

RESEARCH ARTICLE

The Conserved *Dcw* Gene Cluster of *R. sphaeroides* Is Preceded by an Uncommonly Extended 5' Leader Featuring the sRNA UpsM

Lennart Weber¹, Clemens Thoenen², Marcel Volk¹, Bernhard Remes¹, Marcus Lechner², Gabriele Klug^{1*}

1 Institute of Microbiology and Molecular Biology, IFZ, Justus-Liebig-University Giessen, Giessen, Germany, **2** Institute of Pharmaceutical Chemistry, Philipps-University Marburg, Marburg, Germany

* Gabriele.Klug@mikro.bio.uni-giessen.de



OPEN ACCESS

Citation: Weber L, Thoenen C, Volk M, Remes B, Lechner M, 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(11): e0165694. doi:10.1371/journal.pone.0165694

Editor: Roy Martin Roop, II, East Carolina University Brody School of Medicine, UNITED STATES

Received: June 16, 2016

Accepted: October 17, 2016

Published: November 1, 2016

Copyright: © 2016 Weber et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data and accession number(s) are within the paper and its Supporting Information files. Raw files of transcriptome data are accessible via Gene Expression Omnibus (accession number GSE71844) and NCBI BioProject (accession number PRJNA343088).

Funding: This work was supported by the Deutsche Forschungsgemeinschaft (<http://www.dfg.de/>) (DFG KI563/20) and the DFG funded International Research Training Group (GRK1384)

Abstract

Cell division and cell wall synthesis mechanisms are similarly conserved among bacteria. Consequently some bacterial species have comparable sets of genes organized in the *dcw* (*division and cell wall*) gene cluster. *Dcw* genes, their regulation and their relative order within the cluster are outstandingly conserved among rod shaped and gram negative bacteria to ensure an efficient coordination of growth and division. A well studied representative is the *dcw* gene cluster of *E. coli*. The first promoter of the gene cluster (*mraZ1p*) gives rise to polycistronic transcripts containing a 38 nt long 5' UTR followed by the first gene *mraZ*. Despite reported conservation we present evidence for a much longer 5' UTR in the gram negative and rod shaped bacterium *Rhodobacter sphaeroides* and in the family of *Rhodobacteraceae*. This extended 268 nt long 5' UTR comprises a Rho independent terminator, which in case of termination gives rise to a non-coding RNA (UpsM). This sRNA is conditionally cleaved by RNase E under stress conditions in an Hfq- and very likely target mRNA-dependent manner, implying its function in *trans*. These results raise the question for the regulatory function of this extended 5' UTR. It might represent the rarely described case of a *trans* acting sRNA derived from a riboswitch with exclusive presence in the family of *Rhodobacteraceae*.

Introduction

There are only rare cases for highly conserved gene clusters throughout bacterial genomes due to evolutionary dynamics. Examples for such clusters are genes for ribosomal proteins, the *atp* operon or the *dcw* (*division and cell wall*) gene cluster [1]. However, conservation of *dcw* genes, their regulation and especially their arrangement within the cluster are outstandingly conserved within bacterial groups of similar taxon and cell shape [2]. Besides regulatory mechanisms the conserved order of the genes may ensure an efficient coordination of growth and

(LW). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

division as assumed by the *genomic channeling hypothesis* [3]. A well described example for such conservation is the *dcw* gene cluster of gram negative and rod shaped bacteria.

Dcw gene regulation was studied intensively in *E. coli*, but is not fully understood due to numerous regulatory features like internal promoters, transcript stabilities and protein ratios. It is well known that the first promoter (*mraZ1p*) of the gene cluster in *E. coli* (16 genes in total) gives rise to polycistronic transcripts containing a 38nt long 5' UTR followed by the first gene *mraZ* [4]. Downstream of *mraZ* transcription can potentially continue up to the last gene of the locus (*envA*) that harbors a Rho independent terminator [5].

Here we present evidence for a much longer 5' UTR in the gram negative and rod shaped bacterium *Rhodobacter sphaeroides* also present in other members of *Rhodobacteraceae*. In *R. sphaeroides* this 268 nt long 5' UTR features a Rho independent terminator 84 nt upstream of *mraZ*, which in case of transcriptional termination gives rise to a non-coding RNA of 206 nt length. This transcript was described as an orphan sRNA named RSs0682 [6], henceforth renamed UpsM (upstream sRNA of *mraZ*). Here we also show that conditional processing of UpsM requires the RNA chaperon Hfq, the endoribonuclease RNase E and induction of the RpoHI/II regulon.

Our results raise the question for the complex regulatory function of this extended 5' UTR, which might represent a rare dual function riboswitch exclusively present in the family of *Rhodobacteraceae*.

Results

dRNA-seq Hints at an Extended 5' UTR of the *Dcw* Gene Cluster of *R. sphaeroides*

The sRNA UpsM (previously RSs0682) was described as the most abundant orphan sRNA of *R. sphaeroides*, which is encoded in the intergenic region (IGR) upstream of *mraZ*, the first gene of the *dcw* (division and cell wall) gene cluster. A Rho-independent terminator was predicted at the 3' end of the sRNA locus. Identification of UpsM was based on deep sequencing of cDNA libraries using 454 pyrosequencing. The coverage of UpsM was comparably low (~2,000 reads per library) in the initial sequencing study [6]. Since sequencing technologies have rapidly evolved and nowadays generate millions of reads concomitant with visualisation of transcripts with low abundance, we re-analysed the UpsM locus in a dRNA-seq dataset that has been generated using Illumina sequencing technology (Fig 1 and S1 Fig) without and with prior TEX (terminator-5' phosphate dependent exonuclease) treatment of the RNA to enrich primary transcripts. This confirmed the presence and high abundance of UpsM and the transcriptional start site (TSS). In contrast to the low-coverage 454 pyrosequencing study, additional observations were possible. First, the processing site within UpsM, which was already detected by Northern blot analysis and 5' RACE [6], becomes apparent by a sudden decrease of reads especially after TEX treatment. Secondly, the downstream gene *mraZ* is not preceded by a separate TSS. This is surprising since *mraZ* is the first gene of the *dcw* gene cluster. Therefore it should be expressed in exponentially growing cells in the course of cell wall synthesis and cell division as described for other bacteria [4, 7–11]. This observation led to the following assumptions: *MraZ* transcription depends on the UpsM promoter and there is no additional promoter/TSS exclusively present for *mraZ*. If this is true, the terminator of UpsM has to allow read-throughs in order to guarantee transcription of *mraZ*. This has several consequences: 1.) *mraZ* or the *dcw* gene cluster have an uncommonly long 5' UTR of 268 nt in length, which has not been reported for other bacterial *dcw* clusters, which show high conservation among rod shaped and gram negative bacteria. 2.) UpsM is not an sRNA derived from an IGR (intergenic region), but rather an sRNA which is generated by transcription termination within the 5'

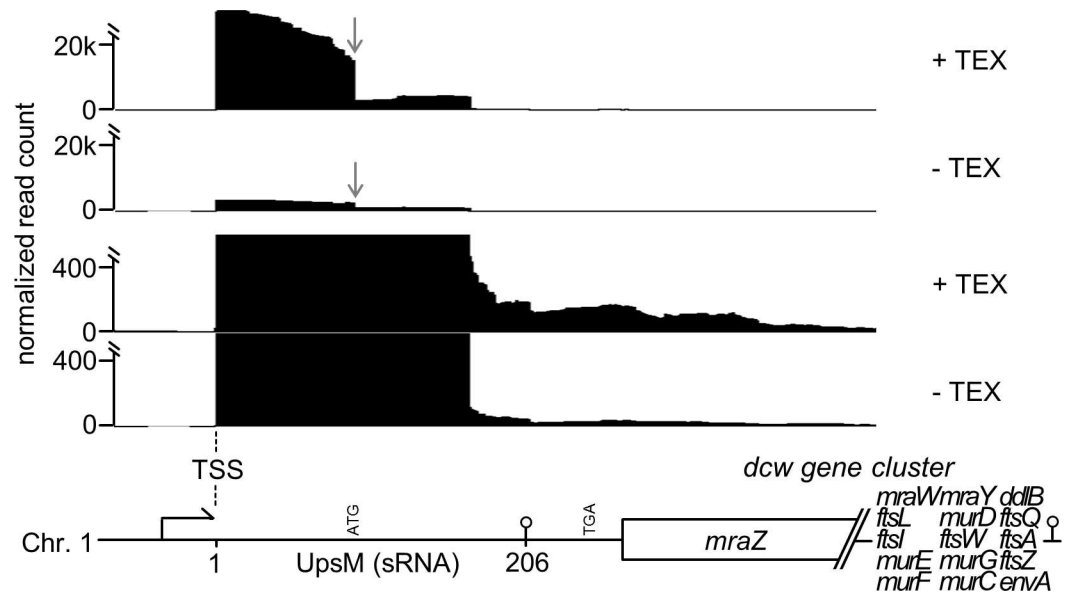


Fig 1. dRNA-seq shows a long 5' UTR of the *mraZ* gene in *R. sphaeroides*. Modified screenshots taken from IGB (integrated genome browser) visualizing the coverage at the genetic locus of *mraZ*. Shown are normalized cDNA reads on a large scale (upper two panels) and a smaller scale (lower two panels) obtained from TEX treated and untreated total RNA isolated from an exponentially and microaerobically grown *R. sphaeroides* 2.4.1 culture. The genetic context is displayed at the bottom. *mraZ* is the first gene of the *dcw* gene cluster. Position 1 reflects the TSS of sRNA UpsM (206 nt) 268 nucleotides upstream of *mraZ*. The terminator of UpsM is indicated as hairpin structure and a processing site within the sRNA is highlighted by an arrow.

doi:10.1371/journal.pone.0165694.g001

UTR of *mraZ*. 3.) In case the terminator allows read-through, the 5' UTR contains a start codon upstream and a stop codon in frame downstream of the terminator and therefore might encode a leader peptide here designated as sORF (small open reading frame).

UpsM Is a 5' UTR Derived sRNA, Which Is Conditionally Cleaved during Stress Conditions in an RNase E, Hfq and RpoHI/RpoHII-Dependent Manner

UpsM was originally identified as an sRNA processed upon ¹O₂ stress in *Rhodobacter sphaeroides* and the RNA chaperone Hfq was shown to be required for this processing [6]. UpsM is the most abundant sRNA in *R. sphaeroides* and represents about 60% of all Hfq bound sRNAs [12]. In addition to the strong interaction of UpsM with Hfq we observed a negative growth effect under anaerobic and aerobic conditions (S2B Fig) and changes in the transcriptome in an initial microarray analysis (S1 Table) for a strain overexpressing UpsM (S2A Fig). Therefore we conclude that UpsM is functional as a *trans* acting sRNA.

To further prove that the stable 130 nt UpsM 3' fragment is generated by processing, total RNA from *R. sphaeroides* isolated 90 min after addition of methylene blue in the light to generate ¹O₂ was treated with TEX (terminator-5' phosphate dependent exonuclease), which degrades RNAs with a monophosphate at the 5' end but not primary transcripts that carry a 5' triphosphate. The TEX treated RNA was compared to untreated RNA on a Northern blot (Fig 2A). The lack of the 130 nt band after TEX treatment strongly supports the assumption that this fragment is a processing product.

In gram-negative bacteria the endoribonuclease RNase E has a major role in initiation of mRNA decay [13, 14] and was also shown to be involved in the generation or processing of several sRNAs [15–18]. To test the involvement of RNase E in UpsM processing we constructed

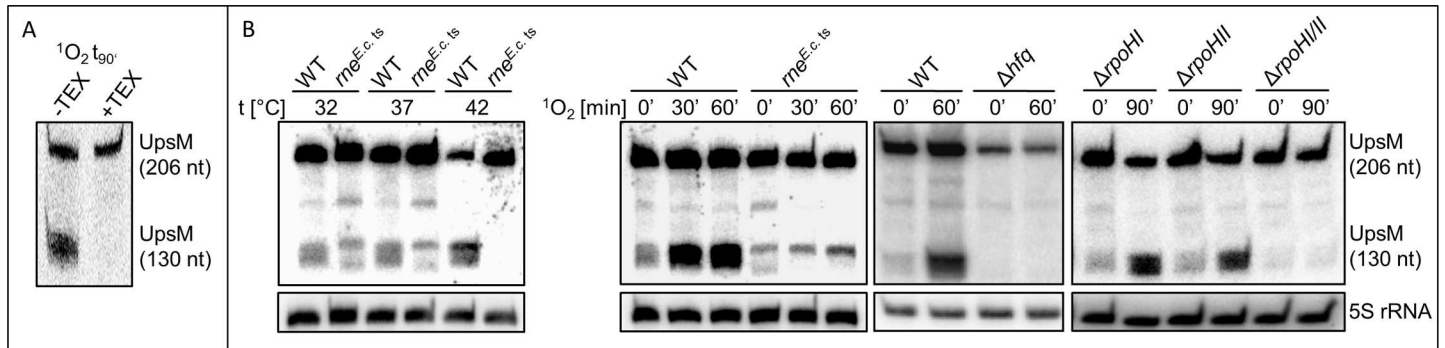


Fig 2. Northern blot analysis reveals Hfq and target mRNA dependent processing of UpsM by RNase E. (A) Detection of UpsM (206 nt) and the UpsM processing product (130 nt) by Northern blot analysis of TEX treated and untreated total RNA isolated from *R. sphaeroides* after 90 min $^1\text{O}_2$ stress. (B) Left: Comparison of the UpsM processing pattern via Northern blot analysis of total RNA isolated from *R. sphaeroides* 2.4.1 (WT) to RNA isolated from a mutant strain expressing a thermosensitive RNase E variant from *E. coli* ($rne^{E.c.ts}$) after growth at 32, 37 or 42°C for 30 min or during $^1\text{O}_2$ stress. Right: Comparison of the processing pattern via Northern blot analysis to strains lacking Hfq, RpoHI, RpoHII or RpoHI and RpoHII after 0 and 60 or 0 and 90 min of $^1\text{O}_2$ stress. Signals of 5S rRNA serve as loading control.

doi:10.1371/journal.pone.0165694.g002

an *R. sphaeroides* strain with impaired RNase E activity. The endogenous *rne* gene (RSP_2131) was replaced by the *rne* gene from *E. coli* N3431 (46% blastp identity) [19, 20], which produces a temperature-sensitive RNase E due to a point mutation. As seen in Fig 2B processing of UpsM in strain *R. sphaeroides rne^{E.c.ts}* is already impaired at 32°C or 37°C, indicating that the *E. coli* enzyme is less active than the endogenous enzyme of *R. sphaeroides*. When cells are shifted to 42°C no UpsM processing occurs in the *R. sphaeroides* strain expressing the temperature-sensitive RNase E variant. When methylene blue was added to the cultures in the light growing at 32°C accumulation of the 130 nt UpsM fragment was only weak in the *rne* mutant strain, while a strong accumulation was observed in the wild type. As shown previously a lack of Hfq also abolished UpsM processing (Fig 2B) [6].

Deletion of the endonuclease RNase III or the 3' to 5' exonuclease RNase J does not lead to an altered processing (S3A Fig), showing that those nucleases are not involved. These results demonstrate that processing of UpsM is catalyzed by RNase E. This raises the question, why RNase E dependent processing occurs only under certain growth conditions. It is highly unlikely that all conditions leading to UpsM processing go along with increased RNase E levels, increased RNase E activity or structural changes in UpsM that promote cleavage by RNase E. It is known that sRNAs are often processed together with their target mRNA [15–18]. The Hfq protein can favor sRNA-mRNA interaction and the subsequent processing [15, 21]. Thus, it is conceivable that induction of the UpsM target(s) is responsible for appearance of the 130 nt processing product under the given conditions. To date (the) target mRNA(s) of UpsM are not identified. To better define the conditions, which promote UpsM processing, we tested the effect of further stress factors and growth conditions on the UpsM pattern (S3B and S3C Fig). Neither SDS, nor ethanol, hydrogen peroxide, tBOOH nor superoxide induced processing of the UpsM transcript (S3B Fig). High NaCl concentrations induced slight processing, while CdCl₂ had a similar effect on processing as $^1\text{O}_2$. Heat shock resulted in very fast UpsM processing (S3C Fig). The processing product was also clearly visible in RNA isolated from stationary phase cultures.

In *R. sphaeroides* the alternative sigma factors RpoHI and RpoHII stimulate many genes in response to stress conditions including $^1\text{O}_2$, CdCl₂, heat and stationary phase [22–26]. Northern blots revealed that normal UpsM processing in presence of $^1\text{O}_2$ occurs in mutants either lacking RpoHI or RpoHII. However, a mutant lacking both sigma factors fails to process UpsM even under $^1\text{O}_2$ stress (Fig 2B). We conclude that interaction with a target RNA very likely

promotes UpsM cleavage by RNase E, whereas the target RNA is transcribed from a promoter, which is recognized by RpoHI as well as by RpoHIII. It is known that the regulons of these two sigma factors indeed overlap [23]. No genes for known RNases are part of the RpoHI/RpoHIII regulon [23, 24, 27] and UpsM heterologously expressed in *E. coli* does not show any induced processing even under stress conditions (data not shown) supporting the assumption of target-dependent processing.

The *Dcw* Gene Cluster Features a Long 5' UTR

We performed reporter assays, 5' RACE (*rapid amplification of cDNA ends*) and RT-PCRs to address the question whether *mraZ* transcription is exclusively dependent on the UpsM promoter and whether the UpsM terminator allows read-throughs in order to guarantee transcription of *mraZ*. Conversely, this means that no additional promoter is localised between the UpsM promoter and *mraZ*.

To test this we performed β -galactosidase activity assays by using reporter plasmids with *mraZ::lacZ* translational fusion and *mraZ* upstream regions of varying length in *R. sphaeroides* (Fig 3A). Strong activity of 130 Miller units was observed with plasmid pPHUmraZUpsM containing the UpsM promoter. However, all shortened upstream regions of *mraZ* (188 nt or 67 nt) led to β -galactosidase activities similar to that observed for the empty vector control (pPHU235) proving the absence of any additional promoter closer to *mraZ*.

To further support this assumption we conducted a 5' RACE to determine 5' ends of *mraZ* mRNAs, whereby cDNA synthesis was enabled by a primer (pUpsM_mraZ_B) binding within the coding region of the *mraZ* gene. Further amplification of cDNA was done with a primer located upstream (pUpsM_B). The resulting PCR products are visible as one clear band on a gel and show the migration behaviour of fragments with 194 bp length which corresponds to the 5' end of UpsM (Fig 3B). The amplified DNA fragments were also subcloned into the pDrive vector without any further purification to determine the 5' ends precisely. The 5' end of UpsM was found in five of ten sequences, whereas the other 5' ends were distributed randomly probably due to technical reasons (Fig 3B). This experiment confirms, 1) that *mraZ* transcription depends on the promoter of UpsM, 2) the terminator of UpsM allows read-throughs leading to *dcw* transcription and 3) the 5' UTR of *mraZ* respectively *dcw* mRNAs is not processed like UpsM, since we used RNA from cells after 90 min $^1\text{O}_2$ stress and did not detect the 5' end of UpsM (130nt).

Furthermore we demonstrated read-throughs at the UpsM terminator and estimated the frequency of such events by RT-PCRs and an unconventional qRT-PCR approach with DNA free RNA from unstressed exponentially grown cultures. The two primer pairs pUpsM_A/B and pmraZ_A/B specifically amplify an UpsM or *mraZ* segment respectively (155 bp and 153 bp), whereas primer pair pUpsM_mraZ_A/B amplifies a segment (143 bp) spanning from UpsM to *mraZ* and therefore detects the read-through. RT-PCR products for all primer pairs are clearly visible on a gel, whereas none of those fragments emerge in control samples without prior reverse transcription (Fig 3C). For a rough estimate of the read-through frequency we compared amplification cycles or rather Cq values in qRT-PCR between all primer pairs in the same RNA samples instead of between different RNA samples with the very same primer pair. Since product sizes are very similar the fluorescent dye will intercalate comparably into *de novo* DNA. The mRNA level of the segment representing the read-through seems to have the lowest abundance. However, this is probably due to a terminator mediated bias and the smaller product size. In Fig 3C the relative transcript levels calculated in comparison to 16S rRNA levels are shown in relation to the transcript level detected by primer pair pUpsM_mraZ. Approximately 700 and 7 times higher RNA levels corresponding to UpsM and *mraZ* respectively were

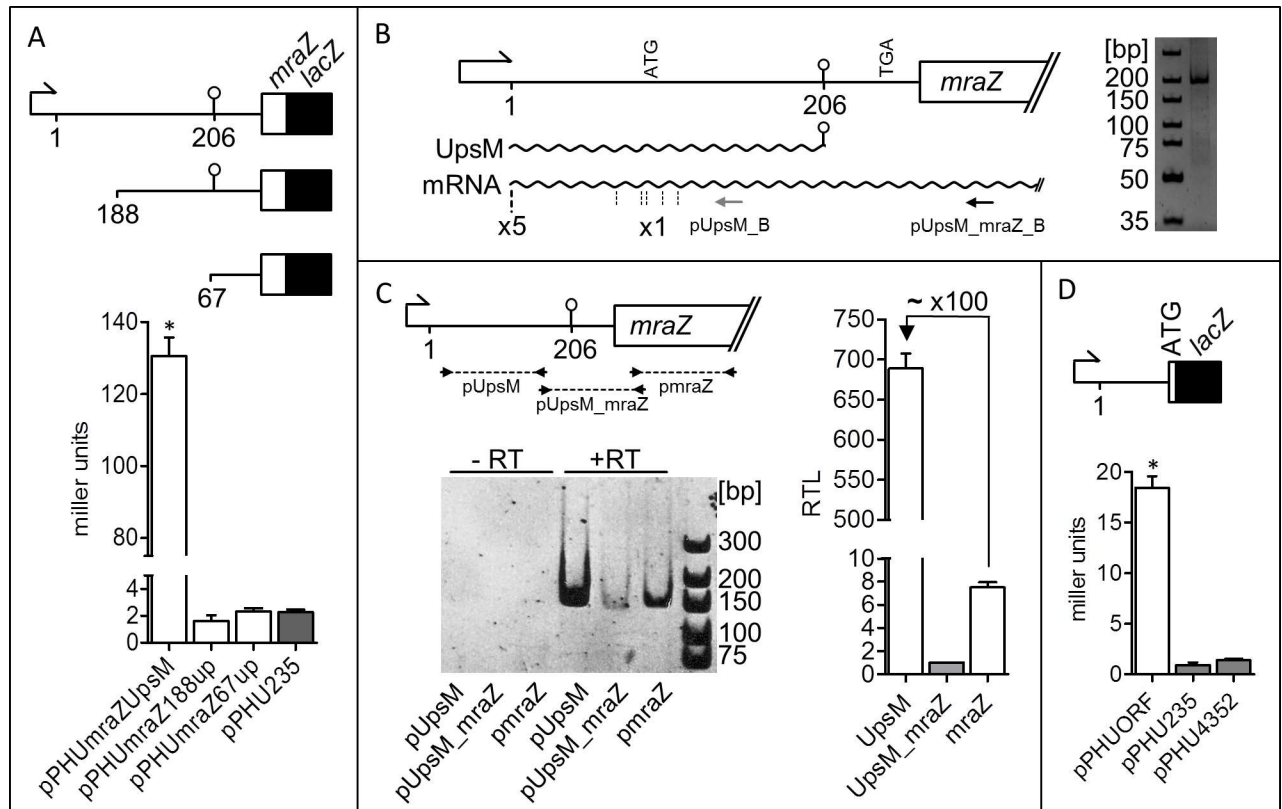


Fig 3. Transcription of *mraZ* is enabled by the UpsM promoter. (A) β -galactosidase activity assays of *R. sphaeroides* 2.4.1 with reporter plasmids with *mraZ*::*lacZ* translational fusion and *mraZ* upstream regions of varying length (long upstream region including the promoter of UpsM, 188 and 67 upstream nucleotides). pPHU235 represents the empty vector control. For each strain, three independent biological experiments with technical duplicates were performed. Error bars indicate standard deviations and an asterisk a significance level of $P < 0.01$ compared to pPHU235. (B) 5' RACE with RNA from *R. sphaeroides* 2.4.1 after 90 min of $^1\text{O}_2$ stress. cDNA was generated with the primer depicted as black arrow, whereas cDNAs were amplified by the primer shown as grey arrow. The PCR product was visualized on a gel (10% PAA/TBE) by ethidium bromide staining. 5' ends (dashed lines) identified by subcloning and sequencing and their corresponding frequencies are highlighted. (C) qRT-PCR products of primer pairs pUpsM, pmraZ (155 bp and 153 bp, both specific for the corresponding mRNA segments) and pUpsM_ *mraZ* (143 bp, spanning from UpsM to *mraZ*) visualized on a gel (10% PAA/TBE) by ethidium bromide staining. Samples without initial RT step were loaded as control. On the right relative transcript levels are shown in relation to the product quantity of primer pair pUpsM_ *mraZ*. qRT-PCRs were performed in technical duplicates with RNA from three biological independent and unstressed *R. sphaeroides* 2.4.1 cultures. Error bars indicate standard deviations. (D) β -galactosidase activity assays of *R. sphaeroides* 2.4.1 conjugated with a reporter plasmid with translational *lacZ* fusion to the start codon (ATG) within the UpsM gene in comparison to the promoter-less empty vector control (pPHU235) and a control plasmid (pPHU4352) containing a strong 16S rRNA promoter. For each strain, three independent biological experiments with technical duplicates were performed. Error bars indicate standard deviations and an asterisk a significance level of $P < 0.01$ compared to both controls.

doi:10.1371/journal.pone.0165694.g003

detected. As demonstrated *mraZ* transcripts have long 5' UTRs and contain the UpsM locus. Therefore part of the DNA amplified by the primer pair pUpsM originates from full-length transcripts also encoding *mraZ*. In other words, *mraZ* transcripts are amplified not only by the primer pair pmraZ but also by pUpsM primers. By taking this into consideration we estimate that one read-through event or rather *mraZ* transcription takes place about once in 100 transcription events under the given experimental conditions.

We were able to show convincingly that *mraZ* has a long 5' UTR under transcriptional control of the UpsM promoter. In addition the 5' UTR contains a start codon upstream and a stop codon in frame downstream of the terminator of UpsM and therefore might encode a leader peptide here designated as sORF (small open reading frame). To test translation initiating at this start codon we performed β -galactosidase activity assays of a reporter plasmid with an

sORF::*lacZ* translational fusion containing the upstream region including P_{UpsM} in *R. sphaeroides*. The sORF::*lacZ* fusion on plasmid pPHUORF resulted in low but significant higher β-galactosidase activities of approximately 18 Miller units in comparison to the promoter-less empty vector control pPHU235 and control plasmid pPHU4352 containing a strong 16S rRNA promoter (Fig 3D). This indicates translation of the ORF, but final proof by a direct detection of the hypothetical peptide is missing.

The 5' UTR of the *Dcw* Gene Cluster in *Rhodobacteraceae* Differs from Other Bacteria

The upstream region of the *mraZ* gene was compared to that of other species. Using public available deep sequencing data obtained from the NCBI SRA database [28], we annotated TSS (transcription start sites) within 300 nt upstream of the gene locus. Reads were mapped to the respective genomes using segemehl [29] with default parameters after quality trimming using Trimmomatic [30] at a quality threshold of 25 in a sliding window of size 3. Only reads of size >14 nt were considered. See S4 Fig for details. Rho-independent Terminators in the 5'-region were predicted using TransTermHP [31]. The highest scoring hit was assumed to be the terminator whenever its MFE (minimum free energy) was below -10 kcal/mol according to RNAfold [32].

A summary is shown in Fig 4. *Rhodobacteraceae* consistently show a long 5' UTR with a strong terminator and no additional TSS close to *mraZ* in a particularly striking manner. In

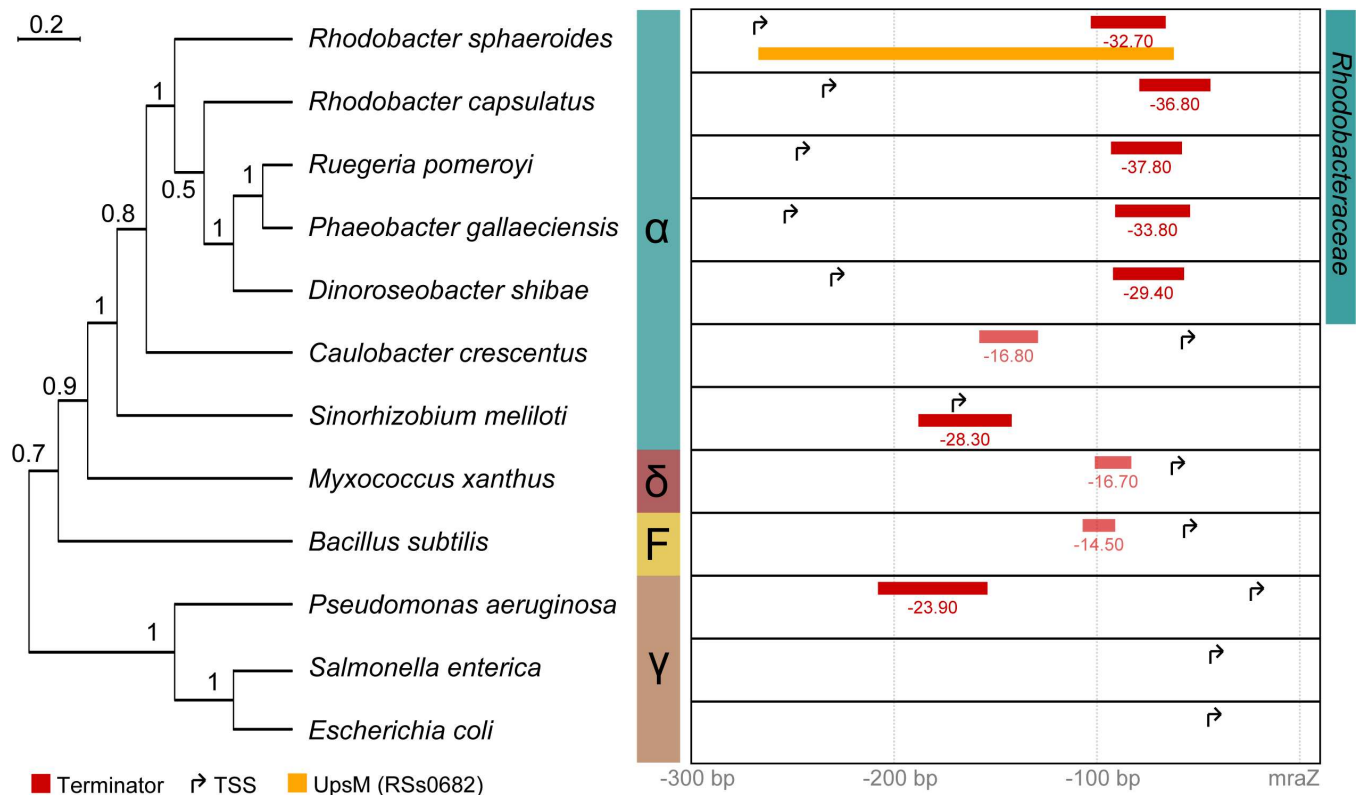


Fig 4. Comprehensive view of *mraZ* upstream regions in different species. Terminator predictions are indicated in red. Respective energies are given in kcal/mol. Regions between Start and Stop codons in frame are shown as grey bars. Transcription start sites are derived from public available deep sequencing data (see S4 Fig for details). The phylogenetic tree was build using clustalx [58] (NJ, 10000 bootstraps) based on a clustalOmega [59] alignment of the respective *mraZ* coding regions. Bootstrap support values are indicated. Seemingly the long *mraZ* 5' UTR with an intrinsic terminator is special to the family of *Rhodobacteraceae*.

doi:10.1371/journal.pone.0165694.g004

contrast, other species apart from *Rhodobacteraceae* typically have an exclusive TSS closer to *mraZ* resulting in 5' URTs of ≤ 70 nt, including the closely related Alphaproteobacterium *Caulobacter crescentus*. Moreover upstream terminators are not found at all or are predicted to be rather weak in those species. A notable exception is *Sinorhizobium meliloti* by exhibiting 5' UTR mediocre in length (170 nt) whereby the TSS is located within a predicted terminator site. However, whether this upstream TSS contributes to the *mraZ* expression as shown for *R. sphaeroides* remains speculative but has never been reported. Taken together our data suggests that a long *mraZ* 5' UTR with intrinsic terminator generating an sRNA combined with no separate TSS for *mraZ* is an exclusive feature of the family of *Rhodobacteraceae*.

Predicted Secondary Structure and Folding Landscape of UpsM

The secondary structure prediction of UpsM was evaluated *in silico* using RNAfold [32]. It consists of four structured regions (R1-4). R1 is located at the 5' end. It consists of several short hairpins isolated from the remaining RNA by a ~11 nt long stem followed by an unpaired region of 10 nt. R2 is a long hairpin with a small bulge, R3 a short hairpin with a stem of only 4 nt. R4 is located at the 3' end of the molecule and corresponds to the predicted, strong terminator structure (Fig 5A).

The secondary structure elements R2/3 and R4 show similarities to typical riboswitches (R2/3 aptamer region, R4 terminator) [33]. Hence we constrained the fold such that the terminator would not form. Strikingly, the refolding event leads to an interaction of R2 and R4 owing a loss of ~20% energy (15.9 kcal/mol) (Fig 5B). Based on this analysis, the *dcw* 5' UTR has the potential to represent a riboswitch.

To verify our assumption, we aligned the 5'-UTR regions of all *Rhodobacteraceae* in our dataset (cropped at the predicted terminator) using mlocarna [34] which respects sequence

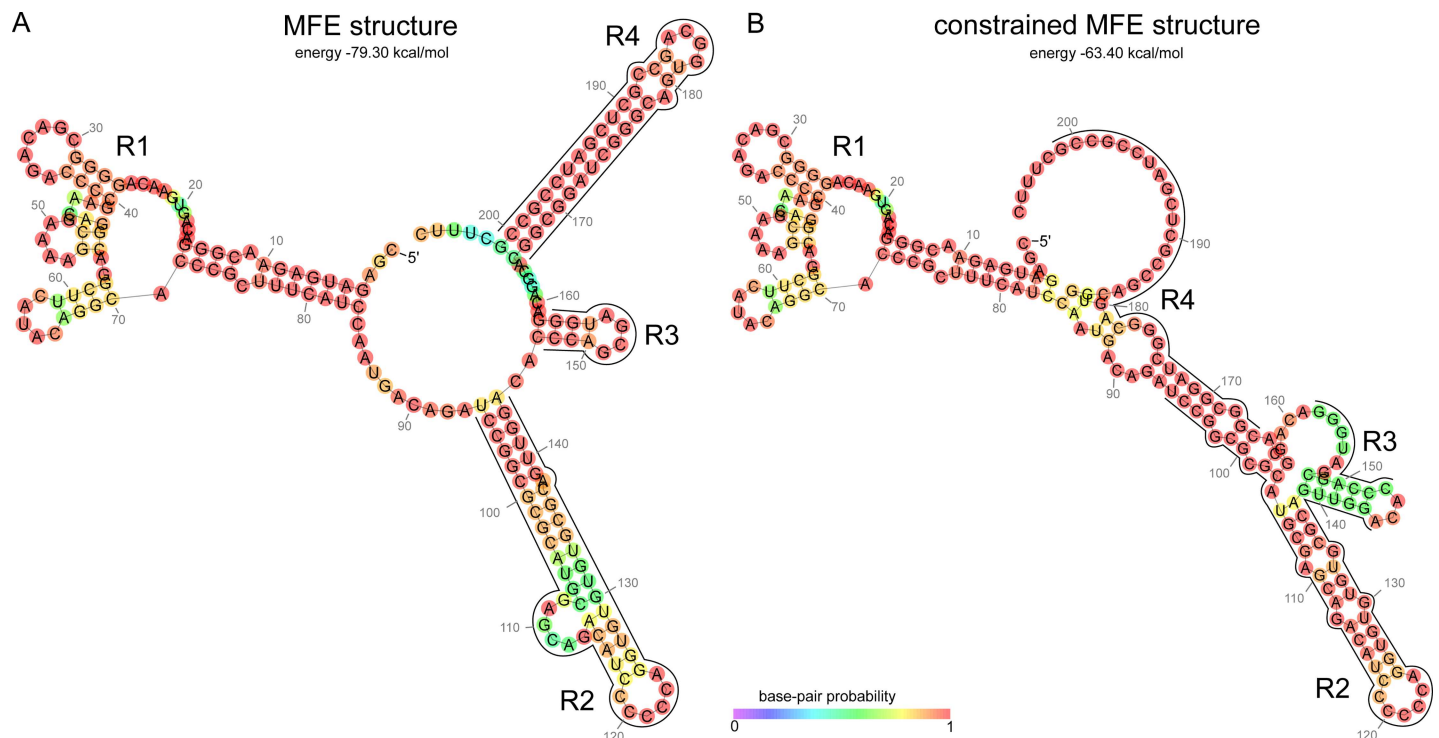


Fig 5. Structural analysis of UpsM. Analogous structured regions are indicated as R1-R4. RNAfold structure of UpsM in *R. sphaeroides* with and without constraint terminator (R4).

doi:10.1371/journal.pone.0165694.g005

and secondary structure at the same time. The alignment was folded using RNAalifold from the Vienna RNA 2.0 package [32] with no constraints and, analogously to above, constraint such that the terminator would not be allowed to form. S5A Fig shows the resulting structures with conservation. S5B Fig indicates the underlying alignment with secondary structure indications.

The consensus structure implies that R1 is hardly conserved. Moreover, the predicted transcription start site is further downstream in other species, leading to shorter transcripts (e.g. 19 nt of R1 are missing in *D. shibae*). We thus assume that R1 is relevant for the function of UpsM. Other than that, the consensus structure well resembles the UpsM structure, indicating that R2-4 are conserved among *Rhodobacteraceae*. Constraint folding however does not induce an interaction of R2 and R4 in this case. Hence, the putative riboswitch function is seemingly not conserved in *Rhodobacteraceae* but, if so, an evolutionary trait special to *R. sphaeroides*.

The folding landscape of UpsM was predicted using RNAsubopt, RNAfold and barriers from the Vienna RNA 2.0 package [32] as well as treekin [35]. A summary is shown in S6 Fig. The four most probable structures of UpsM mainly differ in the terminator hairpin (R4, states 1, 4 and 12) while one variant slightly differs in R1. States 1 and 4 represent the most pronounced folding minima with the highest probability (S6A and S6B Fig). Potential interactions between the structured regions are not observed. These findings indicate a rather stable structure that requires an additional partner (e.g RNA, protein or ligand) for substantial refold.

Discussion

In this study we further characterize the sRNA UpsM (previously RSs0682) and demonstrate that it is derived from the 5' UTR of the mRNA for *mraZ*, the first gene of the *dcw* gene cluster. Overexpression of UpsM leads to a mild growth defect and to a change in the global gene expression pattern as well under aerobic conditions as under photooxidative stress conditions, proving that UpsM is also functional in *trans*. Further experiments need to clarify whether altered mRNA levels are due to direct base pairing or rather to binding of high amounts of Hfq. Analysis of a dRNA-seq dataset of high coverage and Northern blot analysis of TEX treated RNA from $^1\text{O}_2$ stressed cultures confirmed a processing step from UpsM (206nt) to UpsM (130nt). The UpsM processing pattern under different stress conditions and in various mutant strains showed that processing requires the RNA chaperon Hfq, the endoribonuclease RNase E and the alternative sigma factors RpoHI/II. mRNAs being part of the RpoHI/II controlled stress regulon are induced similar to the processing of UpsM upon $^1\text{O}_2$ stress [6, 12, 23], heat stress [24] and in stationary phase (unpublished). Therefore we assume that UpsM is bound and stabilized by Hfq prior to target recognition. However, under stress conditions a target mRNA is expressed which may form a duplex with UpsM mediated by Hfq. In the course of base pairing the structure of UpsM might be altered and becomes susceptible to RNase E cleavage. A similar mechanism was first described for the Hfq dependent sRNA RyhB, which is degraded in an RNase E dependent manner upon binding to the target mRNA *sodB* in *E. coli* [15]. Moreover, an interaction of Hfq with the scaffolding domain of RNase E for the purpose of a recruitment of RNase E to sRNA-mRNA hybrids has been discussed [36] and might also be true for *R. sphaeroides*.

The dRNA-seq data of high coverage based on exponentially growing *R. sphaeroides* cultures, did not show a TSS exclusive for *mraZ*, despite the fact that *mraZ* represents the first gene of the *dcw* gene cluster. Therefore and based on dRNA-seq data we hypothesized that transcription of *mraZ* depends on the UpsM promoter, which implicates that the terminator of UpsM allows read-throughs in order to guarantee transcription of *mraZ* with a long 5' UTR of 268 nt in length. We were able to proof this assumption by 5' RACE and reporter assays with

different fragments of the *mraZ* 5' UTR and estimated the read-through at the UpsM terminator to take place once in 100 transcription events under our experimental conditions. This might ensure sufficient proximal *dcw* gene transcription, since UpsM is the most frequently transcribed sRNA in *R. sphaeroides* [6]. Taken together our data demonstrates that *mraZ* or polycistronic *dcw* gene mRNAs of *R. sphaeroides* feature long 5' UTRs, with the consequence that UpsM has to be reckoned as a 5' UTR derived sRNA rather than an orphan sRNA derived from an IGR as previously assumed [6].

In recent years transcriptome wide identification of sRNAs revealed that apart from intergenic regions especially 3' UTRs serve as a reservoir for sRNAs being part in the Hfq network [37, 38] and their functions *in trans* was reported [39–41]. In comparison, 5' UTR- or riboswitch-derived sRNAs have been described only occasionally and in general without clear functional assignment *in trans*. A handful of putative sRNAs resulting from processed 5' UTRs are mentioned in a transcriptome study of *Yersinia pseudotuberculosis* [42]. For *E. coli* prematurely terminated transcripts from 5' leader sequences of *ybjM*, *ynaE*, *ydfK*, *mdtJ*, *typA*, *yhiL*, and *dinQ* were detected [43], among which *ynaE*, *ydfK* and *ybjM* 5' leader fragments co-immunoprecipitate with Hfq [44]. In the same study an sRNA corresponding to a 5' UTR segment of the *adhE* mRNA was identified, possibly generated by RNase III [44]. A cloning based screen for ncRNAs in *E. coli* led to the discovery of a few 5' UTR derived RNA fragments, which correspond to riboswitches or to be precise L-box, THI box and RFN box elements known to be required for the regulation of mRNA or protein synthesis by attenuation or RBS accessibility [45]. However, neither an association to Hfq nor any other proof for a function of those RNA fragments *in trans* was provided in this study. Recently 14 sRNAs which also might serve as 5' UTRs were predicted in *Vibrio cholera*, but only for Vcr043 an Hfq dependent stability was shown indicating a function *in trans* [46]. To our knowledge and surprisingly only once dual function 5' UTRs were described so far, which act as riboswitch and generate sRNAs in the course of attenuation with regulatory function *in trans*. This was described for SreA and SreB, two S-adenosylmethionine (SAM) riboswitches in *Listeria monocytogenes*. SreA and SreB are transcribed together with downstream genes encoding proteins involved in methionine and cysteine transport or metabolism. Upon binding of the ligand SAM transcription is attenuated causing expression of SreA and SreB as short transcripts, which control expression of the virulence regulator PrfA precisely as sRNAs by binding to the 5' region of its mRNA [47]. The authors assumed SreA/SreB to be just the first example of a novel distinct class of riboswitch derived sRNAs, without expecting that no further example was reported over the past seven years. In this study we were able to describe a similar example, since UpsM can be transcribed with the downstream gene, but is also generated as stable sRNA by an intrinsic terminator in the 5' leader. However, we were only able to provide evidence for a function of UpsM *in trans* by strong Hfq dependency and target dependent cleavage of the sRNA, whereas we cannot present any experimental evidence for a riboswitch at the intrinsic terminator in the leader sequence so far. Its prominent position in the 5' UTR of the *dcw* gene cluster and combined with the potential breakup the terminator structure R4 due to an interaction with R2, however, makes UpsM a reasonable candidate for a riboswitch. Regulatory features and a potential function as riboswitch will be subject to future investigation.

The regulatory features of the *mraZ/dcw* 5' UTR might be even more complex, since we have indications for weak translational activity at an sORF overlapping with the terminator of UpsM. *Cis*-regulatory elements as such sORFs sometimes also termed leader peptides or uORFs (upstream open reading frame) in eukaryotes are often involved in transcriptional or translation attenuation [48, 49]. However, this sORF is only present in *R. sphaeroides* species and presence of the resulting peptide is not unequivocally proven.

The *mraZ/dcw* 5' UTR or rather UpsM may represent not only the second example for a class of riboswitch derived sRNAs, but also is the first extended 5' UTR described for the *dcw* gene cluster, despite outstanding conservation among rod shaped and gram negative bacteria [1–3]. In *E. coli* the first promoter (*mraZ1p*) of the gene cluster (16 genes in total) gives rise to polycistronic transcripts containing a short 38nt long 5' UTR followed by the first gene *mraZ* [4, 7]. However, in this study we have been able to describe a much longer 5' UTR of 268 nt in length featuring *trans*-regulatory and potential *cis*-regulatory elements (summarized in S7 Fig). Therefore we were interested whether we would find similar 5' leaders in the same genetic context of other bacteria. Our results suggest that long 5'UTRs with intrinsic terminators are exclusively present in members of *Rhodobacteraceae*.

Material and Methods

Bacterial Strains and Growth Conditions

Bacterial strains used in this study are listed in S2 Table. Details on their construction are given in S1 File. *R. sphaeroides* strains were cultivated at 32°C in malate minimal-salt medium [50]. To grow the cells aerobically, cultures were either gassed with air in Meplat bottles to attain a concentration of 160 to 180 μM of dissolved oxygen or by continuous shaking of Erlenmeyer flasks containing 20% culture by volume at 140 rpm. For microaerobic growth conditions, having a dissolved oxygen concentration of about 25 μM , Erlenmeyer flasks containing 80% culture by volume were shaken at 140 rpm. For anaerobic growth in the dark in the presence of 60 mM DMSO as electron acceptor we used completely filled screw-cap Meplat bottles, completely filled and sealed with Parafilm. When necessary tetracycline (2 $\mu\text{g ml}^{-1}$), kanamycin (25 $\mu\text{g ml}^{-1}$) or spectinomycin (10 $\mu\text{g ml}^{-1}$) was added to liquid and solid growth media (1.6% agar). Photooxidative stress conditions were generated as described earlier [51], except the final concentration of methylene blue (0.2 μM) (Sigma-Aldrich; M9140). Other stress conditions were generated by a final concentration of 250 mM NaCl, 10 μM CdCl₂, 0.005% SDS, 2.5% ethanol, 300 μM tBOOH, 1 mM H₂O₂ and 250 μM paraquat (O₂⁻) or by temperature shift to 42°C. To culture *E. coli* strains, cells were continuously shaken at 180 rpm in Luria–Bertani medium at 37°C or grown on solid growth media containing 1.6% (w/v) agar. When necessary kanamycin (25 $\mu\text{g ml}^{-1}$) or tetracycline (20 $\mu\text{g ml}^{-1}$), ampicillin (200 $\mu\text{g ml}^{-1}$) or spectinomycin (10 $\mu\text{g ml}^{-1}$) was added to the media.

Northern Blot Analysis

Northern blots were performed as described earlier [6]. Oligodeoxynucleotides used for end-labeling with [γ -³²P]-ATP (Hartmann Analytic; SRP-301) by T4 polynucleotide kinase (Fermentas; #EK0031) are listed in S3 Table. A low stringency Church buffer was used for hybridization [52]. Membranes were washed in 5x SCC buffer + 0.1% SDS. After exposure on phosphoimaging screens (Bio-Rad), images were analyzed by the 1D-Quantity One software (Bio-Rad).

Isolation of Total RNA

Total RNA used for Northern blot, 5' RACE and real time RT-PCR was isolated by the hot phenol method [53]. To remove remaining traces of DNA, samples were treated with 6 U of DNaseI (Invitrogen; #18047019) per 1 μg of RNA. Absence of DNA contamination was confirmed by PCR with primers targeting *gloB* (RSP_0799) (S3 Table).

5' RACE

To determine 5' mRNA ends of *mraZ* using 5' rapid amplification of cDNA ends (RACE), 3 μ g of DNA free total RNA isolated from wild type cells after 90 minutes of $^1\text{O}_2$ stress were reverse transcribed into cDNA by using avian myeloblastosis virus reverse transcriptase (Promega) and gene-specific primer pUpsM_ *mraZ*_B (S3 Table). A second amplification was done with primer pUpsM_B (S3 Table). The 5'RACE protocol was performed as described previously [24].

qRT-PCR

The One-Step Brilliant III QRT-PCR Master Mix Kit (Agilent) was used for reverse transcription and following PCR as described in the manufacturer's manual but in 10 μ l volumes containing 2 μ l DNA free RNA in the concentration 0.2 ng/ μ l. Runs in independent biological triplicates with technical duplicates were done by the use of a Bio-Rad CFX96 Real Time System. Cq values at the auto calculated RFU were extracted with the corresponding software Bio-Rad CFX Manager. mRNA levels were calculated in relation to the mRNA levels of 16S rRNA similar to Pfaffl [54], but with fixed primer efficiencies of 2.0 and a fixed denominator of 1.0, since Cq values of different primer pairs were compared and the very same RNA sample was used. Therefore the resulting formula is: Ratio = $2^{\Delta\text{Cq}(16\text{S} - \text{A})}$, whereas A is Cq of primer pair pUpsM, pUpsM_ *marZ* or pmraZ. Primers are listed in S3 Table.

β -Galactosidase Activity Assay

β -galactosidase activity was measured in conjugants obtained after transferring the respective reporter plasmids via di-parental conjugation from *E. coli* to *R. sphaeroides*. Three independent liquid cultures, inoculated with equal numbers of colonies, were grown microaerobically and diluted to OD₆₆₀ 0.2, before reaching stationary phase. Three samples of 1 ml were collected in early exponential growth phase (OD₆₆₀ 0.4). Measurements of β -galactosidase activity were carried out as described previously [22].

RNA Treatment, Library Preparation and Sequencing

DNA free total RNA isolated from exponential and microaerobic cultures were treated with TEX (Epicentre #TER51020) to enrich primary transcripts [55]. For *R. sphaeroides* RNA Illumina cDNA libraries were prepared by vertis Biotechnology AG, Germany (<http://www.vertis-biotech.com/>) as described before without prior RNA fragmentation or size fractionation [50]. Illumina cDNA libraries resulting from TEX treated *R. capsulatus* RNA were generated as described before without prior rRNA depletion [42]. cDNA libraries were sequenced on a HiSeq 2000 or HiSeq 2500 machine in single-read mode running 100 cycles. Raw files for *R. sphaeroides* have been deposited in the National Center for Biotechnology Information Gene Expression Omnibus (GEO) [56] and are accessible via the GEO accession GSE71844. Raw files for *R. capsulatus* are accessible via BioProject Accession PRJNA343088.

Microarray Analysis

Microarray analysis was performed as described before [26, 57]. Total RNA, obtained from 6 independent cultures per strain later hybridized with a duplicate of arrays was chemically labeled with Cy3 and Cy5 (Kreatech; EA-022/EA-023), respectively. Multiarray analysis was performed with the Bio-conductor package Limma for R. On the basis of calculated MA plots, genes were considered reliable if the average signal intensity [A-value: $1/2 \log_2(\text{Cy3} \times \text{Cy5})$] was ≥ 12 . To filter out potentially insignificant changes among genes that passed the reliability

criterion, a cutoff value was applied; i.e., those genes were retained whose average expression value of the overexpression strain (a) compared with the average value of the control treatment (b) was either a \log_2 fold change of ≥ 0.65 or ≤ -0.65 . Microarray data are deposited in the Gene Expression Omnibus (GSE87789).

Supporting Information

S1 Fig. Nucleotide sequence of the *mraZ* 5' UTR. Sequence of UpsM is shaded in grey. The corresponding terminator is underlined. The hypothetical sORF coding region starting with ATG and the *mraZ* coding region starting with GTG are depicted by bold letters.

(TIF)

S2 Fig. (A) Altered UpsM transcript level shown by Northern blot analysis of total RNA of the overexpression strain *R. sphaeroides* 2.4.1 pBBRUpsMx2 after 60 min $^1\text{O}_2$ stress in comparison to the wild-type strain harboring the empty vector (pBBR1MCS2). Signals of 5S rRNA serve as loading control. (B) Aerobic, microaerobic and anaerobic growth of the overexpression strain *R. sphaeroides* 2.4.1 pBBRUpsMx2 in comparison to the wild-type strain harboring the empty vector (pBBR1MCS2). The optical density at 660 nm (OD_{660}) was determined over time, and growth is indicated as continuous line. All graphs represent the mean of three biological independent experiments. Error bars indicate the standard deviation at each time point measured.

(TIF)

S3 Fig. Processing pattern of UpsM under various ongoing stress conditions and in strains lacking RNases shown by Northern blot analysis of total RNA isolated from *R. sphaeroides*.

(A) Processing pattern of UpsM in strains lacking RNase III (Δrnc) or RNase J (Δrnj). Signals of 5S rRNA serve as loading control. (B) Stress conditions were generated by a final concentrations of 0.005% SDS, 2.5% ethanol, 300 μM tBOOH, 1mM H_2O_2 and 250 μM paraquat (O_2^-). (C) Stress conditions were generated by a final concentrations of 0.2 μM methylene blue in the presence of 800 Wm^{-2} white light ($^1\text{O}_2$), 250 mM NaCl and 10 μM CdCl_2 or by stationary phase or growth under heat stress at 42°C.

(TIF)

S4 Fig. Genomic regions upstream of *mraZ* in other species. Underlying deep sequencing data was used to predict transcriptional start sites. Layout resembles Fig 4. Genome IDs are indicated at the y-axis. Sources are indicated at the right.

(TIF)

S5 Fig. Structural analysis of UpsM. Analogous structured regions are indicated as R1-R4.

(A) RNAfold structure of UpsM in *R. sphaeroides* with and without constraint terminator (R4) and consensus structure of aligned sequences for all *Rhodobacteraceae* without constraint terminator (R4). (B) RNAalifold alignment with structural annotation and indicated terminator constraint (x = bases forced to be unpaired). An interaction of R2 and R4 occurs when applying terminator constraints. This is however not resembled by the consensus structure.

(TIF)

S6 Fig. Folding state analysis of UpsM. (a) Barrier tree of all suboptimal structures of the unprocessed sRNA. Four main species with favorable energies were identified (1, 4, 8, 12). (B) Population density of these structures over time (no unit). (C) Structural representations of the four most favorable states.

(TIF)

S7 Fig. Model summarizing the results of this publication. The 268 nt long 5' UTR of *mraZ*, first gene of the *dcw* (division and cell wall) gene cluster, comprises a Rho independent terminator, which in case of termination gives rise to the 206 nt long non-coding RNA UpsM (upstream srRNA *mraZ*). Under stress conditions this sRNA is conditionally cleaved by RNase E in an Hfq- and likely in a target mRNA-dependent manner, whereas the corresponding target mRNA is controlled by an RpHI/II dependent promoter.
(TIF)

S1 File. Strain construction.

(PDF)

S1 Table. Gene expression of an UpsM overexpression strain *R. sphaeroides* 2.4.1 pBBRUpsMx2 was analysed in comparison to the strain *R. sphaeroides* 2.4.1 pBBR1MCS2 harbouring the empty vector to get first insights into the biological function of UpsM. The transcriptome of both strains was compared by microarray analysis during exponential growth under aerobic and non-stress conditions and after 90 min of $^1\text{O}_2$ stress. For both conditions a biological duplicate of arrays was hybridized with RNA from three biological independent cultures per strain. A Pearson correlation coefficient between the replica of 0.97 and 0.95 was calculated. Changes in expression levels of protein-coding genes passing the selection criteria of microarray analysis, which is a reliable A-value ≥ 12 and a log2 fold change of > 0.65 or < -0.65 between the two strains, are shown.

(DOCX)

S2 Table. Strains and plasmids used in this study

(DOCX)

S3 Table. Oligonucleotides used in this study.

(DOCX)

Acknowledgments

We thank Kerstin Habertzettl for help with construction of the *rne* mutant strain and Sven Findeiss for discussion of the folding state analysis.

Author Contributions

Conceptualization: LW GK.

Data curation: GK ML.

Formal analysis: LW CT ML BR.

Funding acquisition: GK.

Investigation: LW BR MV CT ML.

Methodology: LW.

Project administration: GK LW.

Resources: GK ML.

Supervision: LW GK.

Validation: GK LW.

Visualization: LW.

Writing – original draft: LW ML GK.

Writing – review & editing: LW GK.

References

1. Tamames J. Evolution of gene order conservation in prokaryotes. *Genome Biol.* 2001; 2(6): RESEARCH0020. PMID: [11423009](#); PubMed Central PMCID: PMC33396.
2. Tamames J, Gonzalez-Moreno M, Mingorance J, Valencia A, Vicente M. Bringing gene order into bacterial shape. *Trends Genet.* 2001; 17(3):124–6. PMID: [11226588](#).
3. Mingorance J, Tamames J, Vicente M. Genomic channeling in bacterial cell division. *J Mol Recognit.* 2004; 17(5):481–7. doi: [10.1002/jmr.718](#) PMID: [15362108](#).
4. Vicente M, Gomez MJ, Ayala JA. Regulation of transcription of cell division genes in the *Escherichia coli* *dcw* cluster. *Cell Mol Life Sci.* 1998; 54(4):317–24. PMID: [9614967](#). doi: [10.1007/s000180050158](#)
5. de la Fuente A, Palacios P, Vicente M. Transcription of the *Escherichia coli* *dcw* cluster: evidence for distal upstream transcripts being involved in the expression of the downstream *ftsZ* gene. *Biochimie.* 2001; 83(1):109–15. PMID: [11254983](#).
6. 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. doi: [10.1111/j.1365-2958.2009.06949.x](#) PMID: [19906181](#).
7. Eraso JM, Markillie LM, Mitchell HD, Taylor RC, Orr G, Margolin W. The Highly Conserved MraZ Protein Is a Transcriptional Regulator in *Escherichia coli*. *J Bacteriol.* 2014; 196(11):2053–66. doi: [10.1128/Jb.01370-13](#) PMID: [WOS:000335909500014](#).
8. Hara H, Yasuda S, Horiuchi K, Park JT. A promoter for the first nine genes of the *Escherichia coli* *mra* cluster of cell division and cell envelope biosynthesis genes, including *ftsI* and *ftsW*. *J Bacteriol.* 1997; 179(18):5802–11. PMID: [9294438](#); PubMed Central PMCID: PMC179470.
9. Selinger DW, Saxena RM, Cheung KJ, Church GM, Rosenow C. Global RNA half-life analysis in *Escherichia coli* reveals positional patterns of transcript degradation. *Genome Res.* 2003; 13(2):216–23. doi: [10.1101/gr.912603](#) PMID: [12566399](#); PubMed Central PMCID: PMC420366.
10. Mengin-Lecreulx D, Ayala J, Bouhss A, van Heijenoort J, Parquet C, Hara H. Contribution of the *Pmra* promoter to expression of genes in the *Escherichia coli* *mra* cluster of cell envelope biosynthesis and cell division genes. *J Bacteriol.* 1998; 180(17):4406–12. PMID: [9721276](#); PubMed Central PMCID: PMC107448.
11. Carrion M, Gomez MJ, Merchante-Schubert R, Dongarra S, Ayala JA. *mraW*, an essential gene at the *dcw* cluster of *Escherichia coli* codes for a cytoplasmic protein with methyltransferase activity. *Biochimie.* 1999; 81(8–9):879–88. PMID: [10572301](#).
12. 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. doi: [10.1111/j.1365-2958.2011.07658.x](#) PMID: [21535243](#).
13. Belasco JG. All things must pass: contrasts and commonalities in eukaryotic and bacterial mRNA decay. *Nat Rev Mol Cell Biol.* 2010; 11(7):467–78. doi: [10.1038/nrm2917](#) PMID: [20520623](#); PubMed Central PMCID: PMC3145457.
14. Carpousis AJ. The RNA degradosome of *Escherichia coli*: An mRNA-degrading machine assembled on RNase E. *Annu Rev Microbiol.* 2007; 61:71–87. doi: [10.1146/annurev.micro.61.080706.093440](#) PMID: [WOS:000250965600006](#).
15. Masse E, Escorcía FE, Gottesman S. Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli*. *Genes Dev.* 2003; 17(19):2374–83. doi: [10.1101/gad.1127103](#) PMID: [12975324](#); PubMed Central PMCID: PMC218075.
16. Pfeiffer V, Papenfort K, Lucchini S, Hinton JC, Vogel J. Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nat Struct Mol Biol.* 2009; 16(8):840–6. doi: [10.1038/nsmb.1631](#) PMID: [19620966](#).
17. Prevost K, Desnoyers G, Jacques JF, Lavoie F, Masse E. Small RNA-induced mRNA degradation achieved through both translation block and activated cleavage. *Genes Dev.* 2011; 25(4):385–96. doi: [10.1101/gad.2001711](#) PMID: [21289064](#); PubMed Central PMCID: PMC3042161.
18. Saramago M, Barria C, Dos Santos RF, Silva IJ, Pobre V, Domingues S, et al. The role of RNases in the regulation of small RNAs. *Curr Opin Microbiol.* 2014; 18:105–15. doi: [10.1016/j.mib.2014.02.009](#) PMID: [24704578](#).

19. Apirion D. Isolation, genetic mapping and some characterization of a mutation in *Escherichia coli* that affects the processing of ribonucleic acid. *Genetics*. 1978; 90(4):659–71. PMID: [369943](#); PubMed Central PMCID: PMC1213911.
20. Goldblum K, Apirion D. Inactivation of the ribonucleic acid-processing enzyme ribonuclease E blocks cell division. *J Bacteriol*. 1981; 146(1):128–32. PMID: [6163761](#); PubMed Central PMCID: PMC217061.
21. Waters LS, Storz G. Regulatory RNAs in bacteria. *Cell*. 2009; 136(4):615–28. doi: [10.1016/j.cell.2009.01.043](#) PMID: [19239884](#); PubMed Central PMCID: PMC3132550.
22. Nuss AM, Adnan F, Weber L, Berghoff BA, 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(11):e79520. doi: [10.1371/journal.pone.0079520](#) PMID: [24223961](#); PubMed Central PMCID: PMC3818230.
23. Nuss AM, Glaeser J, Berghoff BA, Klug G. Overlapping alternative sigma factor regulons in the response to singlet oxygen in *Rhodobacter sphaeroides*. *J Bacteriol*. 2010; 192(10):2613–23. doi: [10.1128/JB.01605-09](#) PMID: [20304993](#); PubMed Central PMCID: PMC2863570.
24. Nuss AM, 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(1):220–30. doi: [10.1128/JB.00925-08](#) PMID: [18978062](#); PubMed Central PMCID: PMC2612413.
25. Billenkamp F, Peng T, 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. doi: [10.1128/JB.02475-14](#) PMID: [25777678](#); PubMed Central PMCID: PMC4402390.
26. 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. doi: [10.1080/15476286.2015.1031948](#) PMID: [25833751](#); PubMed Central PMCID: PMC4615379.
27. Dufour YS, Imam S, Koo BM, Green HA, Donohue TJ. Convergence of the transcriptional responses to heat shock and singlet oxygen stresses. *PLoS Genet*. 2012; 8(9):e1002929. doi: [10.1371/journal.pgen.1002929](#) PMID: [23028346](#); PubMed Central PMCID: PMC3441632.
28. Leinonen R, Sugawara H, Shumway M, International Nucleotide Sequence Database C. The sequence read archive. *Nucleic Acids Res*. 2011; 39(Database issue):D19–21. doi: [10.1093/nar/gkq1019](#) PMID: [21062823](#); PubMed Central PMCID: PMC3013647.
29. Hoffmann S, Otto C, Kurtz S, Sharma CM, Khaitovich P, Vogel J, et al. Fast mapping of short sequences with mismatches, insertions and deletions using index structures. *PLoS Comput Biol*. 2009; 5(9):e1000502. doi: [10.1371/journal.pcbi.1000502](#) PMID: [19750212](#); PubMed Central PMCID: PMC2730575.
30. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014; 30(15):2114–20. doi: [10.1093/bioinformatics/btu170](#) PMID: [24695404](#); PubMed Central PMCID: PMC4103590.
31. Kingsford CL, Ayanbule K, Salzberg SL. Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. *Genome Biol*. 2007; 8(2):R22. doi: [10.1186/gb-2007-8-2-r22](#) PMID: [17313685](#); PubMed Central PMCID: PMC1852404.
32. Lorenz R, Bernhart SH, Honer Zu Siederdisen C, Tafer H, Flamm C, Stadler PF, et al. ViennaRNA Package 2.0. *Algorithms Mol Biol*. 2011; 6:26. doi: [10.1186/1748-7188-6-26](#) PMID: [22115189](#); PubMed Central PMCID: PMC3319429.
33. Garst AD, Edwards AL, Batey RT. Riboswitches: structures and mechanisms. *Cold Spring Harb Perspect Biol*. 2011; 3(6). doi: [10.1101/cshperspect.a003533](#) PMID: [20943759](#); PubMed Central PMCID: PMC3098680.
34. Will S, Joshi T, Hofacker IL, Stadler PF, Backofen R. LocARNA-P: accurate boundary prediction and improved detection of structural RNAs. *RNA*. 2012; 18(5):900–14. doi: [10.1261/rna.029041.111](#) PMID: [22450757](#); PubMed Central PMCID: PMC3334699.
35. Wolfinger MT, Svrcek-Seiler WA, Flamm C, Hofacker IL, Stadler PF. Efficient computation of RNA folding dynamics. *J Phys A Math Gen*. 2004; 37(17):4731–41. Pii S0305-4470(04)73887-1 doi: [10.1088/0305-4470/37/17/005](#) PMID: [WOS:000221482800006](#).
36. Ikeda Y, Yagi M, Morita T, Aiba H. Hfq binding at RhlB-recognition region of RNase E is crucial for the rapid degradation of target mRNAs mediated by sRNAs in *Escherichia coli*. *Mol Microbiol*. 2011; 79(2):419–32. doi: [10.1111/j.1365-2958.2010.07454.x](#) PMID: [WOS:000286114200013](#).
37. Chao Y, Papenfort K, Reinhardt R, Sharma CM, Vogel J. An atlas of Hfq-bound transcripts reveals 3' UTRs as a genomic reservoir of regulatory small RNAs. *EMBO J*. 2012; 31(20):4005–19. doi: [10.1038/emboj.2012.229](#) PMID: [22922465](#); PubMed Central PMCID: PMC3474919.

38. Miyakoshi M, Chao Y, Vogel J. Regulatory small RNAs from the 3' regions of bacterial mRNAs. *Curr Opin Microbiol.* 2015; 24:132–9. doi: [10.1016/j.mib.2015.01.013](https://doi.org/10.1016/j.mib.2015.01.013) PMID: [25677420](https://pubmed.ncbi.nlm.nih.gov/25677420/).
39. 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(3):352–63. doi: [10.1016/j.molcel.2015.12.023](https://doi.org/10.1016/j.molcel.2015.12.023) PMID: [26805574](https://pubmed.ncbi.nlm.nih.gov/26805574/).
40. Miyakoshi M, Chao YJ, Vogel J. Cross talk between ABC transporter mRNAs via a target mRNA-derived sponge of the GcvB small RNA. *EMBO J.* 2015; 34(11):1478–92. doi: [10.15252/embj.201490546](https://doi.org/10.15252/embj.201490546) PMID: [WOS:000355996900007](https://pubmed.ncbi.nlm.nih.gov/WOS:000355996900007/).
41. Kim HM, Shin JH, Cho YB, Roe JH. Inverse regulation of Fe- and Ni-containing SOD genes by a Fur family regulator Nur through small RNA processed from 3'UTR of the *sodF* mRNA. *Nucleic Acids Res.* 2014; 42(3):2003–14. doi: [10.1093/nar/gkt1071](https://doi.org/10.1093/nar/gkt1071) PMID: [24234448](https://pubmed.ncbi.nlm.nih.gov/24234448/); PubMed Central PMCID: [PMC3919588](https://pubmed.ncbi.nlm.nih.gov/PMC3919588/).
42. Nuss AM, Heroven AK, Waldmann B, Reinkensmeier J, Jarek M, Beckstette M, et al. Transcriptomic profiling of *Yersinia pseudotuberculosis* reveals reprogramming of the Crp regulon by temperature and uncovers Crp as a master regulator of small RNAs. *PLoS Genet.* 2015; 11(3):e1005087. doi: [10.1371/journal.pgen.1005087](https://doi.org/10.1371/journal.pgen.1005087) PMID: [25816203](https://pubmed.ncbi.nlm.nih.gov/25816203/); PubMed Central PMCID: [PMC4376681](https://pubmed.ncbi.nlm.nih.gov/PMC4376681/).
43. Raghavan R, Groisman EA, Ochman H. Genome-wide detection of novel regulatory RNAs in *E. coli*. *Genome Res.* 2011; 21(9):1487–97. doi: [10.1101/gr.119370.110](https://doi.org/10.1101/gr.119370.110) PMID: [21665928](https://pubmed.ncbi.nlm.nih.gov/21665928/); PubMed Central PMCID: [PMC3166833](https://pubmed.ncbi.nlm.nih.gov/PMC3166833/).
44. Bilusic I, Popitsch N, Rescheneder P, Schroeder R, Lybecker M. Revisiting the coding potential of the *E. coli* genome through Hfq co-immunoprecipitation. *RNA Biol.* 2014; 11(5):641–54. PMID: [24922322](https://pubmed.ncbi.nlm.nih.gov/24922322/); PubMed Central PMCID: [PMC4152368](https://pubmed.ncbi.nlm.nih.gov/PMC4152368/). doi: [10.4161/rna.29299](https://doi.org/10.4161/rna.29299)
45. Kawano M, Reynolds AA, Miranda-Rios J, Storz G. Detection of 5'- and 3'-UTR-derived small RNAs and cis-encoded antisense RNAs in *Escherichia coli*. *Nucleic Acids Res.* 2005; 33(3):1040–50. doi: [10.1093/nar/gki256](https://doi.org/10.1093/nar/gki256) PMID: [15718303](https://pubmed.ncbi.nlm.nih.gov/15718303/); PubMed Central PMCID: [PMC549416](https://pubmed.ncbi.nlm.nih.gov/PMC549416/).
46. Papenfort K, Forstner KU, Cong JP, Sharma CM, Bassler BL. Differential RNA-seq of *Vibrio cholerae* identifies the VqmR small RNA as a regulator of biofilm formation. *Proc Natl Acad Sci U S A.* 2015; 112(7):E766–75. doi: [10.1073/pnas.1500203112](https://doi.org/10.1073/pnas.1500203112) PMID: [25646441](https://pubmed.ncbi.nlm.nih.gov/25646441/); PubMed Central PMCID: [PMC4343088](https://pubmed.ncbi.nlm.nih.gov/PMC4343088/).
47. Loh E, Dussurget O, Gripenland J, Vaitkevicius K, Tiensuu T, Mandin P, et al. A trans-acting riboswitch controls expression of the virulence regulator PrfA in *Listeria monocytogenes*. *Cell.* 2009; 139(4):770–9. doi: [10.1016/j.cell.2009.08.046](https://doi.org/10.1016/j.cell.2009.08.046) PMID: [19914169](https://pubmed.ncbi.nlm.nih.gov/19914169/).
48. Lovett PS, Rogers EJ. Ribosome regulation by the nascent peptide. *Microbiol Rev.* 1996; 60(2):366–85. PMID: [8801438](https://pubmed.ncbi.nlm.nih.gov/8801438/); PubMed Central PMCID: [PMC239448](https://pubmed.ncbi.nlm.nih.gov/PMC239448/).
49. Henkin TM, Yanofsky C. Regulation by transcription attenuation in bacteria: how RNA provides instructions for transcription termination/antitermination decisions. *Bioessays.* 2002; 24(8):700–7. doi: [10.1002/bies.10125](https://doi.org/10.1002/bies.10125) PMID: [WOS:000177092500005](https://pubmed.ncbi.nlm.nih.gov/WOS:000177092500005/).
50. Remes B, Berghoff BA, Forstner KU, Klug G. Role of oxygen and the OxyR protein in the response to iron limitation in *Rhodobacter sphaeroides*. *BMC Genomics.* 2014; 15:794. doi: [10.1186/1471-2164-15-794](https://doi.org/10.1186/1471-2164-15-794) PMID: [25220182](https://pubmed.ncbi.nlm.nih.gov/25220182/); PubMed Central PMCID: [PMC4176601](https://pubmed.ncbi.nlm.nih.gov/PMC4176601/).
51. 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.
52. Church GM, Gilbert W. Genomic sequencing. *Proc Natl Acad Sci U S A.* 1984; 81(7):1991–5. PMID: [6326095](https://pubmed.ncbi.nlm.nih.gov/6326095/); PubMed Central PMCID: [PMC345422](https://pubmed.ncbi.nlm.nih.gov/PMC345422/).
53. 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. doi: [10.1111/j.1574-6968.1986.tb01270.x](https://doi.org/10.1111/j.1574-6968.1986.tb01270.x)
54. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 2001; 29(9):e45. PMID: [11328886](https://pubmed.ncbi.nlm.nih.gov/11328886/); PubMed Central PMCID: [PMC55695](https://pubmed.ncbi.nlm.nih.gov/PMC55695/).
55. Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiss S, Sittka A, et al. The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature.* 2010; 464(7286):250–5. doi: [10.1038/nature08756](https://doi.org/10.1038/nature08756) PMID: [20164839](https://pubmed.ncbi.nlm.nih.gov/20164839/).
56. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 2002; 30(1):207–10. PMID: [11752295](https://pubmed.ncbi.nlm.nih.gov/11752295/); PubMed Central PMCID: [PMC99122](https://pubmed.ncbi.nlm.nih.gov/PMC99122/).
57. Mank NN, Berghoff BA, Hermanns YN, Klug G. Regulation of bacterial photosynthesis genes by the small noncoding RNA PcrZ. *Proc Natl Acad Sci U S A.* 2012; 109(40):16306–11. doi: [10.1073/pnas.1207067109](https://doi.org/10.1073/pnas.1207067109) PMID: [22988125](https://pubmed.ncbi.nlm.nih.gov/22988125/); PubMed Central PMCID: [PMC3479615](https://pubmed.ncbi.nlm.nih.gov/PMC3479615/).

58. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. *Bioinformatics*. 2007; 23(21):2947–8. doi: [10.1093/bioinformatics/btm404](https://doi.org/10.1093/bioinformatics/btm404) PMID: [17846036](https://pubmed.ncbi.nlm.nih.gov/17846036/).
59. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol*. 2011; 7:539. doi: [10.1038/msb.2011.75](https://doi.org/10.1038/msb.2011.75) PMID: [21988835](https://pubmed.ncbi.nlm.nih.gov/21988835/); PubMed Central PMCID: PMC3261699.