



HAL
open science

Function and mechanism of action of the small regulatory RNA ArcZ in Enterobacterales .

Quentin Dubois, Typhaine Brual, Charlotte Oriol, Pierre Mandin, Guy Condemine, Erwan Gueguen

► To cite this version:

Quentin Dubois, Typhaine Brual, Charlotte Oriol, Pierre Mandin, Guy Condemine, et al.. Function and mechanism of action of the small regulatory RNA ArcZ in Enterobacterales .. RNA, In press, 10.1261/rna.080010.124 . hal-04649459

HAL Id: hal-04649459

<https://hal.science/hal-04649459v1>

Submitted on 16 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 Function and mechanism of action of the small regulatory 2 RNA ArcZ in *Enterobacterales*

3 Quentin Dubois¹, Typhaine Brual¹, Charlotte Oriol², Pierre Mandin², Guy Condemine¹ and Erwan
4 Gueguen^{1*}

5 ¹Univ Lyon, Université Claude Bernard Lyon1, CNRS, INSA Lyon, UMR5240 MAP LYON, France

6 ²CNRS, Aix-Marseille Université, Laboratoire de Chimie Bactérienne, UMR7283, IMM, IM2B, F-13009
7 Marseille, France

8
9 *corresponding author: erwan.gueguen@univ-lyon1.fr

10 11 12 Abstract

13
14
15 ArcZ is a small regulatory RNA conserved in *Enterobacterales*. It is a Hfq-dependent RNA that
16 is cleaved by RNase E in a processed form of 55 to 60 nucleotides. This processed form is highly
17 conserved for controlling the expression of target mRNAs. ArcZ expression is induced by abundant
18 oxygen levels and reaches its peak during the stationary growth phase. This control is mediated by
19 the oxygen-responsive two-component system ArcAB, leading to the repression of *arcZ* transcription
20 under low-oxygen conditions in most bacteria in which it has been studied. ArcZ displays multiple
21 targets, and it can control up to 10% of a genome and interact directly with more than 300 mRNAs in
22 *Escherichia coli* and *Salmonella enterica*. ArcZ displays a multi-faceted ability to regulate its targets
23 through diverse mechanisms such as RNase recruitment, modulation of ribosome accessibility on the
24 mRNA and interaction with translational enhancing regions. By influencing stress response, motility
25 and virulence through the regulation of master regulators such as FlhDC or RpoS, ArcZ emerges as a
26 major orchestrator of cell physiology within *Enterobacterales*.

27
28
29 **Keywords:** ArcZ, sRNA, post-transcriptional regulation, virulence, stress response, RNase E
30

31 INTRODUCTION

32

33 In bacteria, adaptation to environmental changes, production of secondary metabolites or
34 virulence are highly controlled and regulated processes that allow the genes involved in these
35 pathways to be expressed at the right time (Schneider et al. 2009). Such control can occur directly on
36 the DNA via transcriptional regulation, on the mRNA via post-transcriptional regulation, or on the
37 proteins themselves via post-translational regulation. Major classes of post-transcriptional regulators
38 are riboswitches and antisense RNAs (asRNAs) that act in cis or small non-coding RNAs that act in
39 trans (sRNAs). sRNAs are the most abundant class of post-transcriptional regulators (reviewed in
40 Eichner et al. 2022).

41 asRNAs are transcribed in the reverse sense of the target gene. They pair with the mRNA, and
42 this pairing leads to the degradation of both RNAs (Dadzie et al. 2013). On the other hand, sRNAs are
43 small untranslated RNAs, generally short, that regulate their mRNA targets post-transcriptionally
44 (reviewed in Eichner et al. 2022). These sRNAs base pair directly to target mRNAs, often with the help
45 of a chaperone like Hfq or ProQ. In most cases, these sRNAs pair to the 5' untranslated region (UTR)
46 region of the target mRNA. In this way, sRNAs can have a repressive effect on mRNA expression,
47 either by hybridizing to the ribosome binding site (RBS) and/or by recruiting an RNase to degrade the
48 mRNA. They can also have an activating effect on mRNA expression by stabilizing and protecting
49 mRNAs, or by activating translation initiation, for instance by preventing the formation of a
50 translation-inhibiting structure (reviewed in Dutta and Srivastava 2018).

51 The study of sRNA function has taken a major turn since the advent of new high-throughput
52 sequencing methods. These new massive cDNA sequencing techniques have led to the invention of
53 global methods, RIP-seq, CLIP-seq, CLASH, RIL-seq, GRIL-seq, MAPS, and Term-seq, which enable
54 the study of the full range of interactions between Hfq, sRNAs and their target mRNAs that occur *in*
55 *vivo* (reviewed in Saliba et al. 2017).

56 ArcZ is a small RNA found in most *Enterobacteriales* bacteria (Fig. 1A) and its role has been
57 more specifically studied in bacteria such as *Salmonella typhimurium*, *Dickeya dadantii*, *Escherichia*

58 *coli*, *Erwinia amylovora* and *Photorhabdus sp.* ArcZ was identified in initial genome-wide sRNA
59 screenings of *E. coli* and named either SraH (Argaman et al. 2001) or RyhA (Wassarman et al. 2001).
60 The sequence of this small RNA is relatively well conserved among the different species in which it
61 occurs, especially at the 3' end (Fig. 1A).

62 The transcriptional start site (TSS) of ArcZ has been identified precisely in *Pantoea ananatis*,
63 *E. coli* K12, *S. typhimurium*, *Klebsiella pneumoniae*, *Photorhabdus laumondii* and *Xenorhabdus*
64 *nematophila* by 5'RACE or Genome-Wide Transcription Start Site Profiling (Fig. 1A, +1 indicated with
65 a red frame) (Argaman et al. 2001; Kim et al. 2012; Kröger et al. 2012; Shin et al. 2019; Neubacher et
66 al. 2020). It usually ranges in size from 121 to 130 nucleotides, except in *Pantoea ananatis* where the
67 5' end is particularly long (Soper et al. 2010; Shin et al. 2019; Neubacher et al. 2020). On the
68 chromosome, the gene encoding this sRNA is located downstream of the *arcB* gene, in the reverse
69 direction (Fig. 1B). The 3' end of *arcZ* usually overlaps the 3' end of the *arcB* gene by ten bases and
70 ends at a Rho-independent terminator (Argaman et al. 2001) (Fig. 1B). *arcZ* does not overlap *arcB* in
71 *Pectobacterium carotovorum* (Wang et al. 2018). Due to its genomic location, the *sraH/ryhA* gene has
72 been renamed as *arcZ* (referring to arc-associated sRNA Z) (Papenfort et al. 2009). This small
73 regulatory RNA targets key regulators (*rpoS*, *flhD*, *lrp*) involved in various functions in the bacterial
74 cell, playing a pleiotropic role in diverse bacterial species (Papenfort et al. 2009). For instance, a
75 transcriptome analysis carried out in *E. amylovora* revealed that ArcZ regulates the expression of 10%
76 of its genome (Schachterle et al. 2019a). Recent RIL-seq data have shown that ArcZ may interact
77 directly with more than 10% of mRNAs in the *E. coli* and *Salmonella* genomes. ArcZ may have one of
78 the largest target regulons for sRNA (Melamed et al. 2016; Liu et al. 2023). However, the *arcZ* gene is
79 not essential in any of the bacteria where it has been studied, and mutated strains lacking this gene
80 survive in laboratory conditions. The purpose of this review is to compile all presently accessible data
81 on ArcZ in various bacterial species of the *Enterobacteriales* order. We will discuss the identified targets
82 of ArcZ and the molecular process through which ArcZ can regulate their expression.

83 **WHEN AND HOW IS A MATURE FORM OF ArcZ OBTAINED ?**

84 ***arcZ* is expressed during the stationary phase and in aerobic conditions.**

85 The sRNA ArcZ is regulated by the two-component system ArcAB. ArcB is a sensor kinase
86 that can transfer its phosphate to ArcA under anoxic conditions, thereby activating this transcriptional
87 regulator (Brown et al. 2022). Once activated, phosphorylated ArcA represses *arcZ* transcription by
88 binding to the *arcZ* promoter region. Additionally, *arcB* mRNA directly contributes to the repression
89 of ArcZ. It is speculated that *arcB* mRNA acts as an asRNA and destabilizes ArcZ through pairing with
90 it. Therefore, the ArcAB system represses the transcription of ArcZ in situations of limited oxygen
91 supply. ArcZ is expressed under high oxygen conditions (Mandin and Gottesman, 2010). Additionally,
92 ArcZ exhibits maximum levels during the stationary growth phase (Chen and Gottesman 2017). ArcZ
93 also directly represses *arcB* transcription in *E. coli*, providing a negative feedback loop that may affect
94 the function of the ArcA-ArcB regulon (Mandin and Gottesman 2010)

95 After transcription, ArcZ is recognized by the chaperone Hfq and is rapidly cleaved by RNase
96 E into a shorter 56-nucleotide form (Fig. 1A and 1C). This processed form corresponds to the 3' part of
97 ArcZ, highly conserved in *Enterobacteriales*, which is the active form that can enhance or repress
98 expression of its target mRNAs (Fig. 1A and 1C) (Chao et al. 2017).

99

100 **Hfq and RNase E, two key players in ArcZ functions**

101 Hfq binds broadly to mRNAs, sRNAs, and ribosomal RNAs (reviewed in Updegrove et al.
102 2016). It is the first protein chaperone discovered to bind with sRNAs (reviewed in Vogel and Luisi
103 2011). It is a hexameric protein abundant in many bacteria and crucial for the stability of mRNA and
104 sRNA expression. As an RNA chaperone, Hfq binds to UA-rich sequences of sRNAs and promotes
105 hybridization of sRNAs to their target mRNAs, which causes either negative or positive regulation of
106 gene expression. Thus, Hfq could be considered as a catalyst that stabilizes the sRNA and promotes
107 the meeting and binding between the sRNA and the mRNA it regulates. It is also possible that Hfq
108 assists sRNA-mRNA binding by modifying the structure of the mRNA upon its own binding

109 (Geissmann and Touati 2004). Like many other sRNAs, ArcZ relies on Hfq for binding to target mRNAs
110 (Fig. 1C) (Soper et al. 2010). Hfq can increase the rate of binding of ArcZ to its mRNA targets and the
111 stability of the ArcZ/mRNA complex (Soper et al. 2010).

112 Hfq possesses two faces – a proximal face that enables it to bind to sRNAs and a distal face
113 that facilitates its binding to mRNAs. The consensus binding site for Hfq at the mRNA level
114 corresponds to the motif (AAN)₄. This repeat is commonly located in the 5'UTR of target mRNAs, as
115 evidenced by studies carried out on *rpoS* or *mutS* in *E. coli*. It has been proposed that the binding of
116 Hfq alone to *mutS* mRNA creates a structure that inhibits translation, while ArcZ bound to Hfq binds
117 to a sequence near the RBS of *mutS*, preventing ribosome binding (Soper et al. 2010; Chen and
118 Gottesman 2017). Another known function of Hfq is to protect mRNAs and sRNAs from degradation
119 by RNase E. In the case of DsrA and RyhB sRNAs, Hfq binds to these sRNAs at the RNase E cleavage
120 site, preventing their degradation (Moll et al. 2003).

121 RNase E also plays a crucial role in the maturation of specific sRNAs, including MicL and ArcZ,
122 by processing them into functional forms (Guo et al. 2014; Chao et al. 2017). These sRNAs undergo
123 cleavage by RNase E, resulting in a shorter and stable form that corresponds to the 3' end (Updegrove
124 et al. 2019). *In vivo*, two forms of *E. coli* ArcZ have been identified: a full 121-nucleotide form and a
125 processed 56-nucleotide form, which corresponds to the 3' end of ArcZ. Only the short form of ArcZ
126 has the ability to interact with its mRNA targets, resulting in the activation or repression of their
127 translation and/or stability (Fig. 1C) (Chao et al. 2017). The short form was also detected by Northern
128 blotting in *S. typhimurium*, *D. dadantii*, *Photorhabdus* and *Xenorhabdus* sp. (Papenfert et al. 2009;
129 Yuan et al. 2019; Neubacher et al. 2020). The RNase E cleavage site of *E. coli* ArcZ consists of a
130 minimum consensus sequence of five nucleotides: R(G/A)N(G/A/U/C)W(A/U)UU with the cut occurring
131 between the nucleotide sequences RN and WUU (Fig. 1C). Hfq is required to obtain a unique cleavage.
132 In its absence, RNase E cleaves the ArcZ transcript into various fragments, failing to generate the
133 functional 56 nucleotide-long ArcZ (Chao et al. 2017). Another important factor for unique cleavage
134 of the full-length form of ArcZ by RNase E at the consensus site is the presence of the highly conserved

135 sequence CCCUGGUGUUGGCGCA immediately following the consensus cleavage site (Chao et al.
136 2017). Indeed, ArcZ has other potential consensus sites for cleavage by the RNase E, but only one site
137 is cleaved to yield the transcript of 56 nucleotides in *E. coli* (Fig. 1).

138 The most likely hypothesis to explain why RNase E cleavage of ArcZ is essential to produce a
139 functional sRNA is that the 5' region of ArcZ prevents the conserved 3' sequence from being free for
140 base pairing with mRNAs. Cleavage releases the 3' part of ArcZ, enabling its base pairing with its
141 mRNA targets. Therefore, the chaperone Hfq and RNase E are two essential players in the processing
142 of ArcZ into a shorter sRNA of 56 nucleotides, but also for the stabilization and pairing of this sRNA
143 with its targets (Chao et al. 2017).

144

145 **THE mRNA TARGETS OF ArcZ IN ENTEROBACTERIALES**

146

147 **ArcZ regulates the general stress response.**

148 The RpoS-mediated general stress response in *E. coli* has been reviewed extensively (Battesti
149 et al. 2011). The first well-defined target of ArcZ in *E. coli* is the *rpoS* mRNA, as two studies pinpointed
150 this regulation and detailed its molecular mechanism (Mandin and Gottesman 2010; Soper et al. 2010)
151 (Fig. 2). RpoS (σ^{38}) is an alternative sigma factor responsible for activating genes that enhance
152 resistance against various stresses, including the *gad* genes during acid stress and the *ots* genes during
153 cold stress, through its interaction with RNA polymerase. Directly or not, RpoS regulates about 500
154 genes in *E. coli*. During the exponential phase of growth, RpoS production is repressed at various
155 levels, while it is activated during stationary phase. At the post-transcriptional level, the regulation of
156 *rpoS* is complex: the sRNAs ArcZ, DsrA and RprA are involved in activating the translation of the *rpoS*
157 mRNA, while the sRNAs OxyS and CyaR repress it (Mandin and Gottesman 2010). In addition to the
158 aforementioned regulation of *arcZ*, the transcription of these various sRNAs is itself regulated. *dsrA*
159 expression is triggered during low temperature conditions, *rprA* expression is induced through the Rcs
160 phosphorelay, *oxyS* is induced by H₂O₂ oxidative stress through the OxyR regulator, and *cyaR* is

161 regulated by CRP (Colland et al. 2000). These sRNAs, except DsrA, are known to be strictly dependent
162 on the chaperone Hfq, without which they can no longer regulate their target (Soper et al. 2010).

163 During acid stress in *E. coli*, ArcZ, by promoting translation of *rpoS* mRNA, increases the
164 synthesis of GadX, which in turn activates transcription of the *gadE* gene. GadE is the primary
165 transcriptional activator of the *gadA* and *gadB* genes, which encode glutamate decarboxylases, as
166 well as *gadD*, which encodes a glutamate transporter (Fig. 2) (Castanié-Cornet et al. 2010; Bak et al.
167 2014). The harmful impact of acid stress is thus counteracted by the production of GABA via the
168 decarboxylation of glutamic acid (De Biase et al. 1999).

169 Another aspect of ArcZ's influence on *rpoS* mRNA operates through regulatory cascades,
170 specifically through the interaction between ArcZ and CyaR (Fig. 2). CyaR transcription is controlled
171 by the global transcriptional repressor CRP, the CpxAR two-component system, and the sigma factor
172 RpoE (De Lay and Gottesman 2012). CyaR has the ability to interact with *rpoS* mRNA and to down-
173 regulate *rpoS* mRNA expression (Kim and Lee 2020) (Fig. 2). ArcZ has the ability to directly interact
174 with CyaR and inhibit its function by causing its degradation by the RNase E (Kim and Lee 2020).
175 Conversely, CyaR has no effect on the activity of ArcZ (Iosub et al. 2020). Moreover, since the
176 interaction region between ArcZ and CyaR corresponds to their *rpoS* mRNA binding sites, there may
177 be competition between these two sRNAs for interacting with the *rpoS* mRNA (Kim and Lee 2020).
178 CyaR also represses the translation of *nadE*, which encodes an enzyme involved in NAD⁺ biosynthesis
179 (Hughes et al. 1988). Under anaerobic conditions, ArcZ, which is absent, is unable to repress the action
180 of CyaR. Thus, CyaR reduces the concentration of NAD⁺ by repressing *nadE* translation. Conversely,
181 in aerobiosis, ArcZ prevents CyaR action, providing greater NAD⁺ availability (Fig. 2). These ArcZ-
182 CyaR interactions could thus enhance the regulation of *rpoS* expression, allowing *E. coli* to respond
183 more effectively to various stresses (Kim and Lee 2020).

184 In *E. coli*, ArcZ directly represses *mutS* mRNA translation (Fig. 2). Bacteria under stress
185 accumulate mutations to better survive and adapt (reviewed in Foster 2007). The DNA mismatch
186 repair (MMR) system limits the occurrence of mutations (Wyrzykowski and Volkert 2003). The MutS

187 protein, a crucial component of the MMR system, identifies mispaired bases in DNA and triggers the
188 repair process via the MMR system (Su and Modrich 1986). The repression of *mutS* by ArcZ occurs
189 during the stationary growth phase when ArcZ is abundant. Due to this repression, the MMR system
190 is no longer active and an increase in unrepaired mutations is observed (Chen and Gottesman 2017).
191 Furthermore, the sRNA SdsR directly represses *mutS* (Chen and Gottesman 2017). SdsR is transcribed
192 by the RNA polymerase only when this latter is associated with RpoS. Since ArcZ is able to activate
193 the translation of *rpoS* mRNA in *E. coli*, ArcZ represses *mutS* both directly by acting as a post-
194 transcriptional repressor on *mutS* mRNA and indirectly through SdsR. Repression of *mutS* by ArcZ
195 contributes to stress-induced mutagenesis in *E. coli* (Fig. 2).

196 *Erwinia amylovora* is a phytopathogenic bacterium that causes fire blight and has a wide range
197 of host species within the *Rosaceae* family (e.g., apple, pear, raspberry) (reviewed in Piqué et al. 2015).
198 ArcZ is required for the full virulence of *E. amylovora* (Zeng et al. 2013). In *E. amylovora*, ArcZ
199 modulates the levels of the catalase KatA and of the thiol peroxidase Tpx (Fig. 3). These two enzymes
200 are essential for *E. amylovora* to fight against the free radicals produced by the plant during infection
201 such as hydrogen peroxide produced during apple infection (Santander et al. 2018). ArcZ indirectly
202 regulates *katA* at the transcriptional level through the ArcA regulator. By increasing the translation of
203 ArcA, which is a transcriptional activator of *katA*, ArcZ enhances the response to oxidative stress (Fig.
204 3) (Schachterle et al. 2019a). Previously, it was demonstrated that the ArcAB system in *E. coli*
205 represses the expression of *arcZ* (Fig. 3). If such a regulation exists in *E. amylovora*, the three-way
206 interaction of ArcZ-ArcA-KatA should create a positive feedback loop, increasing the amount of
207 KatA. Conversely, ArcZ directly controls *tpx* at the post-transcriptional level by binding to its mRNA,
208 resulting in reduced level of the thiol peroxidase Tpx (Fig. 3). Thus, through these two regulatory
209 modes, ArcZ finely tunes the cellular response to oxidative stress based on oxygen availability and
210 oxidative status (Schachterle et al. 2019a). Post-transcriptional repression of *tpx* was also described
211 in *S. typhimurium* (Papenfort et al. 2009).

212

213 **ArcZ targets mRNAs involved in virulence.**

214 ArcZ is also able to regulate genes important for bacterial virulence, such as motility, biofilm
215 formation and secretion of antimicrobial compounds.

216 *S. typhimurium* is a bacterium that is commonly associated with food poisoning. The
217 bacterium's ability to cause disease is primarily attributed to the secretion of virulence factors through
218 a type III secretion system encoded by a pathogenicity island called SPI1. This T₃SS enables the
219 secretion of effectors that facilitate the internalization of bacterial cells into host cells. The *hilA* gene
220 is crucial for the synthesis of the *S. typhimurium* T₃SS as it activates the T₃SS structural genes present
221 in SPI1. *hilA* transcription is regulated by HilC, HilD and RtsA (Ellermeier et al. 2005). The sRNAs ArcZ
222 and FnrS indirectly repress *hilA* expression by repressing the translation of *hilD* mRNA through direct
223 interaction, but this process relies on the oxygen levels and can be antagonistic. In aerobic conditions,
224 ArcZ is expressed and represses the translation of *hilD*. On the other hand, the two component system
225 Fnr activates the transcription of FnrS in anoxia, which in turn represses the translation of *hilD* (Fig. 4)
226 (Durand and Storz 2010; Kim et al. 2018). This regulatory network enables the most efficient
227 expression of T₃SS genes when exposed to fluctuating oxygen levels (Kim et al. 2018).

228 Interestingly, ArcZ was the first sRNA found to bind to the mRNA of a horizontally acquired
229 virulence gene. This gene, known as STM3216, is specific to *S. typhimurium*, and it is predicted to
230 function as a receptor involved in chemotaxis. ArcZ directly inhibits the translation of STM3216 mRNA
231 (Fig. 4). This discovery highlights the potential of sRNAs to regulate the expression of horizontally
232 acquired genes (Papenfort et al. 2009). ArcZ also controls biofilm formation in *S. typhimurium*,
233 through the transcriptional regulator CsgD. CsgD regulates several genes responsible for curli's
234 assembly, transport, and structural component synthesis. These components are important for
235 biofilm formation (Hammar et al. 1995; Mika and Hengge 2014). CsgD is also the main regulator of
236 the expression of the *rdar* morphotype, which relates to multicellular behavior characterized by the
237 production of adhesive extracellular matrix and curli expression (reviewed in Römling 2005). CsgD is
238 regulated by RpoS, but ArcZ has been shown to partially regulate *csgD* independently of RpoS

239 (Monteiro et al. 2012). Furthermore, ArcZ appears to regulate attachment to surfaces for biofilm
240 formation by repressing the synthesis of type 1 fimbriae (Monteiro et al. 2012).

241 Escape from the host immune system is also a crucial aspect for successful pathogen
242 infection. The bacterial surface lipopolysaccharides (LPS) are perceived by the host immune system
243 as a foreign element. Consequently, bacteria synthesize enzymes that modify LPS to evade
244 recognition. In *E. coli*, *eptB* encodes an LPS-modifying enzyme. The synthesis of EptB is under the
245 control of the sigma factor RpoE. *eptB* mRNA is directly repressed by ArcZ in *E. coli* (Fig. 2) (Moon et
246 al. 2013).

247 Motility is an essential virulence factor for flagellated bacteria such as *E. coli*, *E. amylovora*, *D.*
248 *dadantii* and many others. In *E. coli*, the major regulator of motility is FlhDC, a class I transcriptional
249 activator of the flagellar regulatory cascade (Zhao et al. 2007). In *E. coli*, ArcZ binds directly to *flhD*
250 mRNA and inhibits its translation, along with that of *flhC*, since they are included in the same operon.
251 Consequently, the motility of *E. coli* is reduced when ArcZ is expressed (De Lay and Gottesman 2012)
252 (Fig. 2).

253 A mechanistically identical repression of *flhDC* by ArcZ is also observed in *E. amylovora* (Fig.
254 3). Interestingly, in contrast to results in *E. coli*, this post-transcriptional repression of *flhDC* by ArcZ
255 actually enhances the motility of *E. amylovora*. This difference could be explained by an indirect
256 activation of *flhDC* by ArcZ in *E. amylovora*, not present in *E. coli* (Schachterle and Sundin 2019;
257 Schachterle et al. 2019b). LrhA is a direct transcriptional repressor of *flhDC*. In *E. coli*, the Leucine-
258 responsive Regulatory Protein (Lrp) binds to the *lrhA* promoter, which leads to its transcriptional
259 activation (Kroner et al. 2019). In *E. amylovora*, ArcZ directly interacts with *lrp* mRNA to post-
260 transcriptionally repress it. Thus, in *E. amylovora*, ArcZ directly represses *flhDC* directly while
261 indirectly promoting *flhDC* transcription through the *lrp-lrhA* pair. Therefore, the ArcZ-Lrp-FlhDC trio
262 operates as a 'feed-forward' type of regulatory loop (Schachterle and Sundin 2019). This regulatory
263 loop may accelerate the response through faster FlhDC production when ArcZ levels vary. This effect
264 has been previously demonstrated for the "feed-forward" regulatory loop Fur-SodA-RyhB (Semsey

265 2014). It is worth noting that FlhDC activates levan production and that Lrp represses the production
266 of the exopolysaccharide amylovoran, both of which are key compounds essential for biofilm
267 formation in *E. amylovora*. Thus, ArcZ not only regulates motility through Lrp and FlhDC, but also
268 influences biofilm formation in *E. amylovora* (Fig. 3) (Schachterle and Sundin 2019).

269 In the *Pectobacteriaceae* family of plant pathogenic bacteria, one of the key regulators of
270 virulence factor synthesis is named PecT in *D. dadantii* (Condemine et al. 1999) or HexA in
271 *Pectobacterium carotovorum* (Mukherjee et al., 2000). PecT represses the expression of the sRNA
272 RsmB, a Hfq-independent small regulatory RNA. The *rsmB* RNA binds to RsmA, an RNA binding
273 protein, preventing it from repressing the expression of its target genes, including those encoding the
274 main virulence factors, the T₃SS and pectinases, which are Plant Cell Wall-Degrading Enzymes
275 (PCWDE) (Hyytiäinen et al. 2001; Yang et al. 2008). It has been demonstrated that ArcZ directly
276 represses the translation of *pecT*. This repression results in an increase in the amount of *rsmB* available
277 to titrate RsmA, leading to increased production of the T₃SS and of the PCWDE (Yuan et al. 2019)
278 (Fig. 5). Consequently, an *arcZ* mutant in *D. dadantii* has a drastically reduced virulence (Yuan et al.
279 2019). A mutant of *P. carotovorum* lacking *arcZ* was also found to be less virulent. However, the targets
280 of ArcZ in this bacterium were not investigated (Wang et al. 2018).

281 Similarly, ArcZ induces the expression of the secondary metabolite clusters *sol* and *zms* in
282 *Dickeya solani*, resulting in the production of the antimicrobial molecules solanimycin and zeamine,
283 respectively. Nevertheless, the precise targets of ArcZ in *D. solani* are still unknown, although it is
284 suspected that *pecT* may be involved, as its target site in the 5'UTR of *pecT* is conserved (Brual et al.
285 2023).

286 Secondary metabolites play a significant role in the mutualistic associations of nematodes
287 with *Xenorhabdus* and *Photorhabdus* (Tobias et al. 2017b). HexA, known for its ability to repress
288 secondary metabolite production in *Photorhabdus* (Tobias et al. 2017a), is repressed by ArcZ. This is
289 achieved through direct binding to *hexA* 5'UTR mRNA in *Photorhabdus* and *Xenorhabdus*.

290 Consequently, secondary metabolites that modulate the nematode immune response (Fig. 5) are
291 produced more abundantly when ArcZ represses *hexA* translation (Fig. 5) (Neubacher et al. 2020).

292

293 **ArcZ and the bacterial metabolism.**

294 A few genes targeted by ArcZ that play a role in nutrient metabolism or transport have been
295 identified. In *E. coli*, the *ppsA* gene encodes a phosphoenolpyruvate synthetase that is required for the
296 conversion of pyruvate to phosphopyruvate, which initiates the process of gluconeogenesis. ArcZ
297 directly or indirectly regulates *ppsA* in a positive manner. Hence, ArcZ regulates gluconeogenesis
298 initiation by activating *ppsA* transcription or translation when *E. coli* is grown in the presence of
299 pyruvate (Rachwalski et al. 2022).

300 Another potential target involved in specific serine transport in *S. typhimurium* is *sdaC*, which
301 is directly repressed by ArcZ at the mRNA level. In this case, *sdaC* is co-transcribed with *sdaB*, a serine
302 deaminase, suggesting that SdaCB is involved in serine catabolism (Fig. 4). Thus, ArcZ may play a
303 potential role in repressing serine catabolism (Papenfort et al. 2009).

304

305 **MOLECULAR MECHANISMS USED BY ArcZ TO REGULATE TRANSLATION**

306 As previously shown, ArcZ possesses the ability to inhibit or activate its target mRNAs. The
307 processed 56 nt transcript that binds to mRNAs is highly conserved (Fig. 1A). This transcript is capable
308 of regulating targets in diverse ways (Fig. 6). The various targets can be categorized into five groups
309 based on the ArcZ pairing location in the target mRNA: 1) ArcZ can pair at the RBS or in close
310 proximity to it, such as with the *mutS* mRNA ; 2) it can pair in the coding region of a gene such as with
311 the *tpx* mRNA ; 3) it can pair in the 5'UTR region of the mRNA approximately 50 nucleotides from the
312 AUG, as demonstrated with *pecT* or *flhDC* ; 4) it can prevent Rho-mediated premature termination ;
313 5) it can pair directly with an sRNA like CyaR.

314

315 **Binding of ArcZ near the RBS**

316 In the case of *mutS*, ArcZ pairs near to the ribosome binding site (RBS) to repress translation
317 of *mutS*. In fact, after being made mature and stable by Hfq and RNase E, ArcZ will pair at the level of
318 the 5'UTR of *mutS* between positions -25 to -15 relative to the ATG, upstream of the RBS of *mutS* that
319 is located from -11 to -5. This pairing is proposed to alter the secondary structure of the *mutS*
320 translation initiation region, making it unavailable to ribosomes (Fig. 6) (Chen and Gottesman 2017).

321 ArcZ can bind directly to the RBS region of STM₃₂₁₆ mRNA between position -25 to -5. A
322 similar phenomenon is observed with the *sdaC* mRNA, where ArcZ hybridizes directly with the RBS
323 of the region from -13 to -2. This type of pairing effectively masks the RBS, rendering it inaccessible
324 for translation initiation (Papenfort et al. 2009).

325

326 **Binding of ArcZ in the coding region of the target mRNA**

327 ArcZ can bind to the coding region of the *tpx* gene mRNA, specifically between positions +10
328 to +26. Inhibition may occur through the inhibition of 30S ribosome subunit association, as well as
329 through the degradation of the mRNA via RNase recruitment (Papenfort et al. 2009) (Fig. 6). A similar
330 mechanism is observed with the binding of the MicC sRNA to the coding region of the *ompD* mRNA,
331 resulting in the repression of *ompD* translation due to accelerated degradation by RNase E (Pfeiffer
332 et al. 2009).

333

334 **Binding of ArcZ upstream of the RBS**

335 In *Photothabdus*, ArcZ binds to a 9-nucleotide region located on the 5'UTR of the *hexA* mRNA
336 between positions -50 to -42 (Neubacher et al. 2020). This results in inhibition of *hexA*'s translation.
337 The region of binding is distal to the RBS, making it improbable for ArcZ to prevent the RBS's
338 recognition in a way comparable to that observed for *mutS*. Hence, the inhibition mechanism of ArcZ
339 is different from that of *mutS*. The regions where ArcZ pairs in the 5' UTR of *hexA* are rich in C and A
340 nucleotide bases. This type of sequence in the 5'UTRs of mRNAs can enhance translation (Sharma et
341 al. 2007). Previous research suggests that the GcvB sRNA can regulate several ABC transporter

342 mRNAs using this mechanism (Sharma et al. 2007). The hypothesis is that ArcZ, by binding to these
343 C/A-rich regions, covers up these translation enhancers and thus decreases the translation of *hexA*
344 mRNA (Neubacher et al. 2020).

345 To repress *pecT* in *D. dadantii*, ArcZ pairs at a region located between positions -84 to -58 from
346 the start codon, which is considerably far from the RBS. Additionally, the region involved in the pairing
347 is not rich in C/A. Therefore, this repression could be due to a modification of the secondary structure
348 of *pecT* mRNA leading to the formation of a hairpin structure that masks ribosome access to the RBS
349 (Fig. 6) (Yuan et al. 2019).

350

351 **Dual Site Binding of ArcZ in the 5'UTR *flhDC* mRNA in *E. coli*.**

352 In *E. coli*, the interaction of ArcZ with the 5'UTR region of *flhDC* is more complex than the previously
353 cited cases. Actually, the same region of ArcZ pairs with two sites in the 5'UTR of *flhDC*: one located
354 between nucleotides -23 to -7, in close proximity to the RBS of *flhD*, and the other situated between
355 positions -63 to -47. The farthest region from the RBS corresponds to a region exhibiting high C/A
356 levels. An hypothesis proposes that ArcZ, by pairing, removes a hairpin structure predicted for the
357 *flhDC* mRNA in this region, changing the steric constraints for base stabilization at the RBS level (Fig.
358 6). Nevertheless, it appears that the pairing of ArcZ at the location farthest from the RBS is the most
359 potent in suppressing *flhDC* translation (De Lay and Gottesman 2012; Schachterle et al. 2019b). It is
360 noteworthy that the sRNA McaS activates the translation of *flhD* (Thomason et al. 2012), an effect
361 that is opposite to that of ArcZ. However, McaS and ArcZ share a common binding site between
362 positions -61 to -52 on the *flhD* 5'UTR (De Lay and Gottesman 2012). McaS, like ArcZ, has a second
363 binding site on the 5'UTR of *flhD*, but further upstream of the start codon between positions -86 to -
364 77 on the *flhD* 5'UTR. It was proposed that McaS may initially bind to an overlapping ArcZ site on *flhDC*
365 mRNA, aiding a second McaS site to pair and expose the mRNA for ribosome entry (De Lay and
366 Gottesman 2012; Thomason et al. 2012). This suggests that despite targeting the same area, two
367 sRNAs can regulate differently the same target. Ultimately, ArcZ and McaS likely compete for the

368 same *flhDC* mRNA site, affecting cell motility based on which sRNA prevails (De Lay and Gottesman
369 2012).

370

371 **ArcZ prevents Rho-mediated premature termination and liberates the RBS for *rpoS* translation**

372 ArcZ linked to Hfq binds to the *rpoS* mRNA at the region -91 to -63 relative to the translation
373 start. However, only 23 of the 28 bases in ArcZ match those in *rpoS* mRNA, so this binding is not
374 perfect. In the absence of ArcZ, the RBS of *rpoS* is locked in a hairpin loop, making the RBS
375 inaccessible to the ribosome and preventing translation initiation (Mandin and Gottesman 2010). By
376 pairing in this region of the 5'UTR of *rpoS*, ArcZ unfolds the hairpin loop and frees up the RBS, allowing
377 for translation initiation (Fig. 6). Additionally, binding of DsrA, RprA and ArcZ appears to stabilize *rpoS*
378 mRNA in *E. coli* and protects it from degradation by RNase E (McCullen et al. 2010). In the absence of
379 ArcZ, the Rho factor causes premature termination of *rpoS* transcription by binding to the 5'UTR of
380 *rpoS*. When ArcZ binds to the 5'UTR of *rpoS* mRNA, it prevents interaction between the Rho factor
381 and this region. Therefore, ArcZ functions as an antiterminator for transcription and thereby activates
382 the translation of *rpoS* (Fig. 6) (Sedlyarova et al. 2016).

383

384 **Binding of ArcZ directly to another sRNA**

385 In *E. coli*, CyaR is a CRP-regulated RNA whose pairing to a region close to the ArcZ binding
386 site in the 5'UTR of *rpoS* mRNA leads to its degradation (Kim and Lee 2020). RIL-seq data revealed
387 the presence of ArcZ-CyaR interactions (Melamed et al. 2016). The direct interaction between CyaR
388 and ArcZ results in the degradation of CyaR by the RNase E, thereby alleviating the CyaR-mediated
389 repression of *rpoS* and maximizing the activation of *rpoS* by ArcZ (Kim and Lee 2020).

390

391 **PERSPECTIVES**

392 ArcZ is a major regulator of gene expression in *Enterobacteriales*, as it regulates 10-15% of the
393 genome in various species. The RIL-Seq technique, which allows the identification of sRNA-mRNA

394 target pairs pulled down with the Hfq protein, has been employed to identify more than 300 putative
395 new ArcZ mRNA targets in *E. coli* and *S. enterica* (Melamed et al. 2016; Liu et al. 2023). Of note, 330 of
396 the 335 base-pairing interactions identified involve the seed domain at the 5' end of the processed
397 form of ArcZ (Figure 1A) (Melamed et al. 2016). Known ArcZ targets were identified (*rpoS*, *flhD*, *sdaC*
398 and *tpx*). Although further validation is required for the majority of these targets, these
399 studies provide compelling evidence that ArcZ plays a central role in regulating a wide
400 variety of targets.

401 Determining ArcZ's targets in a new bacterial model remains challenging *in silico*. Prediction
402 softwares of sRNA-mRNA interaction such as CopraRNA or IntaRNA (Wright et al. 2014) may fail to
403 identify complex interaction targets, such as the interaction zones between ArcZ and *flhD*. Therefore,
404 there are still numerous targets and regulatory mechanisms that are yet to be uncovered.
405 Additionally, an interesting hypothesis has been proposed by Papenfort et al. (Papenfort et al. 2009).
406 These authors observed that ArcZ is an sRNA conserved in *Enterobacterales* and that its inhibitory or
407 activating action on mRNA translation is in many cases achieved by pairing at the site of the SD
408 sequence, the most conserved element of bacterial mRNAs. They suggested that when ArcZ is in high
409 intracellular concentration and once ArcZ has paired with all of its primary targets, excess ArcZ
410 molecules will be able to pair with secondary targets at the RBS. This type of regulation has been
411 observed in the post-transcriptional regulatory machinery at the mammalian level (Saxena et al.
412 2003; Papenfort et al. 2009). If confirmed, this hypothesis will likely expand the range of ArcZ-
413 controlled genes beyond expectation.

414 A remarkable observation concerning ArcZ is the remarkable degree of conservation of the
415 seed region located after the cleavage site and, in mirror image, the high degree of conservation of
416 the pairing region with the 5' UTR of *rpoS*. Indeed, an alignment of the *rpoS* 5' UTR in *Enterobacterales*
417 (Fig. 7A) showed that the region of the 5' UTR that pairs with ArcZ is highly conserved. Given that
418 RpoS is a major regulator of the general stress response, it can be assumed that regulation of the
419 stress response by ArcZ via RpoS is also conserved in most *Enterobacterales*. However, some

420 nucleotide changes, of the order of 1 to 2 bases, exist in the seed region of ArcZ (Fig. 1A) as in *D.*
421 *dadantii*, *P. carotovorum*, *E. amylovora* and *E. tarda*. It can be observed that compensatory nucleotide
422 changes for these mutations are not systematically found in the 5'UTR of *rpoS* (Fig. 7A). This suggests
423 that the ArcZ-*rpoS* pairing may be less stable in these bacteria. It would be of interest in the future to
424 conduct a systematic analysis to verify the interactions between ArcZ and *rpoS* 5'UTR in bacteria
425 containing nucleotide modifications in either *arcZ* or *rpoS*. In some cases, the interaction may no
426 longer occur. In addition, the conservation of pairing between the seed region of ArcZ and other
427 targets is less clear. For instance, an alignment of the 5' untranslated region of *flhD* in different
428 *Enterobacterales* species (Fig. 7B) revealed that this pairing is less conserved than the ArcZ/*rpoS*
429 pairing. It is therefore possible that ArcZ does not regulate motility via the control of *flhD* in some of
430 these bacteria, or does not regulate motility at all. Nevertheless, in *E. amylovora*, although the 5'UTR
431 of *flhD* is different from that of *E. coli*, an interaction could be detected (Schachterle et al. 2019b) (Fig.
432 7B). This putative pairing zone is not conserved in *D. dadantii*, whereas ArcZ deletion results in a
433 difference in motility of the mutant compared to a wild-type strain. It was proposed that regulation
434 of motility in *D. dadantii* by ArcZ occurs through an as yet unknown mechanism (Yuan et al. 2019). A
435 significant amount of work remains to be done before a clear picture can be drawn of the manner in
436 which ArcZ fulfills its regulatory role in each of the major bacterial genera of *Enterobacterales*.

437 A single-point mutation was found in the 3' region of ArcZ in the *D. solani* type strain IPO2222
438 (Brual et al. 2023). This G-to-A mutation is positioned 17 bases downstream of the RNase E cleavage
439 site. Despite being distant from the cleavage site, it prevents processing of the full-length ArcZ_{IPO2222}
440 into a short, stable form. Consequently, ArcZ_{IPO2222} is not functional. The predicted secondary structure
441 of the full-size form of ArcZ_{IPO2222} is altered in comparison with non-mutated ArcZ from other *D. solani*
442 strains where ArcZ is active. This indicates the potential of some tertiary structures in small RNAs to
443 prevent cleavage by RNase E. This is supported by the observation that when point mutations are
444 introduced experimentally in ArcZ and alter its secondary structure, ArcZ variants may not be
445 processed by the RNase E (Yuan et al. 2019). Notably, the ArcZ loss-of-function mutation found in *D.*

446 *solani* IPO2222 is shared by other *Dickeya* species, including *D. fangzhongdai* and *D. parazeae* (Brual
447 et al. 2023). Expanding the search for this mutation to other species reveals its presence in strains of
448 *Citrobacter youngae*, *S. typhimurium*, *S. enterica*, *Proteus mirabilis*, *Yersinia pestis*, *Y.*
449 *pseudotuberculosis*, *Y. ruckeri*. Another G-to-T mutation has been found at the same location in other
450 strains (Fig. 8). It would be worthwhile to compile a comprehensive list of *arcZ* alleles in
451 *Enterobacterales* and investigate their systematic processing by the RNase E.

452 Why do certain isolated strains exhibit *arcZ* mutations that result in functional loss? It cannot
453 be ruled out at present that the isolation of these bacteria on a nutrient-rich laboratory medium has
454 resulted in the selection of a mutated strain in *arcZ*. Bacteria cultivated in laboratory conditions are
455 known to accumulate mutations (Nahku et al. 2011; Spira et al. 2011; Chandler et al. 2019; Artuso et
456 al. 2022; Jacques et al. 2023), some of these mutations, particularly those in *arcA* and *rpoS*, which are
457 two prime targets of ArcZ, contribute to enhancing the catabolism of amino acids that are abundant
458 in rich environments (Saxer et al. 2014). Nevertheless, mutations in *arcZ* have not been isolated during
459 experimental evolution assays *in vitro*. Another possible hypothesis is that mutations in *arcZ* can lead
460 to the emergence of cheaters. Cheaters are individuals who do not cooperate with others in the
461 population, but still benefit from the public goods generated by cooperators without contributing to
462 costs of producing those goods (Smith and Schuster 2019). Since ArcZ regulates a variety of genes
463 associated with flagellum apparatus and secretion systems machineries in various bacteria, it is
464 plausible that bacterial cells with *arcZ* defect can benefit from wild-type cooperators' secretion in a
465 host organism. The secretion of virulence factors can be costly for bacteria. For example, *S.*
466 *typhimurium* mutants deficient in T₃SS cannot cause independent mouse infections, but they
467 outcompete isogenic wild-type bacteria during coinfections (Diard et al. 2013). In co-infection
468 experiments within a host, it would be interesting to investigate whether *arcZ*-deficient mutants
469 could outcompete the wild-type strain.

470
471 **FUNDING**

472 This work was supported by a grant from Agence Nationale de la Recherche to G.D. (Project
473 Tn-Phyto, ANR- 19-CE35-0016) and E.G (Project DeCaP, ANR-22-CE18-0004-01). E.G was
474 also supported by the FR BioEEnVis, annual credits from the University Lyon I and the CNRS
475 at regular basis. Q.D. was supported by a PhD grant from INSA de Lyon.

476

477

478 REFERENCES

479 Argaman L, Hershberg R, Vogel J, Bejerano G, Wagner EG, Margalit H, Altuvia S. 2001.
480 Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. *Curr*
481 *Biol* 11: 941–950. doi: 10.1016/s0960-9822(01)00270-6.

482 Artuso I, Lucidi M, Visaggio D, Capecchi G, Lugli GA, Ventura M, Visca P. 2022. Genome
483 diversity of domesticated *Acinetobacter baumannii* ATCC 19606T strains. *Microb*
484 *Genom* 8: 000749. doi: 10.1099/mgen.0.000749.

485 Bak G, Han K, Kim D, Lee Y. 2014. Roles of *rpoS*- activating small RNA s in pathways leading
486 to acid resistance of *Escherichia coli*. *MicrobiologyOpen* 3: 15–28. doi:
487 10.1002/mbo3.143

488 Battesti A, Majdalani N, Gottesman S. 2011. The RpoS-mediated general stress response in
489 *Escherichia coli*. *Annu Rev Microbiol* 65: 189–213. doi: 10.1146/annurev-micro-
490 090110-102946.

491 Brown AN, Anderson MT, Bachman MA, Mobley HLT. 2022. The ArcAB two-component
492 system: Function in metabolism, redox control, and infection. *Microbiol Mol Biol Rev*
493 86: e00110-21. doi: 10.1128/membr.00110-21.

494 Brual T, Effantin G, Baltenneck J, Attaiech L, Grosbois C, Royer M, Cigna J, Faure D,
495 Hugouvieux-Cotte-Pattat N, Gueguen E. 2023. A natural single nucleotide mutation in
496 the small regulatory RNA ArcZ of *Dickeya solani* switches off the antimicrobial
497 activities against yeast and bacteria. *PLoS Genetics* 19: e1010725. doi:
498 10.1371/journal.pgen.1010725.

499 Castanié-Cornet MP, Cam K, Bastiat B, Cros A, Bordes P, Gutierrez C. 2010. Acid stress
500 response in *Escherichia coli*: mechanism of regulation of *gadA* transcription by RcsB
501 and GadE. *Nucleic Acids Res* 38: 3546–3554. doi: 10.1093/nar/gkq097.

502 Chandler CE, Horspool AM, Hill PJ, Wozniak DJ, Schertzer JW, Rasko DA, Ernst RK. 2019.
503 Genomic and phenotypic diversity among ten laboratory isolates of *Pseudomonas*
504 *aeruginosa* PAO1. *J Bacteriol* 201: e00595-18. doi: 10.1128/JB.00595-18

505 Chao Y, Li L, Girodat D, Förstner KU, Said N, Corcoran C, Smiga M, Papenfort K, Reinhardt

- 506 R, Wieden HJ, et al. 2017. *In vivo* cleavage map illuminates the central role of RNase E
507 in coding and non-coding RNA pathways. *Mol Cell* 65: 39–51. doi:
508 10.1016/j.molcel.2016.11.002.
- 509 Chen J, Gottesman S. 2017. Hfq links translation repression to stress-induced mutagenesis in
510 *E. coli*. *Genes Dev* 31: 1382–1395. doi: 10.1101/gad.302547.117.
- 511 Clamp M, Cuff J, Searle SM, Barton GJ. 2004. The Jalview Java alignment editor.
512 *Bioinformatics* 20: 426–427. doi: 10.1093/bioinformatics/btg430.
- 513 Colland F, Barth M, Hengge-Aronis R, Kolb A. 2000. σ factor selectivity of *Escherichia coli*
514 RNA polymerase: role for CRP, IHF and Lrp transcription factors. *The EMBO Journal*
515 19: 3028–3037. doi: 10.1093/emboj/19.12.3028.
- 516 Condemine G, Castillo A, Passeri F, Enard C. 1999. The PecT repressor coregulates synthesis
517 of exopolysaccharides and virulence factors in *Erwinia chrysanthemi*. *Mol Plant*
518 *Microbe Interact* 12: 45–52. doi: 10.1094/MPMI.1999.12.1.45.
- 519 Dadzie I, Xu S, Ni B, Zhang X, Zhang H, Sheng X, Xu H, Huang X. 2013. Identification and
520 characterization of a cis-encoded antisense RNA associated with the replication process
521 of *Salmonella enterica* Serovar *Typhi*. *PLoS ONE* 8: e61308. doi:
522 10.1371/journal.pone.0061308.
- 523 De Biase D, Tramonti A, Bossa F, Visca P. 1999. The response to stationary-phase stress
524 conditions in *Escherichia coli*: role and regulation of the glutamic acid decarboxylase
525 system. *Mol Microbiol* 32: 1198–1211. doi: 10.1046/j.1365-2958.1999.01430.x.
- 526 De Lay N, Gottesman S. 2012. A complex network of small non-coding RNAs regulate motility
527 in *Escherichia coli*. *Mol Microbiol* 86: 524–538. doi: 10.1111/j.1365-
528 2958.2012.08209.x.
- 529 Diard M, Garcia V, Maier L, Remus-Emsermann MNP, Regoes RR, Ackermann M, Hardt WD.
530 2013. Stabilization of cooperative virulence by the expression of an avirulent
531 phenotype. *Nature* 494: 353–356. doi: 10.1038/nature11913.
- 532 Durand S, Storz G. 2010. Reprogramming of anaerobic metabolism by the FnrS small RNA.
533 *Mol Microbiol* 75: 1215–1231. doi: 10.1111/j.1365-2958.2010.07044.x.
- 534 Dutta T, Srivastava S. 2018. Small RNA-mediated regulation in bacteria: A growing palette of
535 diverse mechanisms. *Gene* 656: 60–72. doi: 10.1016/j.gene.2018.02.068.
- 536 Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and
537 space complexity. *BMC Bioinformatics* 5: 113. doi: 10.1186/1471-2105-5-113.
- 538 Eichner H, Karlsson J, Loh E. 2022. The emerging role of bacterial regulatory RNAs in disease.
539 *Trends Microbiol* 30: 959–972. doi: 10.1016/j.tim.2022.03.007.

- 540 Ellermeier CD, Ellermeier JR, Slauch JM. 2005. HilD, HilC and RtsA constitute a feed forward
541 loop that controls expression of the SPI1 type three secretion system regulator *hila* in
542 *Salmonella enterica* serovar *Typhimurium*. *Mol Microbiol* 57: 691–705. doi:
543 10.1111/j.1365-2958.2005.04737.x
- 544 Foster PL. 2007. Stress-induced mutagenesis in bacteria. *Crit Rev Biochem Mol Biol* 42: 373–
545 397. doi: 10.1080/10409230701648494.
- 546 Geissmann TA, Touati D. 2004. Hfq, a new chaperoning role: binding to messenger RNA
547 determines access for small RNA regulator. *EMBO J* 23: 396–405. doi:
548 10.1038/sj.emboj.7600058.
- 549 Guo MS, Updegrove TB, Gogol EB, Shabalina SA, Gross CA, Storz G. 2014. MicL, a new σ^E -
550 dependent sRNA, combats envelope stress by repressing synthesis of Lpp, the major
551 outer membrane lipoprotein. *Genes Dev* 28: 1620–1634. doi: 10.1101/gad.243485.114.
- 552 Hammar M, Arnqvist A, Bian Z, Olsén A, Normark S. 1995. Expression of two *csg* operons is
553 required for production of fibronectin- and congo red-binding curli polymers in
554 *Escherichia coli* K-12. *Mol Microbiol* 18: 661–670. doi: 10.1111/j.1365-
555 2958.1995.mmi_18040661.x.
- 556 Hughes KT, Olivera BM, Roth JR. 1988. Structural gene for NAD synthetase in *Salmonella*
557 *typhimurium*. *J Bacteriol* 170: 2113–2120. doi: 10.1128/jb.170.5.2113-2120.1988
- 558 Hyytiäinen H, Montesano M, Palva ET. 2001. Global regulators ExpA (GacA) and KdgR
559 modulate extracellular enzyme gene expression through the RsmA-rsmB system in
560 *Erwinia carotovora* subsp. *carotovora*. *Mol Plant Microbe Interact* 14: 931–938. doi:
561 10.1094/MPMI.2001.14.8.931.
- 562 Iosub IA, van Nues RW, McKellar SW, Nieken KJ, Marchioretto M, Sy B, Tree JJ, Viero G,
563 Granneman S. 2020. Hfq CLASH uncovers sRNA-target interaction networks linked to
564 nutrient availability adaptation. *eLife* 9: e54655. doi: 10.7554/eLife.54655.
- 565 Jacques LC, Green AE, Barton TE, Baltazar M, Aleksandrowicz J, Xu R, Trochu E, Kadioglu
566 A, Neill DR. 2023. Influence of *Streptococcus pneumoniae* within-strain population
567 diversity on virulence and pathogenesis. *Microbiol Spectr* 11: e0310322. doi:
568 10.1128/spectrum.03103-22.
- 569 Kim D, Hong JS-J, Qiu Y, Nagarajan H, Seo J-H, Cho B-K, Tsai S-F, Palsson BØ. 2012.
570 Comparative analysis of regulatory elements between *Escherichia coli* and *Klebsiella*
571 *pneumoniae* by genome-wide transcription start site profiling. *PLoS Genet* 8: e1002867.
572 doi: 10.1371/journal.pgen.1002867.
- 573 Kim K, Golubeva YA, Vanderpool CK, Slauch JM. 2018. Oxygen-dependent regulation of

- 574 SPI1 type three secretion system by small RNAs in *Salmonella enterica* serovar
575 *Typhimurium*. *Mol Microbiol* mmi.14174. doi: 10.1111/mmi.14174.
- 576 Kim W, Lee Y. 2020. Mechanism for coordinate regulation of *rpoS* by sRNA-sRNA interaction
577 in *Escherichia coli*. *RNA Biol* 17: 176–187. doi: 10.1080/15476286.2019.1672514.
- 578 Kröger C, Dillon SC, Cameron ADS, Papenfort K, Sivasankaran SK, Hokamp K, Chao Y,
579 Sittka A, Hébrard M, Händler K, et al. 2012. The transcriptional landscape and small
580 RNAs of *Salmonella enterica* serovar *Typhimurium*. *Proc Natl Acad Sci U S A* 109:
581 E1277-1286. doi: 10.1073/pnas.1201061109.
- 582 Kroner GM, Wolfe MB, Freddolino PL. 2019. *Escherichia coli* Lrp regulates one-third of the
583 genome via direct, cooperative, and indirect routes. *J Bacteriol* 201: e00411-18. doi:
584 10.1128/JB.00411-18.
- 585 Liu F, Chen Z, Zhang S, Wu K, Bei C, Wang C, Chao Y. 2023. *In vivo* RNA interactome
586 profiling reveals 3'UTR-processed small RNA targeting a central regulatory hub. *Nat*
587 *Commun* 14: 8106. doi: 10.1038/s41467-023-43632-1
- 588 Mandin P, Gottesman S. 2010. Integrating anaerobic/aerobic sensing and the general stress
589 response through the ArcZ small RNA. *EMBO J* 29: 3094–3107. doi:
590 10.1038/emboj.2010.179.
- 591 McCullen CA, Benhammou JN, Majdalani N, Gottesman S. 2010. Mechanism of positive
592 regulation by DsrA and RprA small noncoding RNAs: pairing increases translation and
593 protects *rpoS* mRNA from degradation. *J Bacteriol* 192: 5559–5571. doi:
594 10.1128/JB.00464-10.
- 595 Melamed S, Peer A, Faigenbaum-Romm R, Gatt YE, Reiss N, Bar A, Altuvia Y, Argaman L,
596 Margalit H. 2016. Global mapping of small RNA-target interactions in bacteria. *Mol*
597 *Cell* 63: 884–897. doi: 10.1016/j.molcel.2016.07.026.
- 598 Mika F, Hengge R. 2014. Small RNAs in the control of RpoS, CsgD, and biofilm architecture
599 of *Escherichia coli*. *RNA Biol*. 11: 494–507. doi: 10.4161/rna.28867.
- 600 Moll I, Afonyushkin T, Vytvytska O, Kaberdin VR, Bläsi U. 2003. Coincident Hfq binding and
601 RNase E cleavage sites on mRNA and small regulatory RNAs. *RNA* 9: 1308–1314. doi:
602 10.1261/rna.5850703
- 603 Monteiro C, Papenfort K, Hentrich K, Ahmad I, Le Guyon S, Reimann R, Grantcharova N,
604 Römling U. 2012. Hfq and Hfq-dependent small RNAs are major contributors to
605 multicellular development in *Salmonella enterica* serovar *Typhimurium*. *RNA Biol*. 9:
606 489–502. doi: 10.4161/rna.19682.
- 607 Moon K, Six DA, Lee H-J, Raetz CRH, Gottesman S. 2013. Complex transcriptional and post-

- 608 transcriptional regulation of an enzyme for lipopolysaccharide modification: regulation
609 of *eptB* expression. *Mol. Microbiol.* 89: 52–64. doi: 10.1111/mmi.12257.
- 610 Mukherjee A, Cui Y, Ma W, Liu Y, Chatterjee AK. 2000. *hexA* of *Erwinia carotovora* ssp.
611 *carotovora* strain Ecc71 negatively regulates production of RpoS and *rsmB* RNA, a
612 global regulator of extracellular proteins, plant virulence and the quorum-sensing signal,
613 N-(3-oxohexanoyl)- l-homoserine lactone. *Environ. Microbiol.* 2: 203–215. doi:
614 10.1046/j.1462-2920.2000.00093.x.
- 615 Nahku R, Peebo K, Valgepea K, Barrick JE, Adamberg K, Vilu R. 2011. Stock culture
616 heterogeneity rather than new mutational variation complicates short-term cell
617 physiology studies of *Escherichia coli* K-12 MG1655 in continuous culture.
618 *Microbiology (Reading)* 157: 2604–2610. doi: 10.1099/mic.0.050658-0.
- 619 Neubacher N, Tobias NJ, Huber M, Cai X, Glatter T, Pidot SJ, Stinear TP, Lütticke AL,
620 Papenfort K, Bode HB. 2020. Symbiosis, virulence and natural-product biosynthesis in
621 entomopathogenic bacteria are regulated by a small RNA. *Nat Microbiol* 5: 1481–1489.
622 doi: 10.1038/s41564-020-00797-5.
- 623 Papenfort K, Said N, Welsink T, Lucchini S, Hinton JCD, Vogel J. 2009. Specific and
624 pleiotropic patterns of mRNA regulation by ArcZ, a conserved, Hfq-dependent small
625 RNA. *Mol Microbiol* 74: 139–158. doi: 10.1111/j.1365-2958.2009.06857.x.
- 626 Pfeiffer V, Papenfort K, Lucchini S, Hinton JCD, Vogel J. 2009. Coding sequence targeting by
627 MicC RNA reveals bacterial mRNA silencing downstream of translational initiation.
628 *Nat Struct Mol Biol* 16: 840–846. doi: 10.1038/nsmb.1631.
- 629 Piqué N, Miñana-Galbis D, Merino S, Tomás JM. 2015. Virulence factors of *Erwinia*
630 *amylovora*: A Review. *Int J Mol Sci* 16: 12836–12854. doi: 10.3390/ijms160612836.
- 631 Rachwalski K, Ellis MJ, Tong M, Brown ED. 2022. Synthetic genetic interactions reveal a
632 dense and cryptic regulatory network of small noncoding RNAs in *Escherichia coli*.
633 *mBio* 13: e01225-22. doi: 10.1128/mbio.01225-22.
- 634 Römling U. 2005. Characterization of the rdar morphotype, a multicellular behaviour in
635 *Enterobacteriaceae*. *CMLS, Cell Mol Life Sci* 62: 1234–1246. doi: 10.1007/s00018-
636 005-4557-x.
- 637 Saliba A-E, Santos SC, Vogel J. 2017. New RNA-seq approaches for the study of bacterial
638 pathogens. *Curr Opin Microbiol* 35: 78–87. doi: 10.1016/j.mib.2017.01.001.
- 639 Santander RD, Figàs-Segura À, Biosca EG. 2018. *Erwinia amylovora* catalases KatA and KatG
640 are virulence factors and delay the starvation-induced viable but non-culturable
641 (VBNC) response. *Mol Plant Pathol* 19: 922–934. doi: 10.1111/mpp.12577.

- 642 Saxena S, Jónsson ZO, Dutta A. 2003. Small RNAs with imperfect match to endogenous
643 mRNA repress translation. Implications for off-target activity of small inhibitory RNA
644 in mammalian cells. *J Biol Chem* 278: 44312–44319. doi: 10.1074/jbc.M307089200.
- 645 Saxer G, Krepps MD, Merkley ED, Ansong C, Kaiser BLD, Valovska M-T, Ristic N, Yeh PT,
646 Prakash VP, Leiser OP, et al. 2014. Mutations in global regulators lead to metabolic
647 selection during adaptation to complex environments. *PLoS Genet* 10: e1004872. doi:
648 10.1371/journal.pgen.1004872.
- 649 Schachterle JK, Onsay DM, Sundin GW. 2019a. Small RNA ArcZ regulates oxidative stress
650 response genes and regulons in *Erwinia amylovora*. *Front Microbiol* 10: 2775. doi:
651 10.3389/fmicb.2019.02775.
- 652 Schachterle JK, Sundin GW. 2019. The Leucine-Responsive Regulatory Protein Lrp
653 participates in virulence regulation downstream of small RNA ArcZ in *Erwinia*
654 *amylovora*. *mBio* 10: e00757-19. doi: 10.1128/mBio.00757-19.
- 655 Schachterle JK, Zeng Q, Sundin, G.W. 2019b. Three Hfq-dependent small RNAs regulate
656 flagellar motility in the fire blight pathogen *Erwinia amylovora*. *Mol Microbiol* 111:
657 1476–1492. doi: 10.1111/mmi.14232.
- 658 Schneider P, Jacobs JM, Neres J, Aldrich CC, Allen C, Nett M, Hoffmeister D. 2009. The global
659 virulence regulators VsrAD and PhcA control secondary metabolism in the plant
660 pathogen *Ralstonia solanacearum*. *ChemBioChem* 10: 2730–2732. doi:
661 10.1002/cbic.200900510.
- 662 Sedlyarova N, Shamovsky I, Bharati BK, Epshtein V, Chen J, Gottesman S, Schroeder R,
663 Nudler E. 2016. sRNA-mediated control of transcription termination in *E. coli*. *Cell*
664 167: 111–121.e13. doi: 10.1016/j.cell.2016.09.004.
- 665 Semsey S. 2014. A mixed incoherent feed-forward loop allows conditional regulation of
666 response dynamics. *PLoS One* 9: e91243. doi: 10.1371/journal.pone.0091243.
- 667 Sharma CM, Darfeuille F, Plantinga TH, Vogel J. 2007. A small RNA regulates multiple ABC
668 transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-
669 binding sites. *Genes Dev* 21: 2804–2817. doi: 10.1101/gad.447207
- 670 Shin GY, Schachterle JK, Shyntum DY, Moleleki LN, Coutinho TA, Sundin GW. 2019.
671 Functional characterization of a global virulence regulator Hfq and identification of
672 Hfq-dependent sRNAs in the plant pathogen *Pantoea ananatis*. *Front Microbiol* 10:
673 2075. doi: 10.3389/fmicb.2019.02075.
- 674 Smith P, Schuster M. 2019. Public goods and cheating in microbes. *Curr Biol* 29: R442–R447.
675 doi: 10.1016/j.cub.2019.03.001.

- 676 Soper T, Mandin P, Majdalani N, Gottesman S, Woodson SA. 2010. Positive regulation by
677 small RNAs and the role of Hfq. *Proc Natl Acad Sci U S A* 107: 9602–9607. doi:
678 10.1073/pnas.1004435107.
- 679 Spira B, de Almeida Toledo R, Maharjan RP, Ferenci T. 2011. The uncertain consequences of
680 transferring bacterial strains between laboratories - *rpoS* instability as an example. *BMC*
681 *Microbiol* 11: 248. doi: 10.1186/1471-2180-11-248.
- 682 Su SS, Modrich P. 1986. *Escherichia coli mutS*-encoded protein binds to mismatched DNA
683 base pairs. *Proc Natl Acad Sci U S A* 83: 5057–5061. doi: 10.1073/pnas.83.14.5057.
- 684 Thomason MK, Fontaine F, De Lay N, Storz G. 2012. A small RNA that regulates motility and
685 biofilm formation in response to changes in nutrient availability in *Escherichia coli*.
686 *Mol Microbiol* 84: 17–35. doi: 10.1111/j.1365-2958.2012.07965.x.
- 687 Thompson JD, Gibson TJ, Higgins DG. 2003. Multiple Sequence Alignment Using ClustalW
688 and ClustalX. *Curr Protoc in Bioinform* 00: 2.3.1-2.3.22. doi:
689 10.1002/0471250953.bi0203s00.
- 690 Tobias NJ, Heinrich AK, Eresmann H, Wright PR, Neubacher N, Backofen R, Bode HB. 2017a.
691 *Photorhabdus*-nematode symbiosis is dependent on hfq-mediated regulation of
692 secondary metabolites. *Environ Microbiol* 19: 119–129. doi: 10.1111/1462-
693 2920.13502.
- 694 Tobias NJ, Wolff H, Djahanschiri B, Grundmann F, Kronenwerth M, Shi Y-M, Simonyi S,
695 Grün P, Shapiro-Ilan D, Pidot SJ, et al. 2017b. Natural product diversity associated with
696 the nematode symbionts *Photorhabdus* and *Xenorhabdus*. *Nat Microbiol* 2: 1676–1685.
697 doi: 10.1038/s41564-017-0039-9.
- 698 Troshin PV, Procter JB, Barton GJ. 2011. Java bioinformatics analysis web services for
699 multiple sequence alignment--JABAWS:MSA. *Bioinformatics* 27: 2001–2002. doi:
700 10.1093/bioinformatics/btr304.
- 701 Updegrove TB, Kouse AB, Bandyra KJ, Storz G. 2019. Stem-loops direct precise processing
702 of 3' UTR-derived small RNA MicL. *Nucleic Acids Res* 47: 1482–1492. doi:
703 10.1093/nar/gky1175.
- 704 Updegrove TB, Zhang A, Storz G. 2016. Hfq: the flexible RNA matchmaker. *Curr Opin*
705 *Microbiol* 30: 133–138. doi: 10.1016/j.mib.2016.02.003.
- 706 Vogel J, Luisi BF. 2011. Hfq and its constellation of RNA. *Nat Rev Microbiol* 9: 578–589. doi:
707 10.1038/nrmicro2615.
- 708 Wang C, Pu T, Lou W, Wang Y, Gao Z, Hu B, Fan J. 2018. Hfq, a RNA chaperone, contributes
709 to virulence by regulating Plant Cell Wall-Degrading Enzyme production, Type VI

- 710 Secretion System expression, bacterial competition, and suppressing host defense
711 response in *Pectobacterium carotovorum*. *Mol Plant Microbe Interact* 31: 1166–1178.
712 doi: 10.1094/MPMI-12-17-0303-R.
- 713 Wassarman KM, Repoila F, Rosenow C, Storz G, Gottesman S. 2001. Identification of novel
714 small RNAs using comparative genomics and microarrays. *Genes Dev* 15: 1637–1651.
715 doi: 10.1101/gad.901001.
- 716 Wei X, Tan H, Lobb B, Zhen W, Wu Z, Parks DH, Neufeld JD, Moreno-Hagelsieb G, Doxey
717 AC. 2024. AnnoView enables large-scale analysis, comparison, and visualization of
718 microbial gene neighborhoods. *bioRxiv* doi: 10.1101/2024.01.15.575735
- 719 Wright PR, Georg J, Mann M, Sorescu DA, Richter AS, Lott S, Kleinkayf R, Hess WR,
720 Backofen R. 2014. CopraRNA and IntaRNA: predicting small RNA targets, networks
721 and interaction domains. *Nucleic Acids Res* 42: W119-123. doi: 10.1093/nar/gku359.
- 722 Wyrzykowski J, Volkert MR. 2003. The *Escherichia coli* methyl-directed mismatch repair
723 system repairs base pairs containing oxidative lesions. *J Bacteriol* 185: 1701–1704. doi:
724 10.1128/JB.185.5.1701-1704.2003.
- 725 Yang S, Peng Q, Zhang Q, Yi X, Choi CJ, Reedy RM, Charkowski AO, Yang CH. 2008.
726 Dynamic regulation of GacA in Type III secretion, pectinase gene expression, pellicle
727 formation, and pathogenicity of *Dickeya dadantii* (*Erwinia chrysanthemi* 3937). *Mol*
728 *Plant Microbe Interact* 21: 133–142. doi: 10.1094/MPMI-21-1-0133.
- 729 Yuan X, Zeng Q, Khokhani D, Tian F, Severin GB, Waters CM, Xu J, Zhou X, Sundin GW,
730 Ibekwe AM et al. 2019. A feed-forward signalling circuit controls bacterial virulence
731 through linking cyclic di-GMP and two mechanistically distinct sRNAs, ArcZ and
732 RsmB. *Environmental Microbiology* 21: 2755–2771. doi: 10.1111/1462-2920.14603.
- 733 Zeng Q, McNally RR, Sundin GW. 2013. Global small RNA chaperone Hfq and regulatory
734 small RNAs are important virulence regulators in *Erwinia amylovora*. *J Bacteriol* 195:
735 1706–1717. doi: 10.1128/JB.02056-12.
- 736 Zhao K, Liu M, Burgess RR. 2007. Adaptation in bacterial flagellar and motility systems: from
737 regulon members to 'foraging'-like behavior in *E. coli*. *Nucleic Acids Res* 35: 4441–
738 4452. doi: 10.1093/nar/gkm456.
- 739
740
741
742 **Figure 1 : ArcZ, a cleaved sRNA that is highly conserved in its 3' part.**

743 (A) Alignment of ArcZ sequence from *Pantoea ananatis* (NZ_CM012203), *Escherichia coli*
 744 (NC_000913.3), *Salmonella typhimurium* (NC_003197.2), *Klebsiella pneumoniae* (NC_009648),
 745 *Photobacterium laumondii* (CP024901.1), *Xenorhabdus nematophila* (CP060401.1), *Dickeya dadantii*
 746 (CP002038.1), *Pectobacterium carotovorum* (CP051652.1), *Yersinia enterocolitica* (CP107102.1),
 747 *Serratia marcescens* (CP139958.1), *Proteus mirabilis* (CP045257.1), *Erwinia amylovora* (FN666575.1),
 748 *Providencia alcalifaciens* (CP084296.1), *Edwardsiella tarda* (CP084506.1), *Shigella sonnei*
 749 (CP026802.1), *Citrobacter freundii* (CP049015.1), *Enterobacter cloacae* (CP001918.1), *Cronobacter*
 750 *sakazakii* (CP011047.1). Conservation score is plotted below, and the conserved region is colored in
 751 red. This alignment was carried out using ClustalW and Jalview (Thompson et al. 2003; Clamp et al.
 752 2004; Troshin et al. 2011). The red squares correspond to known transcription starts.
 753 (B) Synteny analysis of chromosomal regions surrounding ArcZ performed using Annoview (Wei et al.
 754 2024). The same genome accession numbers as panel A were used.
 755 (C) Model of ArcZ maturation in *E. coli*. The stem-loop represents the Rho-independent
 756 transcriptional termination site of *arcZ*. The 3' region of ArcZ is recognized by Hfq through binding.
 757 RNase E cleaves ArcZ at a consensus sequence, producing a mature processed form of ArcZ that binds
 758 to target mRNA.

759

760 **Figure 2. ArcZ targets in *Escherichia coli*.**

761 Green arrows indicate activation, red arrows indicate repression, and the black arrows indicate either
 762 translation or activity. CyaR is an sRNA capable of repressing the translation of *nadE* mRNA, which
 763 encodes an enzyme involved in NAD⁺ biosynthesis but also capable of repressing the translation of
 764 *rpoS* mRNA. ArcZ degrades CyaR via RNase E, thereby increasing the translation of *nadE* and the
 765 availability of NAD⁺ and *rpoS*. ArcZ enhances the translation of *rpoS* mRNA, which induces the Gad
 766 pathway, leading to better acid stress resistance. Additionally, ArcZ reduces the translation of *mutS*
 767 mRNA, both directly and indirectly, by activating the translation of *rpoS*, which in turn transcribes
 768 SdsR. SdsR directly represses *mutS* translation. ArcZ also directly represses *flhDC* translation and
 769 competes with the McaS sRNA, which has a common *flhDC* mRNA pairing site with ArcZ. McaS
 770 activates *flhDC* translation, while ArcZ inhibits LPS modification by repressing *eptB* mRNA
 771 translation. Created with BioRender.com.

772

773 **Figure 3. ArcZ targets in *Erwinia amylovora*.**

774 Green arrows indicate activation, red arrows indicate repression, and the black arrows indicate either
 775 translation or activity. ArcZ plays a role in regulating the response to oxidative stress by modulating
 776 the levels of the anti-oxidant enzymes KatA and Tpx. The abbreviation NC stands for 'Non-

777 characterized' protein. ArcZ modulates motility and biofilm formation through the control of *flhDC*
778 and *lrp* translation. Created with BioRender.com.

779

780 **Figure 4. ArcZ targets in *Salmonella typhimurium*.**

781 Green arrows indicate activation, red arrows indicate repression, and the black arrows indicate either
782 translation or activity. ArcZ represses the expression of STM₃₂₁₆, a potential chemoreceptor
783 acquired through horizontal gene transfer. As previously described, ArcZ modulates the response to
784 oxidative stress by repressing *tpx* mRNA. Additionally, ArcZ represses the translation of *sdaC* mRNA,
785 which is involved in serine catabolism. ArcZ and FnrS inhibit the translation of *hilD*, which is
786 responsible for activating the T₃SS synthesis. However, *fnrS* is transcribed under anoxic conditions
787 due to Fnr, whereas *arcZ* is transcribed under strong aerobic conditions. Therefore, the activation of
788 T₃SS is limited by the presence of oxygen. Created with BioRender.com.

789

790 **Figure 5. Regulation of *hexA/pecT* by ArcZ in *Dickeya* and *Photorhabdus*.**

791 Direct and indirect targets of ArcZ are shown in *Dickeya* (green panel) and *Photorhabdus* (brown
792 panel). Green arrows indicate activation, red arrows indicate repression, and the black arrows indicate
793 either translation or activity. In *Dickeya*, ArcZ inhibits the translation of *pecT*, which prevents the
794 inhibition of *rsmB* transcription by PecT. RsmB is a small Hfq-independent regulatory RNA that binds
795 to RsmA and prevents it from repressing the expression of T₃SS and pectinases. In *Photorhabdus*,
796 ArcZ inhibits the translation of *hexA* and thus enhances the production of secondary metabolites,
797 which are essential for nematode symbiosis. Created with BioRender.com.

798

799 **Figure 6. ArcZ's modes of action.**

800 ArcZ can activate the translation of mRNA (green panel). (A) ArcZ interacts with the 5'UTR region of
801 the *rpoS* mRNA and releases its RBS, which is initially involved in a hairpin structure, allowing for
802 translation of *rpoS*. (B) Additionally, during transcription of *rpoS*, ArcZ prevents the Rho termination
803 factor from binding to the mRNA, thus blocking premature termination. ArcZ can also repress the
804 translation of mRNA (red panel), such as with *mutS* or *flhDC*. (C) ArcZ can repress translation by
805 binding to the RBS of the mRNA, preventing access of the ribosome. Additionally, it can bind
806 upstream of the RBS, such as with *pecT* or *flhD*, forming a secondary structure that is incompatible
807 with translation. (D) ArcZ can also bind to the coding region of the mRNA or sRNA, making it more
808 susceptible to RNase such as with *tpx* and *cyaR*. Created with BioRender.com.

809

810 **Figure 7. *rpoS* and *flhD*, two ArcZ targets conserved in *Enterobacteriales*.**

811 Alignment of (A) *rpoS* mRNA sequence and (B) *flhD* mRNA sequence from *Pantoeae ananatis*,
812 *Escherichia coli*, *Salmonella typhimurium*, *Klebsiella pneumoniae*, *Photobacterium laumondii*,
813 *Xenorhabdus nematophila*, *Dickeya dadantii*, *Pectobacterium carotovorum*, *Serratia marcescens*,
814 *Proteus mirabilis*, *Erwinia amylovora*, *Providencia alcalifaciens*, *Edwardsiella tarda*, *Shigella sonnei*,
815 *Citrobacter freundii*, *Enterobacter cloacae*, *Cronobacter sakazakii* (genomes are the same as those used
816 in Figure 1). Conservation score is plotted below, and on this plot the interaction zone between ArcZ
817 and *rpoS* / ArcZ and *flhD* shown in *E. coli* and *E. amylovora* is colored red. The interaction of *E. coli* ArcZ
818 (in red) with *rpoS* and *flhD* mRNA and of *E. amylovora* ArcZ (in green) with *flhD* mRNA is shown
819 (Mandin and Gottesman 2010; De Lay and Gottesman 2012; Schachterle et al. 2019b). The red squares
820 correspond to ArcZ binding site regions conserved between *E. coli* and other bacteria. Alignments
821 were performed using MUSCLE (Edgar 2004) and processed with Jalview (Clamp et al. 2004).

822

823 **Figure 8. Alignments of mutated *arcZ* alleles in *Enterobacteriales*.**

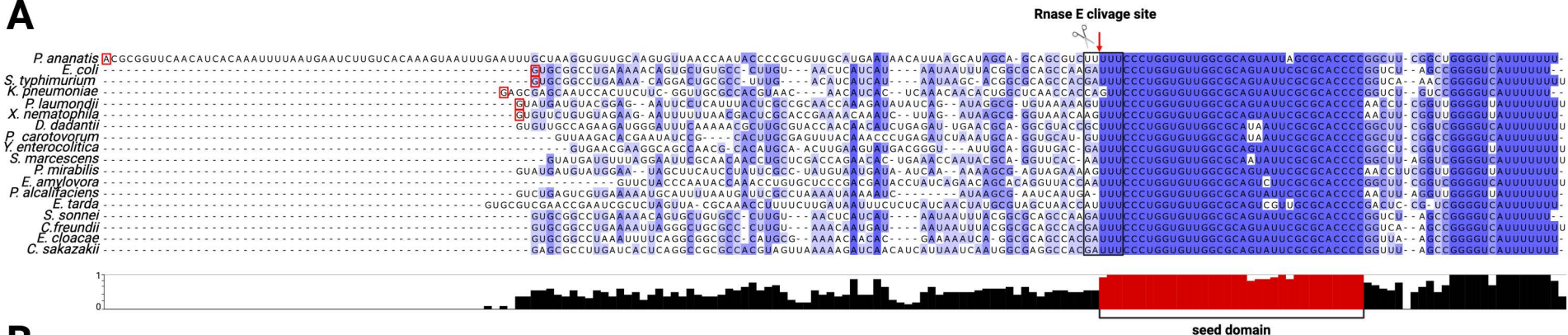
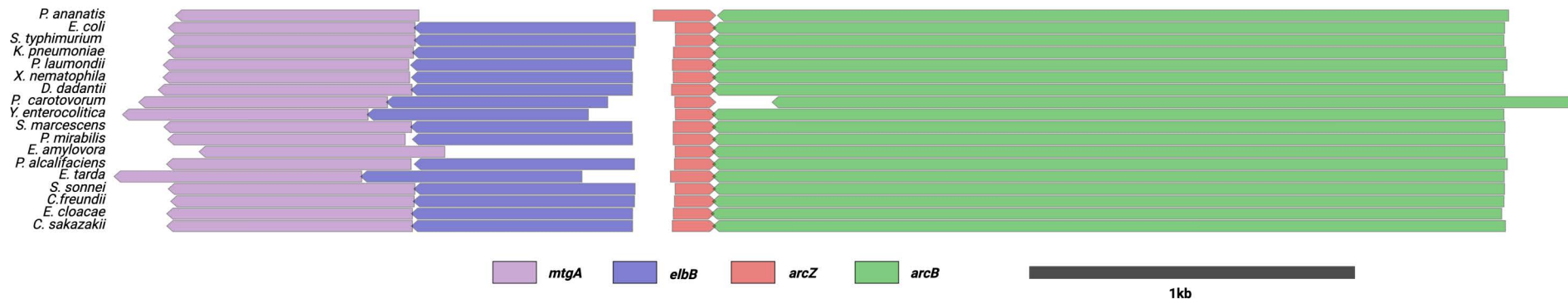
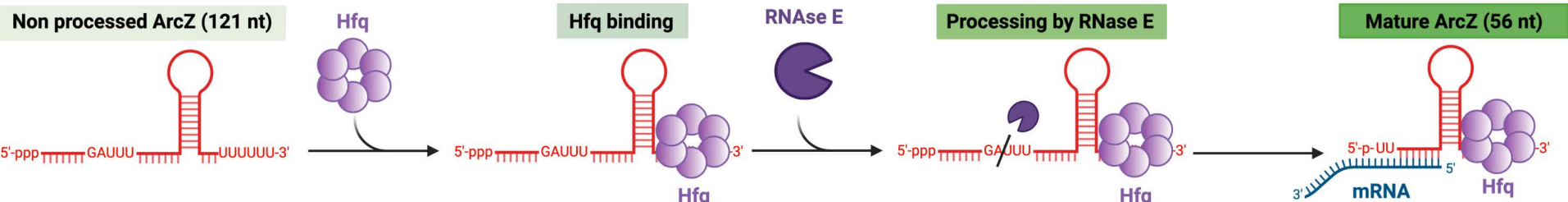
824 A BLASTN analysis was conducted on the *enterobacteriales* NCBI RefSeq genomes using the
825 conserved nucleotide sequence that contains the mutation (A) G to A or (B) G to U at position 90 in
826 comparison with the *E. coli* MG1655 reference genome. The mutation is indicated by a red arrow and
827 surrounded by a red rectangle. Alignments were performed using MUSCLE (Edgar 2004) and
828 processed with Jalview (Clamp et al. 2004).

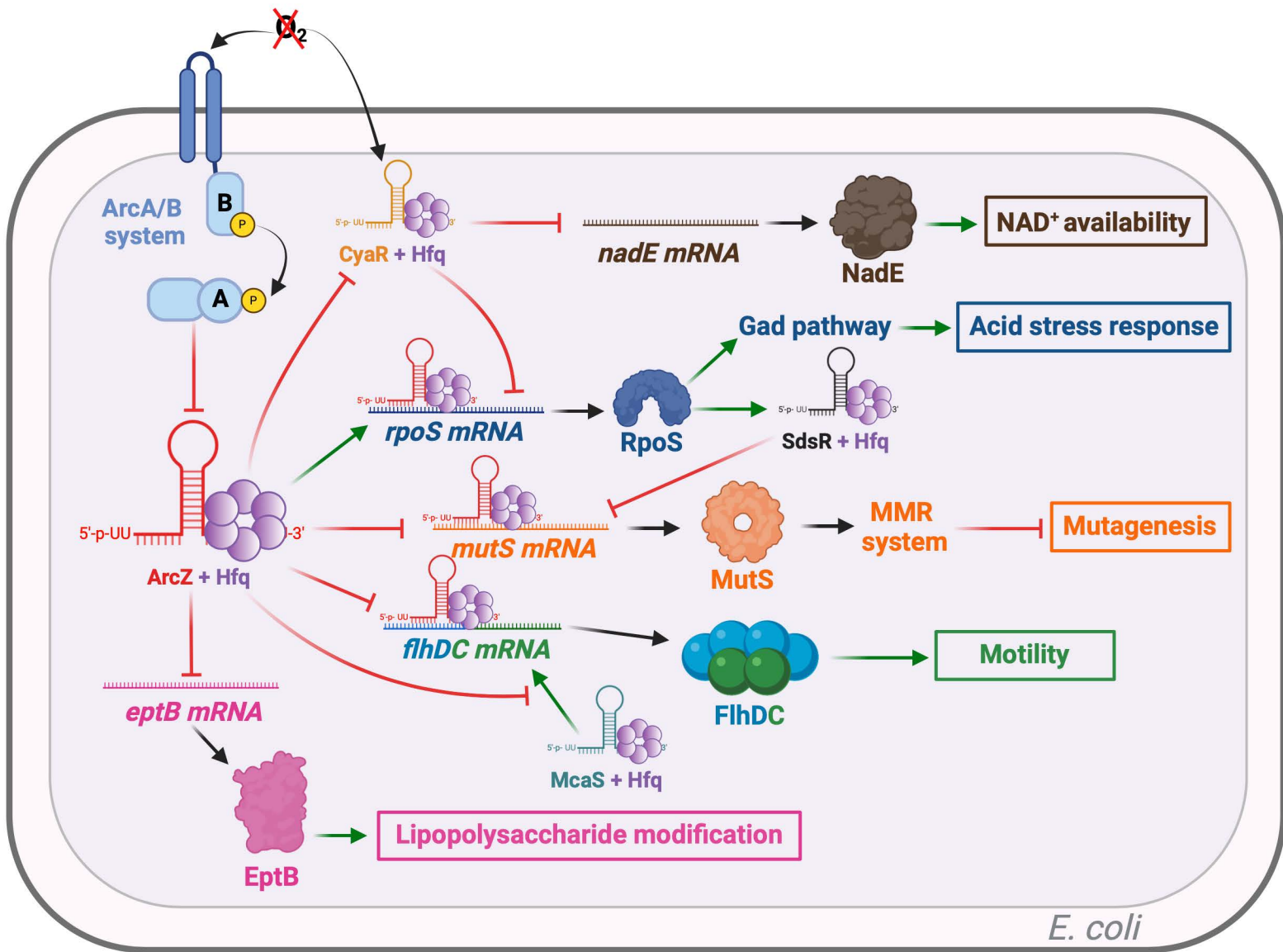
829

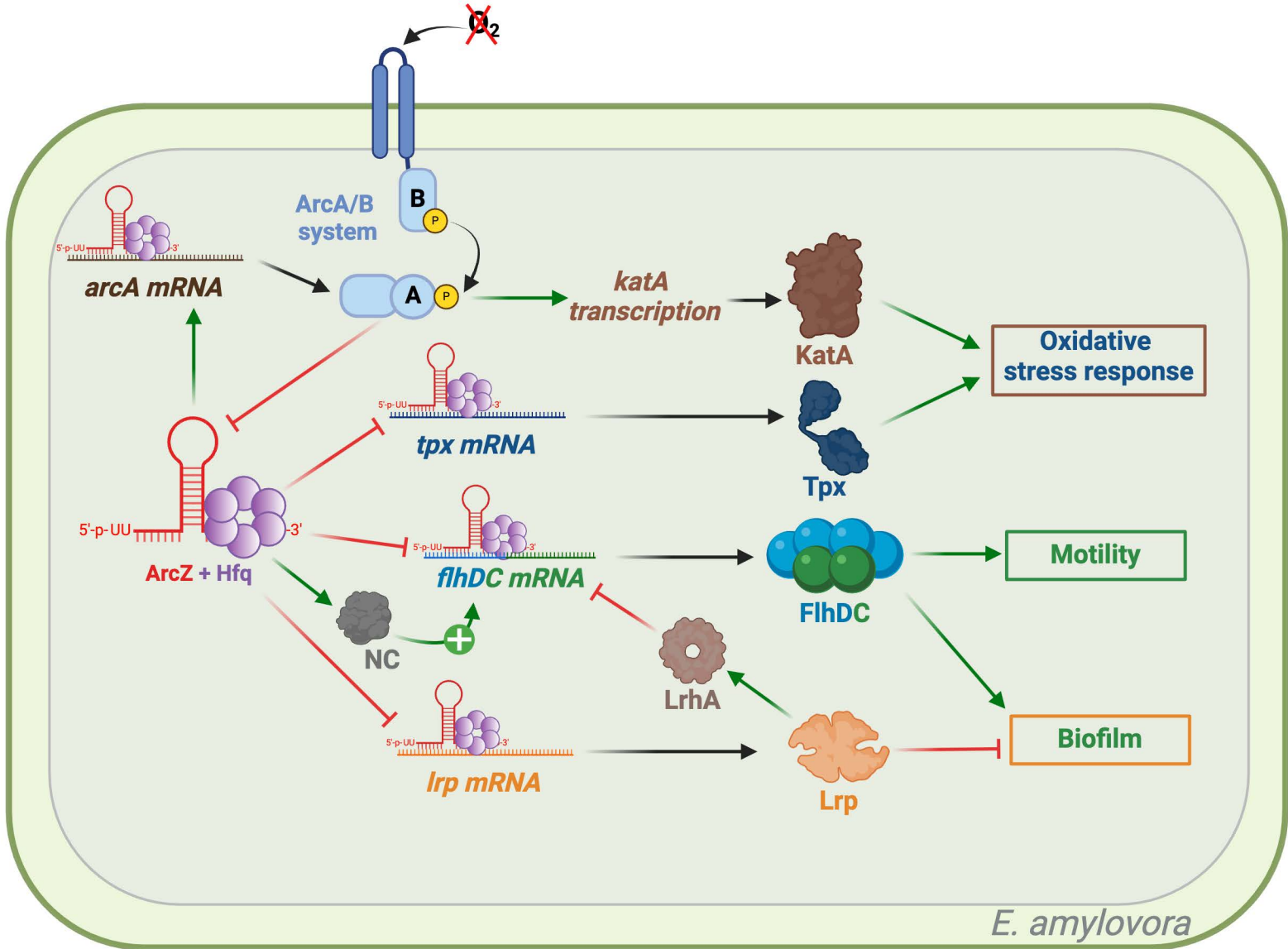
830

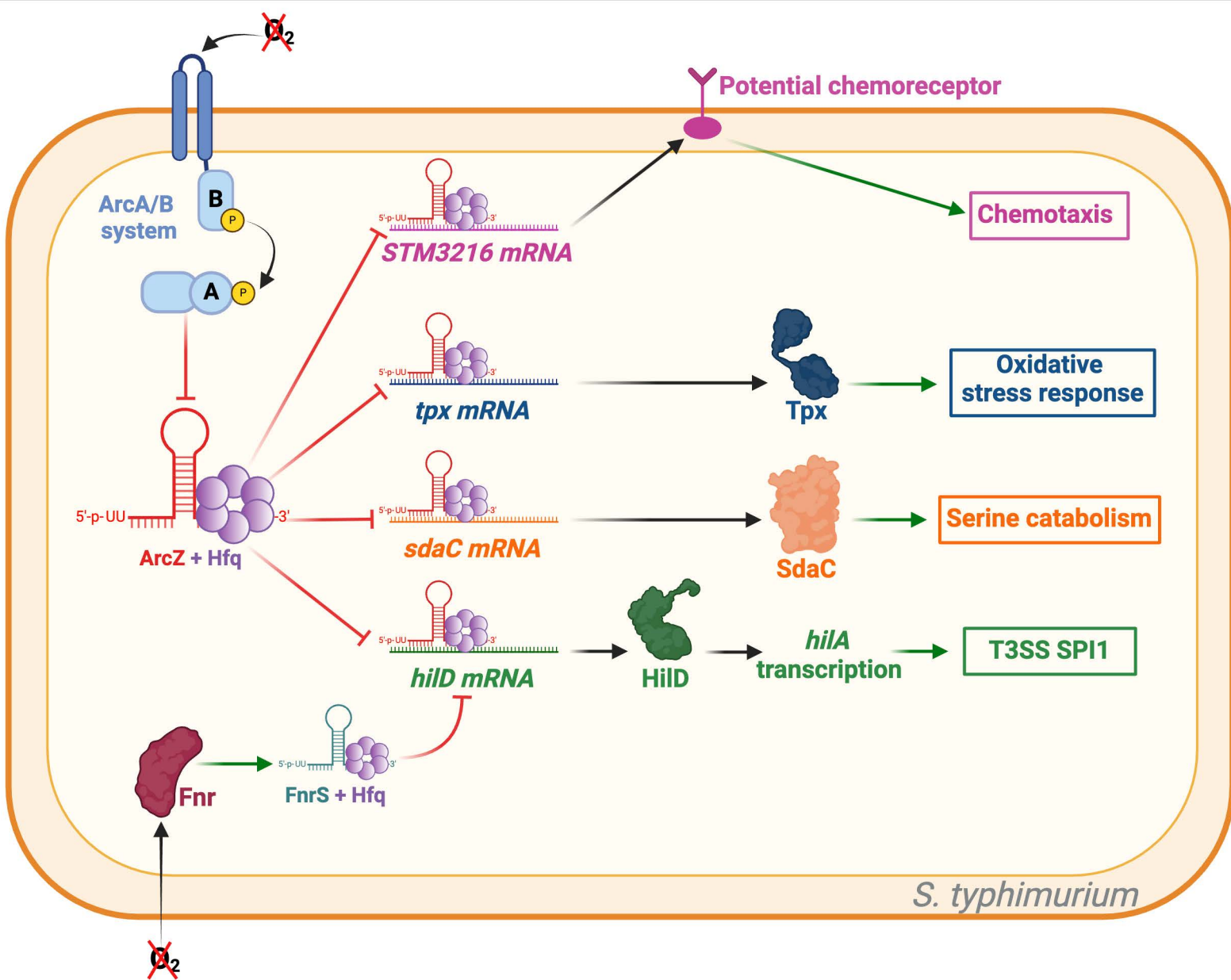
831

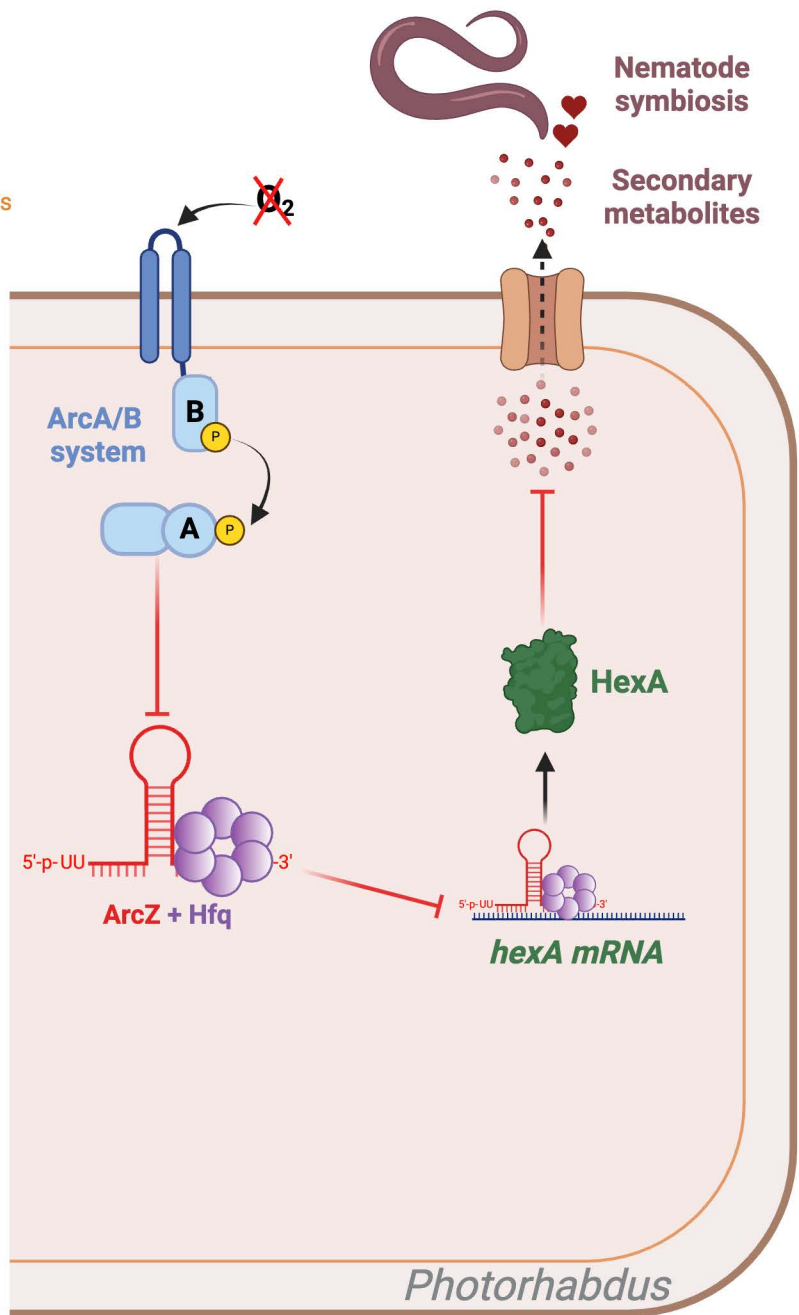
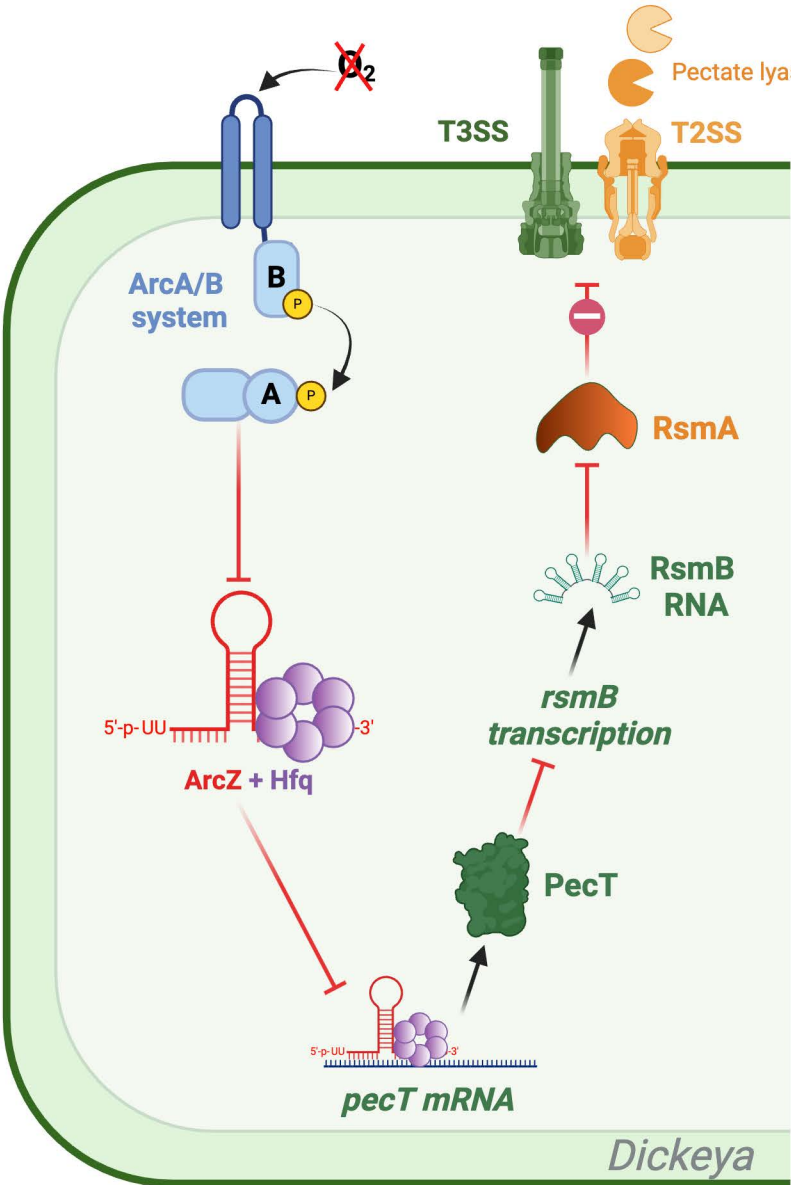
832

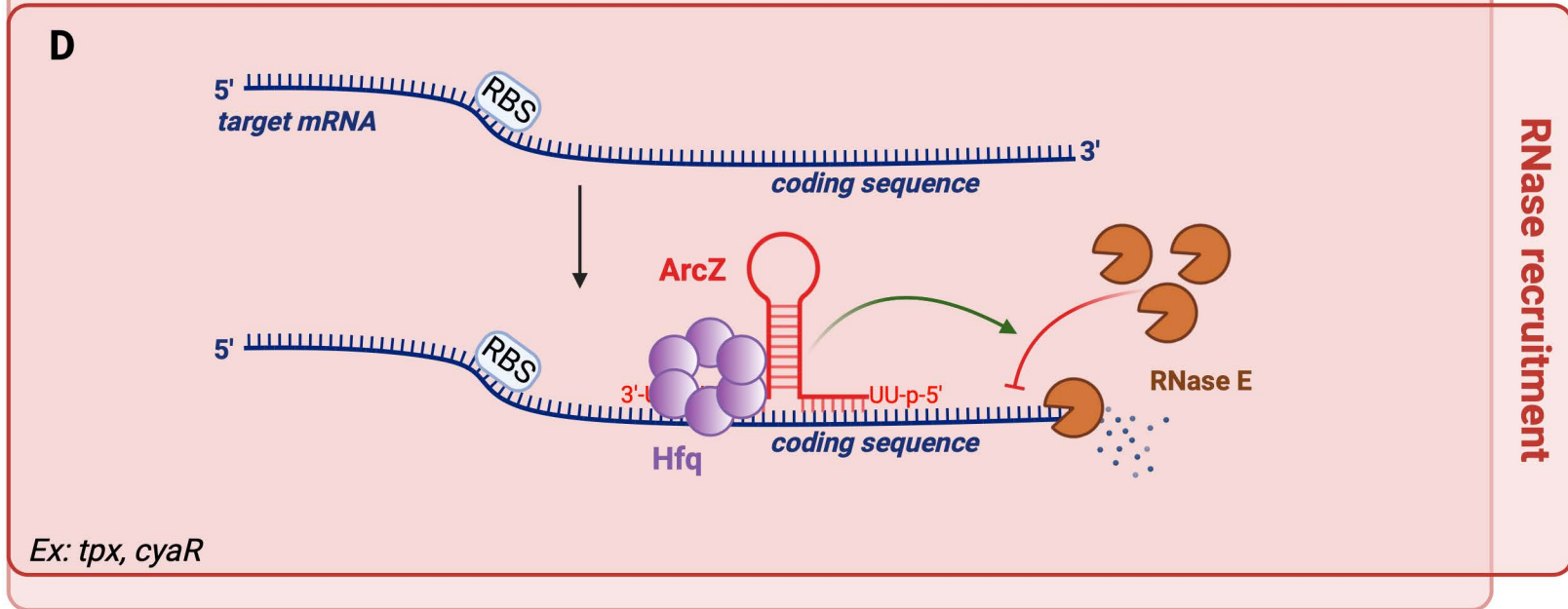
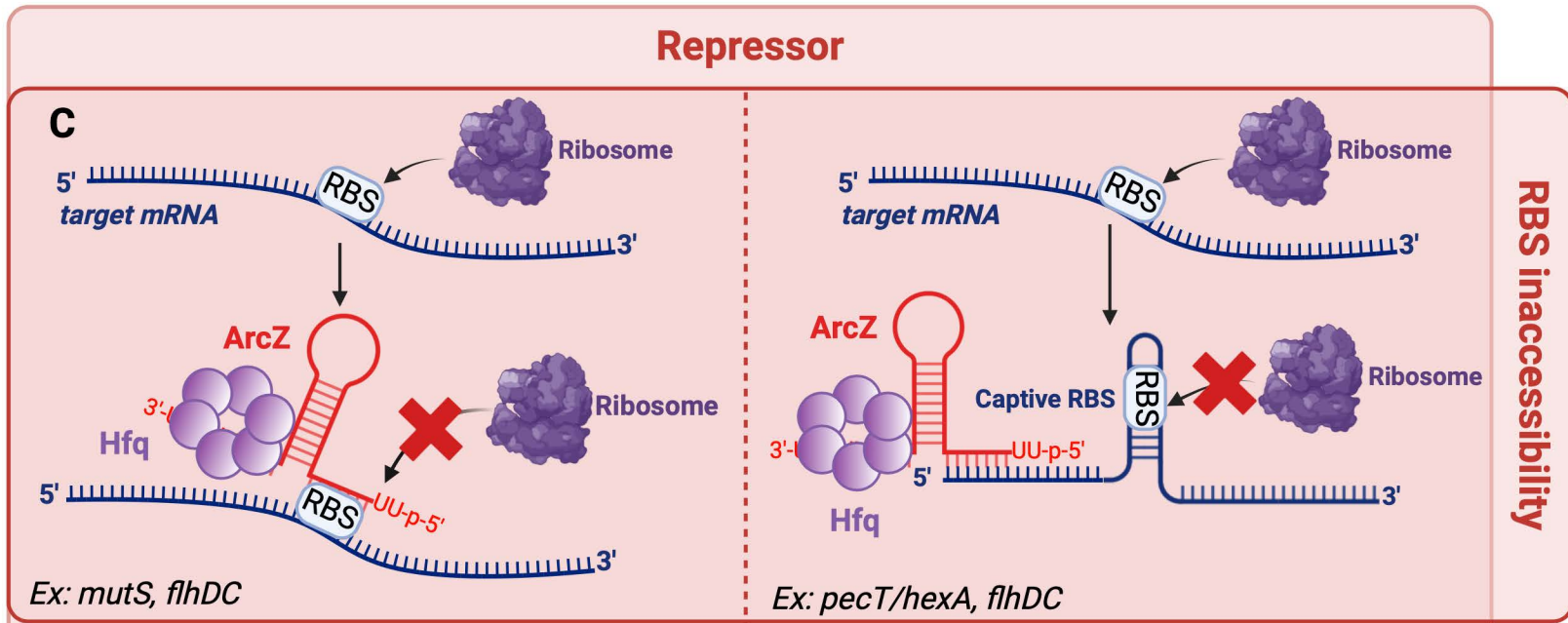
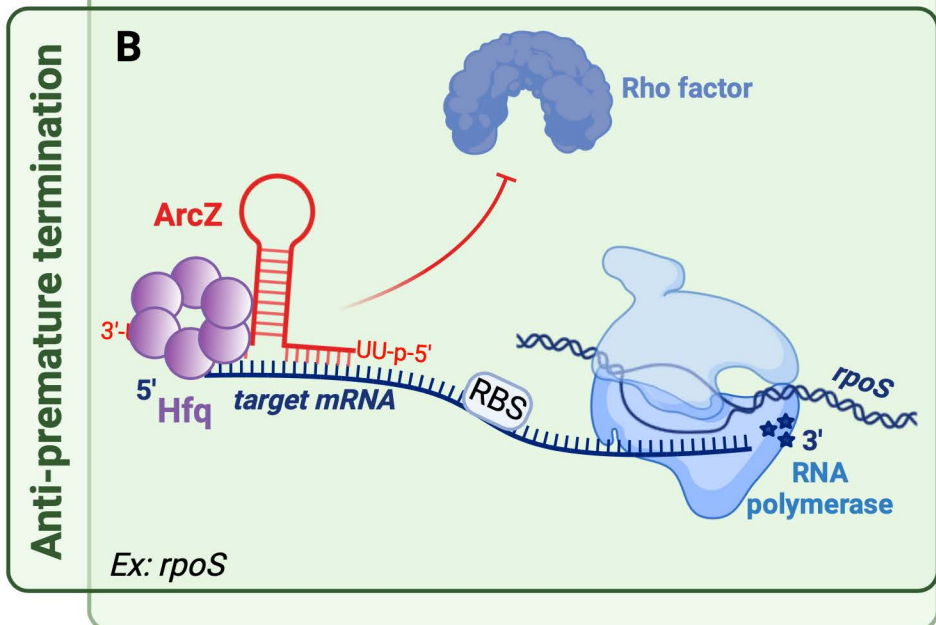
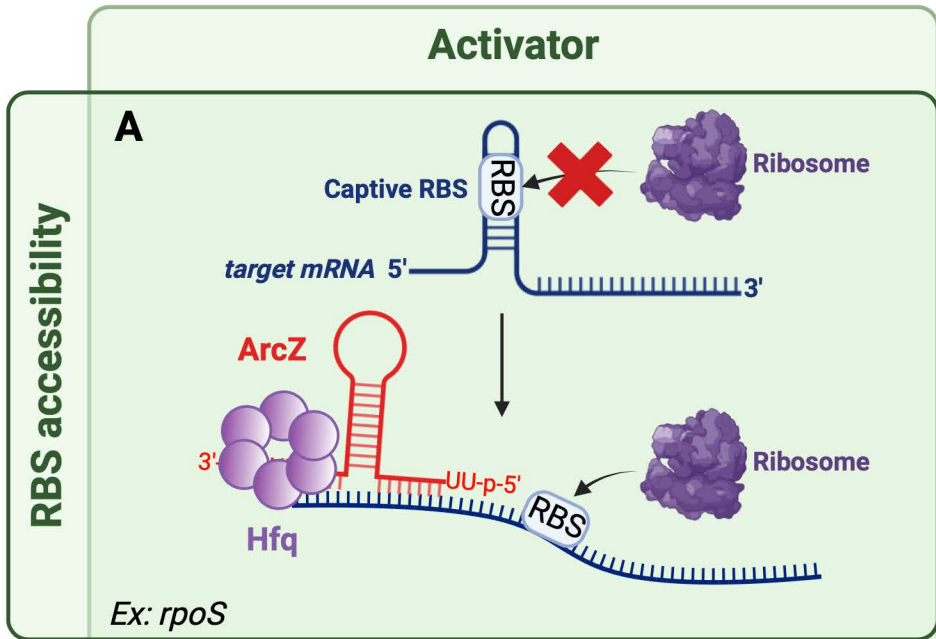
A**B****C**









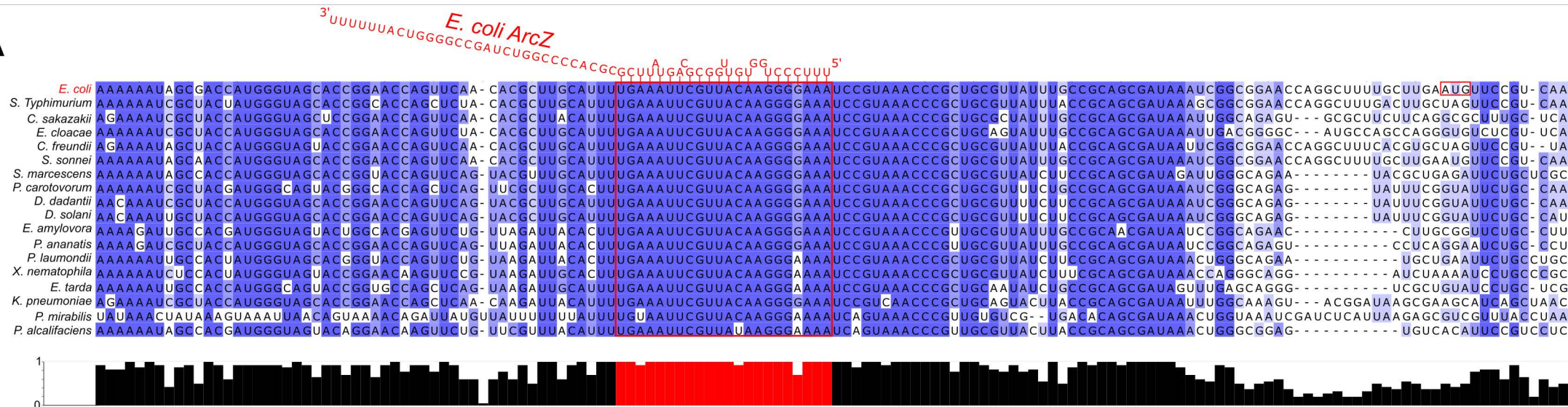
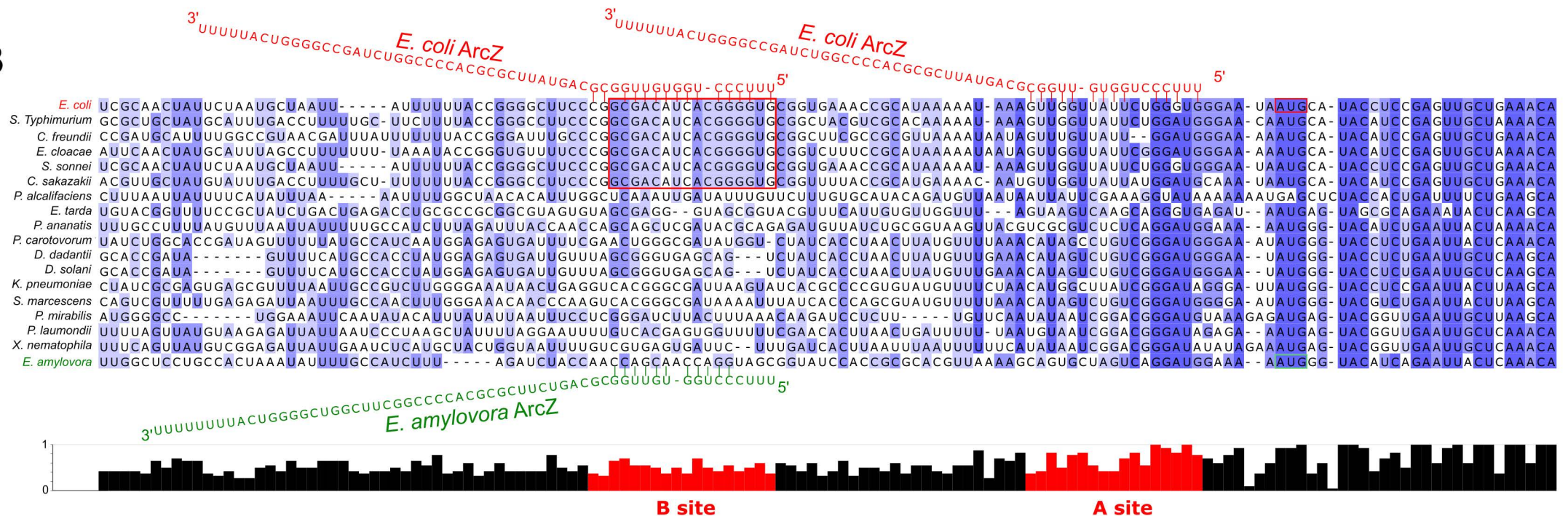


RBS accessibility

Anti-premature termination

RBS inaccessibility

RNase recruitment

A**B**

A

E. coli str. MG1655 GUGCGGCCU-----GAAAAACAGUGCUGUGCCCUUGUAACUCAUCA-UAAUAAUUUACGGCGCAGCCAAGA-UUUUCCUGGUGUUGGC-CAGUAUUCGCGCACCCCGGUC--UAGC-CGGGGUCAUUUUUU--
C. youngae str. CCUG-30791 --CUUAACUGCGUGCGGCCU-----GAAAAUAGGGCUGCGCCCUUGUAACAAUGA-UAAUAAUUUACGGCGCAGCCAAGA-UUUUCCUGGUGUUGGC-CAGUAUUCGCGCACCCCGGUC--AAGC-CGGGGUCAUUUUUUUU--
S. Typhimurium str. STm8-85 --GCUAUCUUAACUGCGUGCGGCCUGAAAAACAGGACUGCGCCUUUGACAUCAUCAU---AAUAAGCACGGCGCAGCCACGA-UUUUCCUGGUGUUGGC-CAGUAUUCGCGCACCCCGGUC--AAAC-CGGGGUCAUUUUUUUU--
S. Moscow str. S-1843 --GCUAUCUUAACUGCGUGCGGCCUGAAAAACAGGACUGCGCCUUUGACAUCAUCAU---AAUAAGCACGGCGCAGCCACGA-UUUUCCUGGUGUUGGC-CAGUAUUCGCGCACCCCGGUC--AAAC-CGGGGUCAUUUUUUUU--
D. parazea str. S31 ----GUGUGU-UGCCAGAUGAUGGGAUUUUCAAACGCUUGCGUACCAUGACAUCUGAGAGCGAACGCAGGUGUACCCAUUUUCCUGGUGUUGGC-CAUAUUAUCGCGCACCCCGGCU-UCGGUCGGGGUCAUUUUUUUU--
D. fangzhongdai str. DSM-101947 ----GUGUGU-UGCCAGAA GAUGGGAUUUUCAAACGCUUGCGUACCAACAACAUCUGAGAUAUGAACGCAGGUGUACCGU-UUUUCCUGGUGUUGGC-CAUAUUAUCGCGCACCCCGGCC-UCGGUCGGGGUCAUUUUUUUU--
D. fangzhongdai str. QZH3 ----GUGUGU-UGCCAGAAGAUGGGAUUUUCAAACGCUUGCGUACCAACAACAUCUGAGAUAUGAACGCAGGUGUACCGU-UUUUCCUGGUGUUGGC-CAUAUUAUCGCGCACCCCGGCC-UCGGUCGGGGUCAUUUUUUUU--
D. fangzhongdai str. LN1 ----GUGUGU-UGCCAGAA GAUGGGAUUUUCAAACGCUUGCGUACCAACAACAUCUGAGAUAUGAACGCAGGUGUACCGU-UUUUCCUGGUGUUGGC-CAUAUUAUCGCGCACCCCGGCC-UCGGUCGGGGUCAUUUUUUUU--
D. fangzhongdai str. CGMCC1.15464 ----GUGUGU-UGCCAGAAGAUGGGAUUUUCAAACGCUUGCGUACCAACAACAUCUGAGAUAUGAACGCAGGUGUACCGU-UUUUCCUGGUGUUGGC-CAUAUUAUCGCGCACCCCGGCC-UCGGUCGGGGUCAUUUUUUUU--
D. solani IPO2222 ----GUGUGU-UGCCAGAA GAUGGGAUUUUCAAACGCUUGCGUACCAACAUAUCUGAGAUAUGAACGCAGGUGUACCGU-UUUUCCUGGUGUUGGC-CAUAUUAUCGCGCACCCCGACC-UCGGUCGGGGUCAUUUUUUUU--
D. solani str. DsR34 ----GUGUGU-UGCCAGAAGAUGGGAUUUUCAAACGCUUGCGUACCAACAUAUCUGAGAUAUGAACGCAGGUGUACCGU-UUUUCCUGGUGUUGGC-CAUAUUAUCGCGCACCCCGACC-UCGGUCGGGGUCAUUUUUUUU--
D. solani str. DsR207 ----GUGUGU-UGCCAGAA GAUGGGAUUUUCAAACGCUUGCGUACCAACAUAUCUGAGAUAUGAACGCAGGUGUACCGU-UUUUCCUGGUGUUGGC-CAUAUUAUCGCGCACCCCGACC-UCGGUCGGGGUCAUUUUUUUU--
P. mirabilis str. NRZ-4144 ----UGCUUGUAUGAUGUAU--GGAAUAGCUUUAUCUUAUUCGCCUAUGUAUGA-UAAUCAAAAAGCGAGUAGAAAAG-UUUUCCUGGUGUUGGC-CAGUAUUCGCGCACCCCAACCUCUGGUUUGGGUUAUUUUUUUU--
S. inhibens str. PRI-2C --UUGGUCUGAUGAUGUUU-AGGAUUUCUCAACACCGUCGACCCAGAACACUG-AAUACCAAUACGCAGGU-UCACAA-UUUUCCUGGUGUUGGC-CAUAUUAUCGCGCACCCCGGCC-UAGGUCGGGGUCAUUUUUUUU--
Y. pseudotuberculosis str. B-7194 --UGCUAGUAUGAUGUAUGAAAGAUAUCUGACAACCUGCGAAUUCACUCGAAAU-CGAAUAUAUACGCAGGU-UACGU-UUUUCCUGGUGUUGGC-CAGUCUUCGCGCACCCCGGCC-UCGGUCGGGGUUAUUUUUUUU--
Y. pestis str. 42028 --UGCUAGUAUGAUGUAUGAAAGAUAUCUGACAACCUGCGAAUUCACUCGAAAU-CGAAUAUAUACGCAGGU-UACGU-UUUUCCUGGUGUUGGC-CAGUCUUCGCGCACCCCGGCC-UCGGUCGGGGUUAUUUUUUUU--
Y. ruckeri str. NVI-701 UUGUUGCUUGAUGAUGUAUCAACGAUUCCGACAACCUGCGAAGACGACCA-----GAAUUAUAUACGCAGGU-UCACGU-UUUUCCUGGUGUUGGC-CAGUAUUCGCGCACCCCGGCC-UCGGUCGGGGUUAUUUUUUUU--
Y. ruckeri str. NVI-494 UUGUUGCUUGAUGAUGUAUCAACGAUUCCGACAACCUGCGAAGACGACCA-----GAAUUAUAUACGCAGGU-UCACGU-UUUUCCUGGUGUUGGC-CAGUAUUCGCGCACCCCGGCC-UCGGUCGGGGUUAUUUUUUUU--

**B**

E. coli str. MG1655 ----GUGCGGCCUGAAAAACAGUGCUGUGCCCUUGUAACUCAUCAUAAUAAUUUACGGCGCAGCCAAGAUUUUCCUGGUGUUGGC-CAGUAUUCGCGCACCCCGGUC--UAGC-CGGGGUCAUUUUUU--
Salmonella sp. str. L-S1477 GCUAUCUUAACUGCGUGCGGCCUGAAAAAC-AGGACUGCGCCUUUG-ACAUCAUCAUAAUAA-GCACGGCGCAGCCACGAUUUCCUGGUGUUGGC-CAGUAUUCGCGCACCCCGGUC--AAAC-CGGGGUCAUUUUUUUU--
C. werkmanii str. FDAARGOS1524 ---UCUUAACUGCGUGCGGCCUGAAAAAUAGGGCUGCGCCUUUGUACACAUAUGACAUAUAAUUUACGGCGCAGCCACGAUUUCCUGGUGUUGGC-CAGUAUUCGCGCACCCCGGUC--AAGC-CGGGGUCAUUUUUUUU--
C. werkmanii str. NBRC105721 --AUCUUAACUGCGUGCGGCCUGAAAAAUAGGGCUGCGCCUUUGUACACAUAUGACAUAUAAUUUACGGCGCAGCCACGAUUUCCUGGUGUUGGC-CAGUAUUCGCGCACCCCGGUC--AAGC-CGGGGUCAUUUUUUUU--
D. parazeae str. 586 GUGUGUUGCCAGAUG-AUGGGAUUUCAAAAACGCUUGCGUAC---CAAUGACAUCUGAGAGCGAACGCAGGUGUACCCAUUUUCCUGGUGUUGGC-CAUAUUAUCGCGCACCCCGGCU-UCGGUCGGGGUCAUUUUUUUU--
Y. enterocolitica str. E701 AGUAAGUGACUGGGU--GAACGAAAGGCAGCCCAACGCACAUGCAACUUGAAGUAUGACGGGUAUUGCAGGU-UAACGAUUUCCUGGUGUUGGC-CAUAUUAUCGCGCACCCCGGCC-UCGGUCGGGGUUAUUUUUUUU--
P. mirabilis str. NRZ-54154b GUUGCUUGAUGAUGUAUGGAAUAGCUUCAUCCUAUUCGCCUAUGUAUGAUAUAUCAAAAAGCGAGUAG----AAAAGUUUCCUGGUGUUGGC-CAGUAUUCGCGCACCCCAACCUCUGGUUUGGGUUAUUUUUUUU--
Y. enterocolitica str. YE519)36/88 --UGCUUGUAUGAUGUAUGGAA-GAAUCCUAACAACCUGCCAAUUCACUCGAAAUUGCAAUUAUAUGCAGGU-UAACGAUUUCCUGGUGUUGGC-CAGUAUUCGCGCACCCCGGCC-UCGGUCGGGGUUAUUUUUUUU--





RNA

A PUBLICATION OF THE RNA SOCIETY

Function and mechanism of action of the small regulatory RNA ArcZ in *Enterobacterales*.

Quentin Dubois, Typhaine Brual, Charlotte Oriol, et al.

RNA published online June 5, 2024

P<P Published online June 5, 2024 in advance of the print journal.

Accepted Manuscript Peer-reviewed and accepted for publication but not copyedited or typeset; accepted manuscript is likely to differ from the final, published version.

Open Access Freely available online through the *RNA* Open Access option.

Creative Commons License This article, published in *RNA*, is available under a Creative Commons License (Attribution 4.0 International), as described at <http://creativecommons.org/licenses/by/4.0/>.

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).



To subscribe to *RNA* go to:
<http://rnajournal.cshlp.org/subscriptions>
