A., Yang, J. H., MacDonald, I. C., Martell, J. D., Takahashi, N., et al. (2014). Antibiotics induce redox-related physiological alterations as part of their lethality. *Proc. Natl. Acad. Sci.U.S.A.* 111, E2100–E2109. doi: 10.1073/pnas.1401876111

Fozo, E. M., Hemm, M. R., and Storz, G. (2008a). Small toxic proteins and the antisense RNAs that repress them. *Microbiol. Mol. Biol. Rev.* 72, 579–589. doi: 10.1128/MMBR.00025-08

Fozo, E. M., Kawano, M., Fontaine, F., Kaya, Y., Mendieta, K. S., Jones, K. L., et al. (2008b). Repression of small toxic protein synthesis by the Sib and OhsC small RNAs. *Mol. Microbiol.* 70, 1076–1093. doi: 10.1111/j.1365-2958.2008.06394.x

Fröhlich, F., Christiano, R., and Walther, T. C. (2013). Native SILAC: metabolic labeling of proteins in prototroph microorganisms based on lysine synthesis regulation. *Mol. Cell. Proteomics* 12, 1995–2005. doi: 10.1074/mcp.M112.025742

Gerdes, K., Christensen, S. K., and Løbner-Olesen, A. (2005). Prokaryotic toxin–antitoxin stress response loci. *Nat. Rev. Microbiol.* 3, 371–382. doi: 10.1038/nrmicro1147

Gerdes, K., and Maisonneuve, E. (2012). Bacterial persistence and toxin-antitoxin loci. *Annu. Rev. Microbiol.* 66, 103–123. doi: 10.1146/annurev-micro-092611-150159

Germain, E., Castro-Roa, D., Zenkin, N., and Gerdes, K. (2013). Molecular mechanism of bacterial persistence by HipA. *Mol. Cell* 52, 248–254. doi: 10.1016/j.molcel.2013.08.045

Gurnev, P. A., Ortenberg, R., Dörr, T., Lewis, K., and Bezrukov, S. M. (2012). Persister-promoting bacterial toxin TisB produces anion-selective pores in planar lipid bilayers. *FEBS Lett.* 586, 2529–2534. doi: 10.1016/j.febslet.2012.06.021

Harrison, J. J., Wade, W. D., Akierman, S., Vacchi-Suzzi, C., Stremick, C. A., Turner, R. J., et al. (2009). The chromosomal toxin gene *yafQ* is a determinant of multidrug tolerance for *Escherichia coli* growing in a biofilm. *Antimicrob. Agents Chemother.* 53, 2253–2258. doi: 10.1128/AAC.00043-09

Hayes, F. (2003). Toxins-antitoxins: plasmid maintenance, programmed cell death, and cell cycle arrest. *Science* 301, 1496–1499. doi: 10.1126/science.1088157

Henry, T. C., and Brynildsen, M. P. (2016). Development of persister-FACSeq: a method to massively parallelize quantification of persister physiology and its heterogeneity. *Sci. Rep.* 6:25100. doi: 10.1038/srep25100

Huang, D. W., Lempicki, R. A., and Sherman, B. T. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57. doi: 10.1038/nprot.2008.211

Imlay, J. A. (2006). Iron-sulphur clusters and the problem with oxygen. *Mol. Microbiol.* 59, 1073–1082. doi: 10.1111/j.1365-2958.2006.05028.x

Imlay, J. A. (2008). Cellular defenses against superoxide and hydrogen peroxide. *Annu. Rev. Biochem.* 77, 755–776. doi: 10.1146/annurev.biochem.77.061606.161055

Imlay, J. A. (2013). The molecular mechanisms and physiological consequences of oxidative stress: lessons from a model bacterium. *Nat. Rev. Microbiol.* 11, 443–454. doi: 10.1038/nrmicro3032

Jacobson, F. S., Morgan, R. W., Christman, M. F., and Ames, B. N. (1989). An alkyl hydroperoxide reductase from *Salmonella typhimurium* involved in the defense of DNA against oxidative damage. *J. Biol. Chem.* 264, 1488–1496.

Jõers, A., Kaldalu, N., and Tenson, T. (2010). The frequency of persisters in *Escherichia coli* reflects the kinetics of awakening from dormancy. *J. Bacteriol.* 192, 3379–3384. doi: 10.1128/JB.00056-10

Kaspy, I., Rotem, E., Weiss, N., Ronin, I., Balaban, N. Q., and Glaser, G. (2013). HipA-mediated antibiotic persistence via phosphorylation of the glutamyl-tRNA-synthetase. *Nat. Commun.* 4:3001. doi: 10.1038/ncomms4001

Keren, I., Minami, S., Rubin, E., and Lewis, K. (2011). Characterization and transcriptome analysis of *Mycobacterium tuberculosis* persisters. *MBio* 2, e00100–e00111. doi: 10.1128/mBio.00100-11

Keren, I., Shah, D., Spoering, A., Kaldalu, N., and Lewis, K. (2004). Specialized persister cells and the mechanism of multidrug tolerance in *Escherichia coli*. *J. Bacteriol.* 186, 8172–8180. doi: 10.1128/JB.186.24.8172-8180.2004

Kim, J.-S., Chowdhury, N., Yamasaki, R., and Wood, T. K. (2018). Viable but non-culturable and persistence describe the same bacterial stress state. *Environ. Microbiol.* 20, 2038–2048. doi: 10.1111/1462-2920.14075

Kim, Y., and Wood, T. K. (2010). Toxins Hha and CspD and small RNA regulator Hfq are involved in persister cell formation through MqsR in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* 391, 209–213. doi: 10.1016/j.bbrc.2009.11.033

Kohanski, M. A., Dwyer, D. J., Hayete, B., Lawrence, C. A., and Collins, J. J. (2007). A common mechanism of cellular death induced by bactericidal antibiotics. *Cell* 130, 797–810. doi: 10.1016/j.cell.2007.06.049

Korch, S. B., Henderson, T. A., and Hill, T. M. (2003). Characterization of the *hipA7* allele of *Escherichia coli* and evidence that high persistence is governed by (p)ppGpp synthesis. *Mol. Microbiol.* 50, 1199–1213. doi: 10.1046/j.1365-2958.2003.03779.x

Langklotz, S., and Narberhaus, F. (2011). The *Escherichia coli* replication inhibitor CspD is subject to growth-regulated degradation by the Lon protease. *Mol. Microbiol.* 80, 1313–1325. doi: 10.1111/j.1365-2958.2011.07646.x

Levin-Reisman, I., Fridman, O., and Balaban, N. Q. (2014). ScanLag: high-throughput quantification of colony growth and lag time. *J. Vis. Exp.* e51456. doi: 10.3791/51456

Levin-Reisman, I., Gefen, O., Fridman, O., Ronin, I., Shwa, D., Sheftel, H., et al. (2010). Automated imaging with ScanLag reveals previously undetectable bacterial growth phenotypes. *Nat. Methods* 7, 737–739. doi: 10.1038/nmeth.1485

Lewis, K. (2007). Persister cells, dormancy and infectious disease. *Nat. Rev. Microbiol.* 5, 48–56. doi: 10.1038/nrmicro1557

Lewis, K. (2010). Persister cells. *Annu. Rev. Microbiol.* 64, 357–372. doi: 10.1146/annurev.micro.112408.134306

Lewis, K. (2013). Platforms for antibiotic discovery. *Nat. Rev. Drug Discov.* 12, 371–387. doi: 10.1038/nrd3975

Lu, J., Yang, J., Tan, G., and Ding, H. (2008). Complementary roles of SufA and IscA in the biogenesis of iron–sulfur clusters in *Escherichia coli*. *Biochem. J.* 409, 535–543. doi: 10.1042/BJ20071166

Manganelli, R., and Gennaro, M. L. (2017). Protecting from envelope stress: variations on the phage-shock-protein theme. *Trends Microbiol.* 25, 205–216. doi: 10.1016/j.tim.2016.10.001

Michalik, S., Bernhardt, J., Otto, A., Moche, M., Becher, D., Meyer, H., et al. (2012). Life and death of proteins: a case study of glucose-starved *Staphylococcus aureus*. *Mol. Cell. Proteomics* 11, 558–570. doi: 10.1074/mcp.M112.017004

Michiels, J. E., Van den Bergh, B., Verstraeten, N., and Michiels, J. (2016). Molecular mechanisms and clinical implications of bacterial persistence. *Drug Resist. Updat.* 29, 76–89. doi: 10.1016/j.drup.2016.10.002

Mok, W. W. K., and Brynildsen, M. P. (2018). Timing of DNA damage responses impacts persistence to fluoroquinolones. *Proc. Natl. Acad. Sci. U.S.A.* 115, E6301–E6309. doi: 10.1073/pnas.1804218115

Moyed, H. S., and Bertrand, K. P. (1983). *hipA*, a newly recognized gene of *Escherichia coli* K-12 that affects frequency of persistence after inhibition of murein synthesis. *J. Bacteriol.* 155, 768–775.

Nikaido, H. (2003). Molecular basis of bacterial outer membrane permeability revisited. *Microbiol. Mol. Biol. Rev.* 67, 593–656. doi: 10.1128/MMBR.67.4.593-656.2003

Orman, M. A., and Brynildsen, M. P. (2013). Dormancy is not necessary or sufficient for bacterial persistence. *Antimicrob. Agents Chemother.* 57, 3230–3239. doi: 10.1128/AAC.00243-13

Page, R., and Peti, W. (2016). Toxin-antitoxin systems in bacterial growth arrest and persistence. *Nat. Chem. Biol.* 12, 208–214. doi: 10.1038/nchembio.2044

Pagès, J.-M., James, C. E., and Winterhalter, M. (2008). The porin and the permeating antibiotic: a selective diffusion barrier in Gram-negative bacteria. *Nat. Rev. Microbiol.* 6, 893–903. doi: 10.1038/nrmicro1994

Pu, Y., Zhao, Z., Li, Y., Zou, J., Ma, Q., Zhao, Y., et al. (2016). Enhanced efflux activity facilitates drug tolerance in dormant bacterial cells. *Mol. Cell* 62, 284–294. doi: 10.1016/j.molcel.2016.03.035

Radzikowski, J. L., Vedelaar, S., Siegel, D., Ortega, Á. D., Schmidt, A., and Heinemann, M. (2016). Bacterial persistence is an active σS stress response to metabolic flux limitation. *Mol. Syst. Biol.* 12:882. doi: 10.15252/msb.20166998

Rappsilber, J., Ishihama, Y., and Mann, M. (2003). Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal. Chem.* 75, 663–670. doi: 10.1021/ac026117i

Roche, B., Aussel, L., Ezraty, B., Mandin, P., Py, B., and Barras, F. (2013). Iron/sulfur proteins biogenesis in prokaryotes: formation, regulation and diversity. *Biochim. Biophys. Acta Bioenerg.* 1827, 455–469. doi: 10.1016/j.bbabio.2012.12.010

Schwanhäusser, B., Gossen, M., Dittmar, G., and Selbach, M. (2009). Global analysis of cellular protein translation by pulsed SILAC. *Proteomics* 9, 205–209. doi: 10.1002/pmic.200800275

Sevcenco, A. M., Pinkse, M. W. H., Wolterbeek, H. T., Verhaert, P. D. E. M., Hagen, W. R., and Hagedoorn, P. L. (2011). Exploring the microbial metalloproteome using MIRAGE. *Metallomics* 3, 1324–1330. doi: 10.1039/c1mt00154j

Shah, D., Zhang, Z., Khodursky, A., Kaldalu, N., Kurg, K., and Lewis, K. (2006). Persisters: a distinct physiological state of *E. coli*. *BMC Microbiol.* 6:53. doi: 10.1186/1471-2180-6-53

Shan, Y., Brown Gandt, A., Rowe, S. E., Deisinger, J. P., Conlon, B. P., and Lewis, K. (2017). ATP-dependent persister formation in *Escherichia coli*. *MBio* 8, e02267–e02216. doi: 10.1128/mBio.02267-16

Storz, G., Jacobson, F. S., Tartaglia, L. A., Morgan, R. W., Silveira, L. A., and Ames, B. N. (1989). An alkyl hydroperoxide reductase induced by oxidative stress in *Salmonella typhimurium* and *Escherichia coli*: genetic characterization and cloning of *ahp*. *J. Bacteriol.* 171, 2049–2055. doi: 10.1128/jb.171.4.2049-2055.1989

Sturm, A., and Dworkin, J. (2015). Phenotypic diversity as a mechanism to exit cellular dormancy. *Curr. Biol.* 25, 2272–2277. doi: 10.1016/j.cub.2015.07.018

Tokumoto, U., and Takahashi, Y. (2001). Genetic analysis of the *isc* operon in *Escherichia coli* involved in the biogenesis of cellular iron-sulfur proteins. *J. Biochem.* 130, 63–71. doi: 10.1093/oxfordjournals.jbchem.a002963

Tripathi, A., Dewan, P. C., Siddique, S. A., and Varadarajan, R. (2014). MazF-induced growth inhibition and persister generation in *Escherichia coli*. *J. Biol. Chem.* 289, 4191–4205. doi: 10.1074/jbc.M113.510511

Unoson, C., and Wagner, E. G. H. (2008). A small SOS-induced toxin is targeted against the inner membrane in *Escherichia coli*. *Mol. Microbiol.* 70, 258–270. doi: 10.1111/j.1365-2958.2008.06416.x

Verstraeten, N., Knapen, W. J., Kint, C. I., Liebens, V., Van den Bergh, B., Dewachter, L., et al. (2015). Obg and membrane depolarization are part of a microbial bet-hedging strategy that leads to antibiotic tolerance. *Mol. Cell* 59, 9–21. doi: 10.1016/j.molcel.2015.05.011

Vila-Sanjurjo, A., Schuwirth, B. S., Hau, C. W., and Cate, J. H. D. (2004). Structural basis for the control of translation initiation during stress. *Nat. Struct. Mol. Biol.* 11, 1054–1059. doi: 10.1038/nsmb850

Vogel, J., Argaman, L., Wagner, E. G. H., and Altuvia, S. (2004). The small RNA istR inhibits synthesis of an SOS-induced toxic peptide. *Curr. Biol.* 14, 2271–2276. doi: 10.1016/j.cub.2004.12.003

Völzing, K. G., and Brynildsen, M. P. (2015). Stationary-phase persisters to ofloxacin sustain DNA damage and require repair systems only during recovery. *MBio* 6, e00731–e00715. doi: 10.1128/mBio.00731-15

Wang, Y., Bojer, M. S., George, S. E., Wang, Z., Jensen, P. R., Wolz, C., et al. (2018). Inactivation of TCA cycle enhances *Staphylococcus aureus* persister cell formation in stationary phase. *Sci. Rep.* 8:10849. doi: 10.1038/s41598-018-29123-0

Weel-Sneve, R., Kristiansen, K. I., Odsbu, I., Dalhus, B., Booth, J., Rognes, T., et al. (2013). Single transmembrane peptide dinq modulates membrane-dependent activities. *PLoS Genet.* 9:e1003260. doi: 10.1371/journal.pgen.1003260

Wilmaerts, D., Bayoumi, M., Dewachter, L., Knapen, W., Mika, J. T., Hofkens, J., et al. (2018). The persistence-inducing toxin HokB forms dynamic pores that cause ATP leakage. *MBio* 9, e00744–e00718. doi: 10.1128/mBio.00744-18

Worzfeld, T., Finkernagel, F., Reinartz, S., Konzer, A., Adhikary, T., Nist, A., et al. (2018). Proteotranscriptomics reveal signaling networks in the ovarian cancer microenvironment. *Mol. Cell. Proteomics* 17, 270–289. doi: 10.1074/mcp.RA117.000400

Yu, D., Ellis, H. M., Lee, E. C., Jenkins, N. A., Copeland, N. G., and Court, D. L. (2000). An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 97, 5978–5983. doi: 10.1073/pnas.100127597

Zeth, K. (2012). Dps biomineralizing proteins: multifunctional architects of nature. *Biochem. J.* 445, 297–311. doi: 10.1042/BJ20120514

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

Appendix

1 Declaration

I declare that I have completed this dissertation single-handedly without the unauthorized help of a second party and only with the assistance acknowledged therein. I have appropriately acknowledged and cited all text passages that are derived verbatim from or are based on the content of published work of others, and all information relating to verbal communications. I consent to the use of an anti-plagiarism software to check my thesis. I have abided by the principles of good scientific conduct laid down in the charter of the Justus Liebig University Giessen "Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis" in carrying out the investigations described in the dissertation.

Place, Date : *Signature* :

2 Authors' contributions

Chapter 2 Daniel-Timon Spanka performed the RNA experiments, developed and programmed XPEAP, contributed to the data analysis and interpretation, and wrote the manuscript. Carina Maria Reuscher performed the physiological experiments and prepared total RNA for RNA sequencing. Gabriele Klug raised the funds, designed and supervised this study, was involved in data interpretation and manuscript writing.

Chapter 3 Conceptualization, Daniel-Timon Spanka and Gabriele Klug; methodology, Daniel-Timon Spanka and Gabriele Klug; software, Daniel-Timon Spanka; validation, Daniel-Timon Spanka and Gabriele Klug; formal analysis, Daniel-Timon Spanka; investigation, Daniel-Timon Spanka; resources, Gabriele Klug; data curation, Daniel-Timon Spanka; writing — original draft preparation, Daniel-Timon Spanka and Gabriele Klug; writing — review and editing, Gabriele Klug; visualization, Daniel-Timon Spanka; supervision, Gabriele Klug; project administration, Gabriele Klug; funding acquisition, Gabriele Klug.

Chapter 4 Julian Grützner: data curation, formal analysis, investigation, validation, methodology, editing draft. Fabian Billenkamp: data curation, formal analysis, investigation, methodology. Daniel-Timon Spanka: data curation, formal analysis, validation, methodology. Tim Rick: data curation; methodology. Vivian Monzon: formal analysis, validation. Konrad U. Förstner: formal analysis, validation. Gabriele Klug: conceptualization, data curation, formal analysis, supervision, funding acquisition, project administration, writing original draft, review and editing.

Chapter 5 Conceptualization, Bernhard Remes, Katrin M. H. Eisenhardt, Julian Grützner, Gabriele Klug; Validation, Bernhard Remes, Katrin M. H. Eisenhardt, Julian Grützner, Daniel-Timon Spanka, Andreas Jäger, Gabriele Klug; investigation: Bernhard Remes, Katrin M. H. Eisenhardt, Julian Grützner, Daniel-Timon Spanka, Andreas Jäger; data curation, Bernhard Remes, Katrin M. H. Eisenhardt, Julian Grützner, Andreas Jäger, Daniel-Timon Spanka; formal analysis, Daniel-Timon Spanka, writing, Gabriele Klug; writing — review and editing, Julian Grützner, Gabriele Klug; visualization, Bernhard Remes, Katrin M. H. Eisenhardt, Julian Grützner, Andreas Jäger, Gabriele Klug; supervision, Bernhard Remes, Katrin M. H. Eisenhardt, Julian Grützner, Gabriele Klug; project administration, Gabriele Klug; funding acquisition, Gabriele Klug.

Chapter 6 Anne Konzer and Bork Berghoff designed the study. Daniel-Timon Spanka performed most of the physiological experiments and established a workflow for analysis of ScanLag data. Daniel Edelmann and Bork Berghoff performed additional physiological experiments and analyzed the data. Anne Konzer performed mass spectrometry, analyzed the data, and contributed to the material and methods section. Bork Berghoff wrote the manuscript.

3 Acknowledgments

Writing a doctoral thesis is much more than a daily business and I am very grateful for so many people who were involved in my projects during the last years.

First I want to thank Prof. Dr. Gabriele Klug for supervising my work, for the always good cooperation on a professional and personal level, and not least for the reliable support. Freedom to act and decide characterized my time at the institute, which I am deeply grateful for. Thank you very much!

Furthermore I am grateful that Prof. Dr. Annegret Wilde agreed to provide a second review of this work, and that Prof. Dr. Reinhard Dammann and Prof. Dr. Kai Thormann have agreed to contribute to the disputation as examiners.

Thank you PD Dr. B. B. for so many deep discussions – about scientific and non-scientific topics, at day- and night-time –, reading the first draft of my thesis, and of course for generously sharing office and coffee with me. Keep on riding (your bicycle).

Moreover I thank Dr. M. M. for giving me advice and hints during the planning phases of so many plasmid cloning strategies, providing several different plasmids for my own research, and giving advice regarding English orthography.

The bioinformatics experts P. Bl., T. A., and P. Ba. from the AG Goesmann helped me a lot, since major parts of my projects involved programming, NGS analysis, and performing data magic of all kind. Thank you guys!

Many experiments would not have been realized without the help of our lab technicians: Special thanks to A. W. for cloning the army of pPHU231-mVenus promoter fusions, K. H. for cloning chromosomal deletions, and A. J. and S. S. for performing all types of experiments I gave them. Furthermore I am grateful for the help provided by the students H. K. and L. Z.

Good projects become even better in the right company, this is why I am grateful for the excellent time spent with M. S., R. S., and our alumni M. L. and C. R., as well as the whole MiMo-Crew.

My parents G. and C. as well as S., T. and G. backed me up a lot and helped me to reach this point. Your support was awesome and without you I would most likely not have written these lines, so thank you!

J., words can hardly express the support I received from you over the years. Together we are sailing the ocean called ‘life’, sometimes anchoring in calm bays, sometimes fighting rough waves and spindrift, and often navigating in unknown waters. No matter what life has still in store for us, together we’ll stay on course. Thank you.