

**The rggC locus, with a frameshift mutation, is involved
in oxidative stress response by
Streptococcus thermophilus**

Annabelle Fernandez, Frédéric Borges, Brigitte Gintz, Bernard Decaris,
Nathalie Leblond-Bourget

► **To cite this version:**

Annabelle Fernandez, Frédéric Borges, Brigitte Gintz, Bernard Decaris, Nathalie Leblond-Bourget. The rggC locus, with a frameshift mutation, is involved in oxidative stress response by Streptococcus thermophilus. Archives of Microbiology, Springer Verlag, 2006, 186 (3), pp.161 - 169. <10.1007/s00203-006-0130-8>. <hal-01631247>

HAL Id: hal-01631247

<https://hal.univ-lorraine.fr/hal-01631247>

Submitted on 9 Nov 2017

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 **The *rggC* locus, with a frameshift mutation, is involved in oxidative**
2 **stress response by *Streptococcus thermophilus***

3

4 **Annabelle Fernandez, Frédéric Borges, Brigitte Gintz, Bernard Decaris and**
5 **Nathalie Leblond-Bourget***

6

7 Laboratoire de Génétique et Microbiologie, UMR INRA 1128, IFR 110, Faculté des Sciences
8 et Techniques de l'Université Henri Poincaré, Campus de Grignard, BP239, 54506
9 Vandœuvre-lès-Nancy, France

10

11 *Corresponding author: Tél: (33)3 83 68 42 10 - Fax: (33)3 83 68 44 99

12

13 email: bourget@nancy.inra.fr

14

15

16

SUMMARY

17
18 In *Streptococcus thermophilus*, the locus *rggC* contains a frameshift mutation and thus
19 consists of two ORFs, *rggC1* and *rggC2* which encode proteins exhibiting similarity with the Rgg
20 transcriptional regulator family. In this work, mutants showing a partial deletion of *rggC1* and
21 *rggC2* were constructed and their response to menadione, a superoxide generating compound,
22 was analysed. These mutants exhibited different behaviour to this oxidative stress compared to
23 the wild-type strain. Analysis of this locus among 21 strains of *S. thermophilus* showed a
24 polythymine tract length variability and a strain dependant adenine residue could be found
25 upstream of this repeat. This inter-strain polymorphism supports evidence for the hypothesis that
26 the *rggC* locus is phase variable.

27

28

29 **Key words:** *Streptococcus thermophilus*, oxidative stress response, transcriptional regulator,
30 Rgg, genetic variability, frameshift mutation

31

32

33

INTRODUCTION

34
35 The genus *Streptococcus* is composed of numerous pathogenic species and the food grade
36 species *S. thermophilus* which is used as a starter of fermentation in yogurt and cheese
37 production. During the biotransformation process, the formation of reactive oxygen species
38 (ROS) may occur and be deleterious to the cells. *S. thermophilus* is an anaerobic aerotolerant
39 organism that can grow in presence of oxygen and survive in low concentrations of superoxide
40 and hydroxyl radicals (Thibessard et al. 2001). This suggests that it has developed strategies for
41 cell detoxication and damage repair caused by ROS.

42 Recently, the genome sequences of two *S. thermophilus* strains LMG18311 and CNRZ1066
43 were completed (Bolotin et al. 2004) and putative genes involved in oxidative stress defence
44 were identified by genomic analysis (Hols et al. 2005). Genes in the disulfide-reducing pathway
45 like glutathione reductase, thioredoxin or thioredoxin reductase have been detected. These genes
46 products help to maintain the redox state of the cell and allow reducing intracellular disulfide.
47 *S. thermophilus* antioxidant defence also comprises genes encoding enzymes that detoxify
48 oxygen or ROS. Thus, the *S. thermophilus* genome encodes an H₂O-forming NADH oxidase, an
49 essential enzyme for the NAD⁺regeneration during aerobic growth (Higuchi et al. 1999). In
50 *Streptococcus mutans*, induction of H₂O-forming NADH oxidase protects cells against oxygen
51 toxicity (Higuchi 1992). *S. thermophilus* also encodes a superoxide dismutase thought to
52 catalyze superoxide radicals dismutation in O₂ and H₂O₂. Although deprived of H₂O₂-
53 detoxifying enzymes like catalase, NADH-peroxidase or alkyl hydroperoxidase, *S. thermophilus*
54 possess an H₂O₂ defence system to reduce damage induced by this compound. This system could
55 include *psaD* and *dpr* genes. *PsaD* is encoding a thiol-peroxidase and *Dpr* is an iron-binding
56 protein required for oxygen tolerance in *S. mutans* (Yamamoto et al. 2004). *Dpr* iron
57 sequestration activity is thought to prevent the formation of toxic hydroxyl radicals via the
58 Fenton reaction (Niven and Ekins 2001; Yamamoto et al. 2002; Pulliainen et al. 2003;
59 Yamamoto et al. 2004; Pulliainen et al. 2005). Valuable information on *S. thermophilus*
60 oxidative stress processes have also acquired by genetic approach using random IS mutagenesis.

61 Inactivation of several genes involved either in cell wall metabolism, exopolysaccharide
62 translocation, tRNA modification, purine metabolism, or iron metabolism have been implicated
63 in cell survival during oxidative stress (Thibessard et al. 2004).

64 Among the well characterised oxidative stress-regulators in other organisms such as OxyR,
65 PerR, sigma(B) and OhrR, controlling the peroxide response (for reviews: Hecker and Volker
66 1998; Mongkolsuk and Helmann 2002) and the superoxide stress response system SoxRS
67 (Dempfle 1996) only *perR*, is present in the *S. thermophilus* genome.

68 In pathogenic streptococci, several regulators involved in oxidative stress have been
69 described. Among these regulatory proteins is RopA, a key regulator of acid and oxidative stress
70 tolerance, genetic competence, biofilm formation and virulence in *S. mutans* (Wen et al. 2005).
71 In *Streptococcus pneumoniae*, the perception of O₂ deprivation is associated with DNA uptake
72 and controlled by CiaRH and MicAB two-component systems (Echenique et al. 2000; Echenique
73 and Trombe 2001). Both *ropA* and *micAB* genes have been identified in the *S. thermophilus*
74 genome (Bolotin et al. 2004). Additional to this list are the regulatory factors belonging to the
75 Rgg family. These proteins are only found in *Streptococcus* and few other low-GC Gram-
76 positive genera, such as *Listeria*, *Lactobacillus* and *Lactococcus*. In *S. pyogenes*, Rgg interacts
77 with regulatory networks to co-regulate virulence factor expression (Chaussee et al. 2001) and
78 has also been implicated in oxidative and thermic stress responses as well as in the synthesis or
79 stability of growth phase-regulated proteins associated with amino-acid metabolism (Chaussee et
80 al. 2004).

81 In order to characterise *S. thermophilus* oxidative stress tolerance molecular basis, we
82 previously used a random IS mutagenesis approach to produce mutants and screened them for
83 their potential impaired tolerance to menadione (a superoxide generating compound). This
84 strategy led us to identify 18 *S. thermophilus* CNRZ368 mutants (Thibessard et al. 2004). One of
85 which is *rggC*. This gene is known to contain a frameshift mutation from the *S. thermophilus*
86 genome sequence. This work is focused on *rggC* and attempts to determine how such a locus
87 inactivated by a frameshift mutation is involved in oxidative stress tolerance. The *rggC* locus

88 analysis was performed by engineering partial deletions of this locus and examining the ability
89 of mutants to survive in the presence of menadione. Our results support an active role for RggC
90 in *S. thermophilus* oxidative stress response.

91

92

MATERIALS AND METHODS

93 1. Bacterial strains growth conditions

94 The bacterial strains used in this study are listed in Table 1. Streptococcal strains including
95 wild-type and mutant strains were cultured either in M17 (Terzaghy and Sandine 1975), or TPPY
96 (a Tryptose, Proteose Peptone, Yeast extract containing medium) (Bracquart 1981) or milk
97 mediums, at 30°C or 42°C, depending on the experiments requirement. When appropriate, these
98 media were supplemented with antibiotics at the following concentrations: erythromycin at
99 2 µg.ml⁻¹, and chloramphenicol at 3 µg.ml⁻¹.

100 *E. coli* EC101 was used for cloning experiments and transformed by electroporation with
101 various plasmids as indicated below. *E. coli* strains were grown on Luria-Bertani (LB) medium
102 (Sambrook et al. 1989) and the following selective antibiotics were added, when needed:
103 erythromycin (150 µg.ml⁻¹), chloramphenicol (15 µg.ml⁻¹) or ampicillin (100 µg.ml⁻¹).

104

105 2. DNA manipulations

106 Isolation of genomic or plasmid DNA, agarose gel electrophoresis, restriction enzyme
107 digestions, Southern hybridisation, PCR, DNA ligation and *E. coli* transformation were
108 performed as described in Sambrook et al. (1989). Sequencing was carried out with the DNA
109 sequencing kit (Applied Biosystems) on an ABI Prism 377 Genetic Analyser (Applied
110 Biosystems). Sequence data were analysed with Blast, ProDom and TIGRfam Model software.
111 The *rggC1* and *rggC2* DNA sequences have previously been deposited in GenBank under
112 accession numbers from AY598750. The *orf2* and *rsp* DNA sequences obtained in this work
113 were deposited under accession number AM087201.

114

115

116 **3. RNA manipulations**

117 RNA was extracted from *S. thermophilus* CNRZ368 overnight (OD₆₀₀ = 1.6) cultures with a
118 single guanidinium thiocyanate step and phenol-chloroform extraction by using the Tri-Reagent
119 (Sigma). Northern blot hybridisation analysis, templates of RNA samples (30 µg) were
120 electrophoresed into a denaturing (i.e. formaldehyde) 1.1 % agarose gel and transferred to a
121 HybondN⁺ (Amersham pharmacia biotech) according to the manufacturer's instructions.
122 Digoxigenin-labelled transcripts complementary to *rggC₂* were used as a probe. These *rggC₂*
123 transcripts were synthesised in vitro using the T3 promoter RNA polymerase system carried by
124 the pBKS-*rggC₂* vector. The latter was constructed by cloning a 533 bp internal fragment of
125 *rggC₂* (nucleotides 932 to 1464) in the pBKS vector. Hybridisation and detection procedures
126 were performed according to manufacturer's instructions (Roche Molecular Biochemicals). RT-
127 PCR experiments were carried out as previously described (Fernandez et al. 2004). Primers used
128 in PCRs were purchased from MWG Biotech AG (Ebersberg) and their sequences were
129 5' -AGGTGACTAGAATGAAATATGG-3' for P₁, 5' -TCCATTTGGCTAATTAATAGC-3' for
130 P₂, 5' -TATTGCTGCCACGAAATTGGT-3' for P₃ and
131 5' -CACCGCCACAAGAATGACTT-3' for P₄. Primers used in *rggC* locus sequencing were
132 5' -AGGAGGTGACTAGAATGAAATATGG-3' for V1 and
133 5' -TCCATTTGGCTAATTAATAGC-3' for V2.

134

135 **4. Constructions of *ΔrggC₁*, *ΔrggC₂*, *orf2* and *rggC_{8T}* mutants**

136 The *ΔrggC₁* gene contained a deletion of 43 % of the CDS. A stop codon (TAA) was inserted
137 between the potential RBS of *rggC₁* and the start codon of *rggC₂* to prevent using this RBS
138 sequence for *RggC₂* translation. The 5' and 3' ends of the *rggC₁* gene were independently
139 amplified with the following sets of primers: 5' -CTTGCTTTGAATTCAGGCT-3' (named I₁)
140 and 5' -TAAAGATAAAGCTTAAGATTCTC-3' for the 5' region;
141 5' -GGCTGATTTAAGCTTCAGCAAATT-3' and 5' -ATCATATGAATTCCTTAGCAT-3',

142 (named I₂), for the 3' region (engineered restriction sites are underlined, stop codon is in
143 boldface type). Amplified fragments were digested with *Hind*III and ligated to each other. The
144 resulting fragment was amplified, digested with *Eco*RI and cloned into pGh9 vector to produce
145 pGh9 Δ *rggC*₁ plasmid. The same procedure was used to produce pGh9 Δ *rggC*₂ and pGh9 Δ *orf2*_{Z1,2}
146 plasmids. pGh9 Δ *rggC*₂ carried a 44 % deletion of *rggC*₂ and was constructed using the primers:
147 5' -TTATAAGAATTCAGGAGGTG-3' and 5' -TGATTTAAAGCTTCTATTGTAT-3' to
148 amplify the 5' region; 5' -AAATGTAAAGCTTCCACGAAA-3' and
149 5' -AAATCCAGAATTCGATTATCC-3' for the 3' region.

150 The construction of the pGh9 Δ *orf2*_{Z1,2} plasmid containing 2 stop codons (TAA, TAA) in the 2
151 first codons of *orf2* was carried out using the following set of primers:
152 5' -TCGAAGAATTCGTTTACG-3' and 5' -TCACCTGAAGCTTATTACATGGT-3' for the 5'
153 region; 5' -AACCATGTAATAAGCTTCAGGTGAT-3' and
154 5' -TTTGATGAATTCTAGTATATGT-3' for the 3' region.

155 The pGh9 Δ *rggC*_{8T} plasmid, carrying a thymine point deletion allowing a *RggC*₁-*C*₂ fusion, was
156 constructed. To generate this mutation, the 5' and 3' regions of *rgg* were amplified with the
157 following sets of primers: I₁ and 5' -CTAACGCTAAAAAAAATTTGCTG-3' for the 5' region;
158 5' -CAAATTTTTTTTAGCGTTAGATGC-3' and I₂ for the 3' region. These 2 products were
159 mixed together, used as templates in a PCR reaction to generate a full length fragment and
160 cloned into pGh9 vector to produce the pGh9 Δ *rggC*_{8T} plasmid. This recombinant plasmid DNA
161 was extracted and used to transform *S. thermophilus* CNRZ368 cells as previously described
162 by (O'Sullivan and Fitzgerald 1999). Allelic gene replacements were done as previously
163 described (Biswas et al. 1993) with the following modifications: the medium used for
164 *S. thermophilus* growth was M17, 3 cycles of growth were carried out with two steps, one at
165 30°C (permissive temperature) and another one at 42°C (restrictive temperature). The
166 chromosomal structure of mutant alleles was checked by PCR, southern blot analysis and
167 sequencing.

168

169 **6. Menadione assay**

170 *S. thermophilus* and its mutant derivatives were grown overnight in milk medium and diluted
171 100 times into TPPY liquid medium. Log-phase cultures ($OD_{600} = 0.6$) were exposed for 3h to
172 menadione concentrations (0 to 72 mM). Aliquots of the culture were then diluted in TPPY and
173 appropriate dilutions plated on TPPY agar plates. The plates were stocked in jars containing a
174 GENbox anaer system (bioMérieux) providing hypoxic conditions. Colony-forming units (CFU)
175 were counted after 20 h of incubation and the percentage of surviving CFUs was determined by
176 comparing viability with or without treatment. All experiments were performed independently at
177 least three times.

178

179 **RESULTS**

180 **1. Characterisation of the disrupted loci in *osrC* genome**

181 In a previous study, the *osrC* mutant of *S. thermophilus* CNRZ368 was characterised in its
182 resistance to menadione, a superoxide-generating agent (Fernandez et al. 2004). In this mutant,
183 pGhost9::ISS1 was inserted in *rggC2* region. The *rggC1* ORF encodes a predicted 65 aa protein
184 exhibiting 77% identity to the C-terminal region of *S. mutans* Rgg. The second ORF, *rggC2*
185 encodes a putative 232 aa protein sharing 54 % identity with the N-terminal region of *S. mutans*
186 Rgg. These two ORFs could initially have constituted a single ORF, having undergone a
187 frameshift mutation (Fig.1A). This mutational event would have occurred between positions 886
188 to 894, where a 9-mer (TTT TTT TTT) known to be genetically unstable is observed. Ribosome-
189 binding site (RBS) consensus sequence (AAAGGAGG in Gram-positive bacteria) scan revealed
190 a potential RBS (AGGAGG) located 8 bp upstream *rggC1*. No canonical RBS sequence-
191 preceding *rggC2* was detected upstream of the first putative start codon or within the CDS of
192 *rggC1*. These results suggest that only RggC1 or a full fusion protein RggC1-2, could be produced,
193 indicating a translational or transcriptional frameshifting.

194 Since Rgg proteins often regulate adjacent genes (Sulavik and Clewell 1996; Lyon et al.
195 1998; Chaussee et al. 1999; Qi et al. 1999; Neely et al. 2003), the ORFs localised downstream of

196 the *rggC* locus were analysed. *RggC₂* is followed by a small ORF, *orf₂* that encodes a putative 35
197 aa protein with no significant homology to any database sequences. *orf₂* is preceded by a
198 canonical RBS with the sequence, AAAGGAGG suggesting it was expressed. Downstream *orf₂*,
199 there is an *rsp* gene (for radical SAM protein) that encodes a putative protein sharing 29 %
200 identity (46% similarity) with a possible Fe-S oxidoreductase of *Thermoanaerobacter*
201 *tengcongensis* belonging to the radical SAM protein family (Sofia et al. 2001). A RBS sequence
202 (GGAGG) was also found upstream *rsp*.

203 The transcriptional signals scan analysis allowed us to identify a potential promoter
204 (TATGATATAAT), located 61 bp upstream of *rggC₁*, and constitute an extended -10 region
205 (Voskuil and Chambliss 1998). Moreover, a putative rho-independant terminator consisting of a
206 perfect inverted repeat of 10 bp followed by a stretch of T's, was visualized 9 bp downstream of
207 the *orf₂* stop codon. A potential promoter located upstream *rsp* was identified including a -10
208 (TGTAAT) and a -35 (TTTAGA) region with a 15 bp spacer.

209

210 **2. Transcriptional analysis of the *rggC* locus**

211 Northern blot analysis using *rggC₂* antisense-RNA (nucleotides 932 to 1464) as a probe was
212 performed on RNA extracted from stationary phase cells. One transcript of approximately 1200
213 nucleotides was detected (Fig. 1B). Since the *rggC₂* gene size is 0.7 kb, this transcript observation
214 is compatible with the polycistronic transcription of *rggC₂* with the neighbouring ORFs.

215 To better characterise this transcript, RT-PCR analyses were carried out. cDNA generated
216 from *S. thermophilus* CNRZ368 RNA were used in PCRs experiments using primers P₁/P₂ and
217 P₃/P₄ localized in *rggC₁/rggC₂* and *rggC₂/orf₂* respectively (Fig. 1C). In both cases, a PCR
218 product of the expected size (310 bp and 265 bp respectively) was amplified suggesting the
219 existence of at least one transcript including *rggC₁*, *rggC₂* and *orf₂*. The transcript size was in
220 agreement with the transcription signals visualized on the sequence.

221

222

223 **3. Involvement of the *rggC* locus in the oxidative stress defence**

224 Since *rggC* locus exhibits operon structure, the resistance phenotype could be either the result
225 of the inactivation of *rggC₂* gene or/and the result of polar effects on the *rggC₁* and *orf2*
226 neighbouring genes. In this study, non-polar Δ *rggC₁*, Δ *rggC₂* and *orf2* mutants were constructed
227 via allelic replacement and their survival in presence of menadione was studied (see materials
228 and methods). Compared to the reference strain, the Δ *rggC₁* mutant displayed a 2-fold higher
229 resistance following exposure to menadione at concentrations of 27 mM and 36 mM (Fig. 2A).
230 The Δ *rggC₂* mutant exhibited a 2-fold decreased of surviving cells at a menadione concentration
231 of 9 mM, 18 mM and 27 mM (Fig. 2B). Thus, a decreased survival was observed for the Δ *rggC₂*
232 mutant, while the survival of the Δ *rggC₁* strain was enhanced indicating the involvement of both
233 of these ORFs in *S. thermophilus* oxidative stress response. In contrast, *orf2* mutant survival was
234 the same as the reference strain (data not shown) suggesting that *orf2* was not involved in this
235 pathway.

236

237 **4. Intraspecific polymorphism of *rggC* potential shift site**

238 The analysis of the DNA sequence of the *rggC* locus predicted 2 ORFs: *rggC₁* and *rggC₂*. A
239 sequence of 9 T's was observed at the 3' end of *rggC₁*. Since such mononucleotide repeats are
240 known to exhibit high frequencies of nucleotide deletion or insertion events, the potential
241 polymorphism of the *rggC* shift site was investigated among 21 strains *S. thermophilus*. The
242 *rggC₁* gene was shown to be present in all 21 studied strains, using PCR amplification with
243 primers V1 and V2 designed to bind the 5' part of *rggC₁* and the central part of *rggC₂*,
244 respectively. The frameshift mutation of *rggC₁* was analysed by sequencing the PCR products
245 with primer V1. Four different sequences were found varying from 7 to 9 T's (Table 2). A
246 deletion of one A upstream of the T tract in *rggC* from CNRZ1402, CNRZ455 and LMD9 strains
247 was also observed. The strain ATCC1958 is characterised by the presence of 8 instead of 9 Ts.

248 This mutation would result in formation of an in-frame ORF, allowing potentially full-length
249 RggC protein translation whereas in all the other cases a RggC₁ protein was expected.

250

251 **5. Involvement of RggC in oxidative stress defence of *S. thermophilus* CNRZ368**

252 To determine the involvement in oxidative stress defence of a full-length RggC protein, an
253 additional *rggC_{8T}* mutant was constructed via allelic replacement. This mutant from CNRZ368
254 contains an eight T chromosomal sequence. Its survival in the presence of different
255 concentrations of menadione was determined and showed to be reduced at a menadione
256 concentration of 18 mM, 27 mM and 36 mM compared to wild-type confirming a role for *rggC*
257 in *S. thermophilus* CNRZ368 oxidative stress response (Fig. 3).

258

259

DISCUSSION

260 Among the fully sequenced streptococcal genomes, that of *S. thermophilus* exhibits the
261 highest proportion of pseudogenes. Indeed, 10% of the *S. thermophilus* genes are characterised
262 by frameshift, nonsense mutation, deletion or truncation and are therefore considered as non
263 functional (Bolotin et al. 2004). Here, we examined *rggC*, considered in *S. thermophilus* as a
264 pseudogene, and we demonstrated its role in *S. thermophilus* oxidative stress response.

265 In *osrC* mutant, the disruption of *rggC₂*, with pGh9:ISS1, confers to the cells a resistance to
266 menadione. This insertion mutant shows a polar mutation affecting the transcript stability of the
267 *rggC* locus and of the neighbouring *rsp* gene.

268 To understand this phenomenon, deletions of *rggC₁* and *rggC₂* were undertaken and
269 demonstrated that both ORFs were involved in oxidative stress response. However, these 2
270 mutants displayed 2 distinct phenotypes: one being resistant and the other sensitive to oxidative
271 stress. Two alternative hypotheses can explain this. Firstly, the 2 proteins RggC₁ and RggC₂
272 could be synthesised and would be characterised by different activities. Nevertheless, the lack of
273 a RBS sequence-preceding *rggC₂* is a strong argument in disfavour of this hypothesis. Secondly,

274 the $\Delta rggC_2$ phenotype could result from the synthesis of a full RggC protein by frameshifting as
275 observed in translational or transcriptional slippage (for review : Baranov et al. 2005).

276 Taking all these data into account, both RggC₁ and RggC proteins are likely to be synthesised
277 in the wild type strain.

278 In CNRZ368 strain, a stretch of 9 T residues could allow a transcriptional slippage, and
279 consequently lead to the RggC fusion protein by translation of a subpopulation of *rggC* transcript
280 in the alternate reading frame. Alternatively, since RNA-DNA hybrids of 9 are described as
281 genetically unstable they could provide RNA polymerase slippage by re-alignment of the nascent
282 RNA chain and its DNA template (Wagner et al. 1990). Few examples of transcriptional
283 slippage are identified in prokaryotes. In *Thermus thermophilus*, τ and γ subunit proteins of DNA
284 polymerase III are synthesized from *dnaX* gene by this mechanism. This ORF shows 9 adenine
285 residues allowing the slippage of RNA polymerase and producing many mRNAs with additional
286 or subtractive adenine residues (Larsen et al. 2000). In *Shigella flexneri*, the *mixE* gene, encoding
287 a transcription activator, consists of 2 overlapping ORFs. A track of 9 thymine residues is
288 observed at the site of overlapping. Transcriptional incorporation of one additional nucleotide
289 allows synthesis of the functional MixE protein (Penno et al. 2005).

290 The putative RggC protein belongs to the family of Rgg-like transcriptional regulators. RggC₁
291 displayed the typical N-terminal helix-turn-helix motif suggestive of a DNA-binding function.
292 Among characterised Rgg-like proteins, *rgg* from *Streptococcus gordonii*, *rgg* and *ropB* from
293 *S. pyogenes* are located immediately adjacent to their regulatory target sequence (Sulavik and
294 Clewell 1996; Lyon et al. 1998; Chaussee et al. 1999; Qi et al. 1999; Neely et al. 2003). Thus,
295 the transcriptional expression of *orf2* and *rsp* ORFs from *S. thermophilus* CNRZ368 was
296 analysed by quantitative RT-PCR in $\Delta rggC_1$, $\Delta rggC_2$ and wild-type genetic contexts (data not
297 shown). Partial deletion of *rggC* locus did not affect the expression of *orf2* and *rsp* ORFs
298 compared with the level seen in the wild type CNRZ368 strain. However, *S. thermophilus*
299 genome contains 7 *rgg* genes (Bolotin et al. 2004). It is therefore possible that one of the Rgg
300 proteins replace RggC function in the defective strain. Moreover, it still is possible that RggC

301 directly or indirectly regulates the transcription of other genes since deletion of *rgg* from
302 *S. pyogenes* has been demonstrated to have genome wide effects on virulence-associated
303 transcription (Chaussee et al. 2001; Chaussee et al. 2002).

304 Analysis of 21 *S. thermophilus* *rgg* sequences revealed a polymorphism of the 9 T residue
305 region. Repeats of a single or several nucleotides, named microsatellite DNA, are involved in
306 numerous phase variation events that could be defined as a high-frequency and reversible
307 mechanisms of gene-switching (for review: (Hallet 2001). This inter-strain length polymorphism
308 supports evidence for the phase variability of the *rggC* locus. Some cases of phase variation have
309 been identified in the *Streptococcus* genus, allowing the control of transcriptional or translational
310 expression (Lukomski et al. 2001; Puopolo and Madoff 2003). A phase variation event at the
311 *rggC* locus could influence a potential programmed frameshifting by inactivation.

312 In 20/21 cases, a +1 or -1 frameshift mutation was observed resulting in a 'gene off'
313 configuration. If the full-length *rggC* ORF was an ancestral allele, it is possible that alleles
314 showing this frameshift mutation may have been selected for a growth advantage associated with
315 the loss of *rggC* function. The menadione sensitivity phenotype observed for the *rggC_{8T}* mutant
316 (whose *rggC* phase is restored) argued in the favour of that hypothesis.

317
318 The results of this study indicate the involvement of the *rggC* locus in *S. thermophilus*
319 CNRZ368 oxidative stress response. Accordingly, the Rgg transcriptional regulatory protein
320 from *S. pyogenes* that coordinates virulence factor synthesis and catabolic activity (Chaussee et
321 al. 2003) was also demonstrated to be involved in thermal and in oxidative stress response
322 (Chaussee et al. 2004). The *S. pyogenes* *rgg* mutant strain was more susceptible than wild-type to
323 paraquat (a hydroxyl ion radicals generating compound) and was characterised by an elevated
324 level of the oxidoreductases AhpC and Nox1 (Chaussee et al. 2004). All these data suggest a role
325 for the Rgg family members in oxidative stress response, and more generally that these proteins
326 may be important in the *streptococcus* adaptation to environmental changes.

327

328

ACKNOWLEDGEMENTS

329 A.F. was supported by grants from the Ministère de la recherche. F.B. was supported by a
330 grant from the Institut National de la Recherche Agronomique. We are grateful to Arran Johnson
331 for his advice regarding the English formulation of the manuscript.

332

REFERENCES

- 333
- 334 Baranov PV, Hammer AW, Zhou J, Gesteland RF, Atkins JF (2005) Transcriptional slippage in bacteria:
335 distribution in sequenced genomes and utilization in IS element gene expression. *Genome Biol* 6:R25
- 336 Biswas I, Gruss A, Ehrlich SD, Maguin E (1993) High-efficiency gene inactivation and replacement system for
337 gram-positive bacteria. *J Bacteriol* 175:3628-3635
- 338 Bolotin A et al. (2004) Complete sequence and comparative genome analysis of the dairy bacterium *Streptococcus*
339 *thermophilus*. *Nat Biotechnol* 22:1554-1558
- 340 Bracquart P (1981) An agar medium for the differential enumeration of *Streptococcus thermophilus* and
341 *Lactobacillus bulgaricus* in yoghurt. *Journal of Applied Bacteriology* 51:303-305
- 342 Chaussee MA, Callegari EA, Chaussee MS (2004) Rgg regulates growth phase-dependent expression of proteins
343 associated with secondary metabolism and stress in *Streptococcus pyogenes*. *J Bacteriol* 186:7091-7099
- 344 Chaussee MS, Ajdic D, Ferretti JJ (1999) The rgg gene of *Streptococcus pyogenes* NZ131 positively influences
345 extracellular SPE B production. *Infect Immun* 67:1715-1722
- 346 Chaussee MS, Somerville GA, Reitzer L, Musser JM (2003) Rgg coordinates virulence factor synthesis and
347 metabolism in *Streptococcus pyogenes*. *J Bacteriol* 185:6016-6024
- 348 Chaussee MS et al. (2002) Rgg influences the expression of multiple regulatory loci to coregulate virulence factor
349 expression in *Streptococcus pyogenes*. *Infect Immun* 70:762-770
- 350 Chaussee MS, Watson RO, Smoot JC, Musser JM (2001) Identification of Rgg-regulated exoproteins of
351 *Streptococcus pyogenes*. *Infect Immun* 69:822-831
- 352 Demple B (1996) Redox signaling and gene control in the *Escherichia coli* *soxRS* oxidative stress regulon--a review.
353 *Gene* 179:53-57
- 354 Echenique JR, Chapuy-Regaud S, Trombe MC (2000) Competence regulation by oxygen in *Streptococcus*
355 *pneumoniae*: involvement of *ciaRH* and *comCDE*. *Mol Microbiol* 36:688-696
- 356 Echenique JR, Trombe MC (2001) Competence modulation by the NADH oxidase of *Streptococcus pneumoniae*
357 involves signal transduction. *J Bacteriol* 183:768-772
- 358 Fernandez A, Thibessard A, Borges F, Gintz B, Decaris B, Leblond-Bourget N (2004) Characterization of oxidative
359 stress-resistant mutants of *Streptococcus thermophilus* CNRZ368. *Arch Microbiol* 182:364-372
- 360 Hallet B (2001) Playing Dr Jekyll and Mr Hyde: combined mechanisms of phase variation in bacteria. *Curr Opin*
361 *Microbiol* 4:570-581
- 362 Hecker M, Volker U (1998) Non-specific, general and multiple stress resistance of growth-restricted *Bacillus*
363 *subtilis* cells by the expression of the *sigmaB* regulon. *Mol Microbiol* 29:1129-1136

364 Higuchi M (1992) Reduced nicotinamide adenine dinucleotide oxidase involvement in defense against oxygen
365 toxicity of *Streptococcus mutans*. Oral Microbiol Immunol 7:309-314

366 Higuchi M et al. (1999) Functions of two types of NADH oxidases in energy metabolism and oxidative stress of
367 *Streptococcus mutans*. J Bacteriol 181:5940-5947

368 Hols P et al. (2005) New insights in the molecular biology and physiology of *Streptococcus thermophilus* revealed
369 by comparative genomics. FEMS Microbiol Rev 29:435-463

370 Larsen B, Wills NM, Nelson C, Atkins JF, Gesteland RF (2000) Nonlinearity in genetic decoding: homologous
371 DNA replicase genes use alternatives of transcriptional slippage or translational frameshifting. Proc Natl
372 Acad Sci U S A 97:1683-1688

373 Leenhouts K (1995) Integration strategies and vectors. Dev Biol Stand 85:523-530

374 Lukomski S et al. (2001) Identification and characterization of a second extracellular collagen-like protein made by
375 group A *Streptococcus*: control of production at the level of translation. Infect Immun 69:1729-1738

376 Lyon WR, Gibson CM, Caparon MG (1998) A role for trigger factor and an rgg-like regulator in the transcription,
377 secretion and processing of the cysteine proteinase of *Streptococcus pyogenes*. Embo J 17:6263-6275

378 Maguin E, Prévost H, Ehrlich SD, Gruss A (1996) Efficient insertional mutagenesis in *Lactococci* and other Gram-
379 positive bacteria. Journal of Bacteriology 178:931-935

380 Mongkolsuk S, Helmann JD (2002) Regulation of inducible peroxide stress responses. Mol Microbiol 45:9-15

381 Neely MN, Lyon WR, Runft DL, Caparon M (2003) Role of RopB in growth phase expression of the SpeB cysteine
382 protease of *Streptococcus pyogenes*. J Bacteriol 185:5166-5174

383 Niven DF, Ekins A (2001) Iron content of *Streptococcus suis* and evidence for a dpr homologue. Can J Microbiol
384 47:412-416

385 O'Sullivan TF, Fitzgerald GF (1999) Electrotransformation of industrial strains of *Streptococcus thermophilus*. J
386 Appl Microbiol 86:275-283

387 Penno C, Sansonetti P, Parsot C (2005) Frameshifting by transcriptional slippage is involved in production of MxiE,
388 the transcription activator regulated by the activity of the type III secretion apparatus in *Shigella flexneri*.
389 Mol Microbiol 56:204-214

390 Pulliainen AT, Haataja S, Kahkonen S, Finne J (2003) Molecular basis of H₂O₂ resistance mediated by
391 Streptococcal Dpr. Demonstration of the functional involvement of the putative ferroxidase center by site-
392 directed mutagenesis in *Streptococcus suis*. J Biol Chem 278:7996-8005

393 Pulliainen AT, Kauko A, Haataja S, Papageorgiou AC, Finne J (2005) Dps/Dpr ferritin-like protein: insights into the
394 mechanism of iron incorporation and evidence for a central role in cellular iron homeostasis in
395 *Streptococcus suis*. Mol Microbiol 57:1086-1100

396 Puopolo KM, Madoff LC (2003) Upstream short sequence repeats regulate expression of the alpha C protein of
397 group B Streptococcus. *Mol Microbiol* 50:977-991

398 Qi F, Chen P, Caufield PW (1999) Purification of mutacin III from group III Streptococcus mutans UA787 and
399 genetic analyses of mutacin III biosynthesis genes. *Appl Environ Microbiol* 65:3880-3887

400 Sambrook J, Fritsch EF, Maniatis T (eds) (1989) *Molecular cloning: a Laboratory Manual*. Cold Spring Harbor
401 Laboratory, New York

402 Sofia HJ, Chen G, Hetzler BG, Reyes-Spindola JF, Miller NE (2001) Radical SAM, a novel protein superfamily
403 linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional
404 characterization using new analysis and information visualization methods. *Nucleic Acids Res* 29:1097-
405 1106

406 Sulavik MC, Clewell DB (1996) Rgg is a positive transcriptional regulator of the *Streptococcus gordonii* *gffG* gene.
407 *J Bacteriol* 178:5826-5830

408 Terzaghy BE, Sandine WE (1975) Improved medium for lactic streptococci. *Current Microbiology* 7:245-250

409 Thibessard A, Borges F, Fernandez A, Gintz B, Decaris B, Leblond-Bourget N (2004) Identification of
410 *Streptococcus thermophilus* CNRZ368 genes involved in defense against superoxide stress. *Appl Environ*
411 *Microbiol* 70:2220-2229

412 Thibessard A, Fernandez A, Gintz B, Leblond-Bourget N, Decaris B (2001) Hydrogen peroxide effects on
413 *Streptococcus thermophilus* CNRZ368 cell viability. *Res Microbiol* 152:593-596

414 Voskuil MI, Chambliss GH (1998) The -16 region of *Bacillus subtilis* and other gram-positive bacterial promoters.
415 *Nucleic Acids Res* 26:3584-3590

416 Wagner LA, Weiss RB, Driscoll R, Dunn DS, Gesteland RF (1990) Transcriptional slippage occurs during
417 elongation at runs of adenine or thymine in *Escherichia coli*. *Nucleic Acids Res* 18:3529-3535

418 Wen ZT, Suntharaligham P, Cvitkovitch DG, Burne RA (2005) Trigger factor in *Streptococcus mutans* is involved
419 in stress tolerance, competence development, and biofilm formation. *Infect Immun* 73:219-225

420 Yamamoto Y, Fukui K, Koujin N, Ohya H, Kimura K, Kamio Y (2004) Regulation of the intracellular free iron pool
421 by Dpr provides oxygen tolerance to *Streptococcus mutans*. *J Bacteriol* 186:5997-6002

422 Yamamoto Y, Poole LB, Hantgan RR, Kamio Y (2002) An iron-binding protein, Dpr, from *Streptococcus mutans*
423 prevents iron-dependent hydroxyl radical formation in vitro. *J Bacteriol* 184:2931-2939

424

425

Strains and plasmids	Relevant feature(s)	Source and/or citation
<i>S. thermophilus</i>		
CNRZ007	Wild-type strain	CNRZ, INRA strain collection
CNRZ302	Wild-type strain	CNRZ, INRA strain collection
CNRZ307	Wild-type strain	CNRZ, INRA strain collection
CNRZ308	Wild-type strain	CNRZ, INRA strain collection
CNRZ368	Wild-type strain	CNRZ, INRA strain collection
CNRZ385	Wild-type strain	CNRZ, INRA strain collection
CNRZ388	Wild-type strain	CNRZ, INRA strain collection
CNRZ391	Wild-type strain	CNRZ, INRA strain collection
CNRZ445	Wild-type strain	CNRZ, INRA strain collection
CNRZ455	Wild-type strain	CNRZ, INRA strain collection
CNRZ702	Wild-type strain	CNRZ, INRA strain collection
CNRZ1066	Wild-type strain	CNRZ, INRA strain collection
CNRZ1068	Wild-type strain	CNRZ, INRA strain collection
CNRZ1402	Wild-type strain	CNRZ, INRA strain collection
ATCC19258	Wild-type strain	American Type Culture Collection
ATCC19987	Wild-type strain	American Type Culture Collection
LMG18311	Wild-type strain	BCCM/LMG strain collection
A054	Wild-type strain	Lab collection
IP6756	Wild-type strain	Institut Pasteur strain collection
IP6757	Wild-type strain	Institut Pasteur strain collection
<i>osrC</i>	Em ^R <i>S. thermophilus</i> CNRZ368 containing integrated pGh9::ISS1 into <i>rggC</i> ₂	((Fernandez et al. 2004)
<i>ΔrggC</i> ₁	CNRZ368 strain carrying a deletion from nt 50 to 127 of <i>rggC</i> ₁	This study
<i>ΔrggC</i> ₂	CNRZ368 strain carrying a deletion from nt 318 to 619 of <i>rggC</i> ₂	This study
<i>orf2</i>	CNRZ368 strain carrying 2 stop codons in the first codon of <i>orf2</i>	This study
<i>rggC</i> _{8T}	CNRZ368 strain presenting an in-frame <i>rggC</i> gene	This study
<i>E. coli</i>		
EC101	TG1 derived strain [<i>supE hsd-5 thi Δ(lac-proAB)</i> F' (<i>traD6proAB⁺ lacI^qlacZΔM15</i>)] containing a chromosomal copy of the pWV01 <i>repA</i> gene	(Leenhouts 1995)
Plasmid		
pBKS	Amp ^R	Stratagene
pGhost9	Em ^R , thermosensitive replication origin from pVE6002	(Maguin et al. 1996)
pGh9 <i>ΔrggC</i> ₁	pGh9 plasmid carrying <i>rggC</i> ₁ deleted from position 50 to 127	This study
pGh9 <i>ΔrggC</i> ₂	pGh9 plasmid carrying <i>rggC</i> ₂ deleted from position 318 to 619	This study
pGh9 <i>orf2</i> _{Z1,2}	pGh9 plasmid carrying <i>orf2</i> with 2 stop codons inserted in the first and the second codon	This study
pGh9 <i>rggC</i> _{8T}	pGh9 plasmid carrying an in-frame <i>rggC</i> gene	This study

427 **Table 2. Analysis of the *rggC* locus in 21 strains of *S. thermophilus***

Sequence	Strain name	Mutation
AGC AAA TTT TTT TTT AGC	A054, ATCC19987, CNRZ007, CNRZ302, CNRZ307, CNRZ368, CNRZ385, CNRZ388, CNRZ391, CNRZ702, LMG18311, IP6756	+1 FS
AGC AAA _TT TTT TTT AGC	ATCC19258	/
AGC AA _ _TT TTT TTT AGC	CNRZ308, CNRZ445, CNRZ1066, CNRZ1068, IP6757	-1 FS
AGC AA _ _T TTT TTT AGC	CNRZ455, CNRZ1402, LMD9 ^a	+1 FS

428

429 ^a according to sequence available <http://www.jgi.doe.gov>.

430

FIGURE LEGENDS

431
432 **Fig. 1** Organisation of the *rggC* locus of *S. thermophilus* CNRZ368. (A) The solid line
433 represents the *S. thermophilus* genomic DNA. ORFs are indicated as open arrows. The broken
434 arrow and the hairpin loop indicate the putative promoter and rho-independant terminator
435 respectively. Putative RBS localisation is indicated by asterisks. Primers used for RT-PCR
436 experiments are represented by small arrows under ORFs. The expected size of the P₁-P₂ and P₃-
437 P₄ PCR fragments is 310 bp and 265 bp, respectively. ARN probe localisation is indicated by the
438 black box. (B) Northern blot analysis of *rggC*₂ transcripts. Left: gel electrophoresis staining with
439 ethidium bromide: lane 1 is RNA size marker and lane 2 is RNA (30µg) total extracted from
440 *S. thermophilus* CNRZ368. 16S and 23S indicate the migrating positions of 16S and 23S rRNA,
441 respectively. Right: result of the hybridisation with the *rggC*₂ RNA probe. (C) Gel
442 electrophoresis showing the RT-PCR products obtained using the primers P₁-P₂ and P₃-P₄. *Pst*I-
443 digested λ DNA used as a DNA fragment size marker is indicated by λ *Pst*I.

444
445 **Fig. 2** Survival of Δ *rggC*₁ and Δ *rggC*₂ mutants and the reference strain with increasing
446 menadione concentrations. The graphs point the average values obtained from experiments
447 performed at least in triplicate. The error bars represent standard deviations.

448
449 **Fig. 3** Survival of *rggC*₈₇ mutant and the reference strain with increasing menadione
450 concentrations. The graphs point the average values obtained from experiments performed at
451 least in triplicate. The error bars represent standard deviations

452

453 **Fig. 1**

454

455

A

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

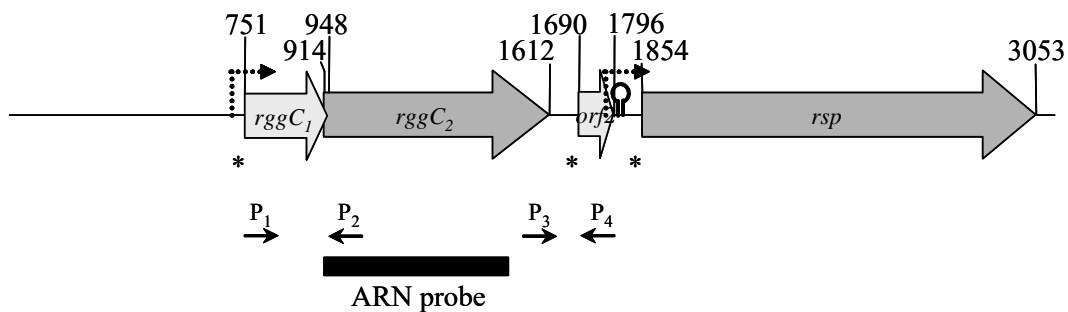
477

478

479

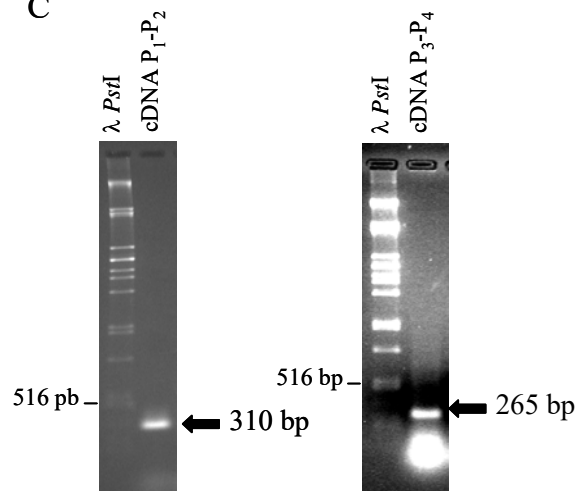
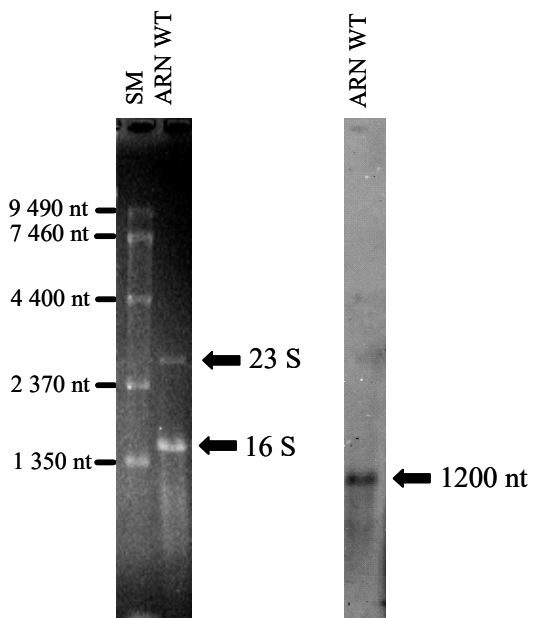
480

481



B

C



482 Fig. 2

483

484 A

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

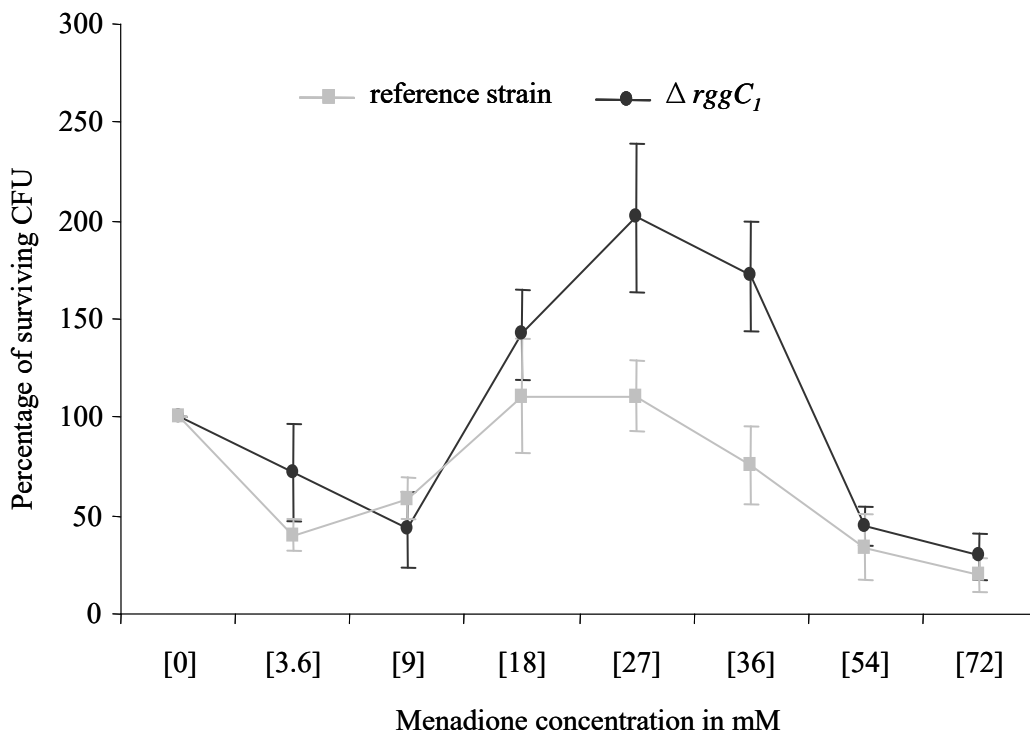
510

511

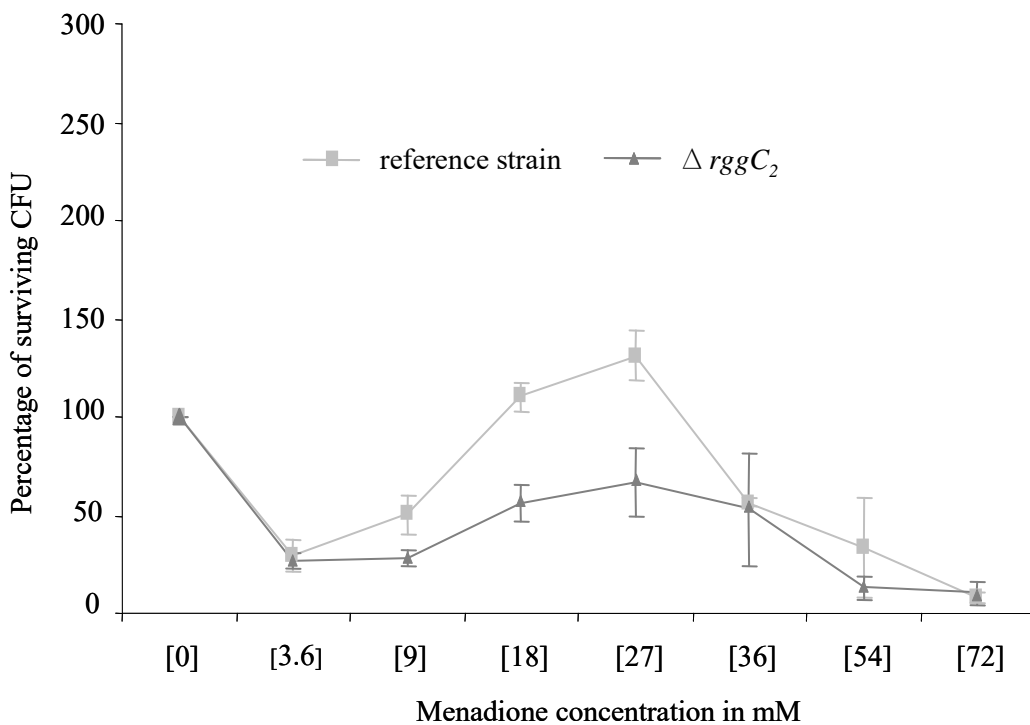
512

513

514



B



515 Fig. 3
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540

