**Applied and Environmental Microbiology**

# CosR is a global regulator of the osmotic stress response with widespread distribution among bacteria

Gwendolyn J. Gregory, Daniel P. Morreale, and E. Fidelma Boyd*

**Department of Biological Sciences, University of Delaware, Newark, DE, 19716**

Corresponding author*

E. Fidelma Boyd

Department of Biological Sciences

341 Wolf Hall, University of Delaware

Newark, DE 19716

Phone: (302) 831-1088. Fax: (302) 831-2281 Email: fboyd@udel.edu

**ABSTRACT** Bacteria accumulate small, organic compounds, called compatible solutes, via uptake from the environment or biosynthesis from available precursors to maintain the turgor pressure of the cell in response to osmotic stress. The halophile *Vibrio parahaemolyticus* has biosynthesis pathways for the compatible solutes ectoine (*ectABCasp_ect)* and glycine betaine (*betIBAproXWV)*, four betaine-carnitine-choline transporters (*bcct1-bcct4*) and a second ProU transporter (*proVWX).* All of these systems are osmotically inducible with the exception of *bcct2.* Previously, it was shown that CosR, a MarR-type regulator, was a direct repressor of *ectABCasp_ect* in *Vibrio* species. In this study, we investigated whether CosR has a broader role in the osmotic stress response. Expression analyses demonstrated that *betIBAproXWV*, *bcct1*, *bcct3*, *bcct4* and *proVWX* are repressed in low salinity. Examination of an in-frame *cosR* deletion mutant showed expression of these systems is de-repressed in the mutant at low salinity compared to wild-type. DNA binding assays demonstrated that purified CosR binds directly to the regulatory region of both biosynthesis systems and four transporters. In *Escherichia coli* GFP reporter assays, we demonstrated that CosR directly represses transcription of *betIBAproXWV*, *bcct3*, and *proVWX*. Similar to *V. harveyi*, we showed *betIBAproXWV* was directly activated by the quorum sensing LuxR homolog OpaR, suggesting a conserved mechanism of regulation among *Vibrio* species. Phylogenetic analysis demonstrated that CosR is ancestral to the *Vibrionaceae* family and bioinformatics analysis showed widespread distribution among *Gamma-Proteobacteria* in general. Incidentally, in *Aliivibrio fischeri, A. finisterrensis, A. sifiae* and *A. wodanis*, an unrelated MarR-type regulator named *ectR* was clustered with *ectABC-asp*, which suggests the presence of another novel ectoine biosynthesis regulator. Overall, these data show that CosR is a global regulator of osmotic stress response that is widespread among bacteria.

**IMPORTANCE** *Vibrio parahaemolyticus* can accumulate compatible solutes via biosynthesis and transport, which allow the cell to survive in high salinity conditions. There is little need for compatible solutes under low salinity conditions, and biosynthesis and transporter systems need to be repressed. However, the mechanism(s) of this repression is not known. In this study, we showed that CosR played a major role in the regulation of multiple compatible solute systems. Phylogenetic analysis showed that CosR is present in all members of the *Vibrionaceae* family as well as numerous *Gamma-Proteobacteria*. Collectively, these data establish CosR as a global regulator of the osmotic stress response that is widespread in bacteria, controlling many more systems than previously demonstrated.

**KEYWORDS** MarR-type regulator CosR osmotic stress response

## [INTRODUCTION]

Halophilic bacteria such as *Vibrio parahaemolyticus* encounter a range of osmolarities and have an absolute requirement for salt. To combat the loss of turgor pressure due to efflux of water in high osmolarity conditions, bacteria have developed a short-term "salt in" strategy requiring the uptake of K+ and a long-term "salt out" strategy that involves the accumulation of compatible solutes in the cell (1-3). Compatible solutes, as the name suggests, are organic compounds that are compatible with the molecular machinery and processes of the cell, and include compounds such as ectoine, glycine betaine, trehalose, glycerol, proline, glutamate, and carnitine, among others (1, 4-9). Compatible solutes are taken up from the environment or biosynthesized from various precursors in response to osmotic stress, which allows cells to continue to grow and divide even in unfavorable environments (2, 4, 10, 11).

Searches of the genome database demonstrated that ectoine biosynthesis genes are present in over 500 bacterial species (12). Most of the species that contain ectoine biosynthesis genes are halotolerant or halophiles. Previously, it was shown that ectoine biosynthesis is present in all halophilic *Vibrio* species including *Vibrio parahaemolyticus,* and this species also possesses the genes for glycine betaine biosynthesis and multiple compatible solute transporters (13). *De novo* biosynthesis of ectoine requires aspartic acid as the precursor, which can be supplied by the cell (14). Aspartic acid is converted to ectoine by four enzymes, EctA, EctB, EctC and Asp_Ect, encoded by the operon *ectABCasp_ect* (15). Ectoine biosynthesis begins with L-aspartate-β-semialdehyde, which is also pivotal to bacterial amino acid and cell wall synthesis (15). Asp_Ect is a specialized aspartokinase dedicated to the ectoine pathway that, among Proteobacteria, is present only in alpha, gamma and delta species (16). Our recent study showed that the quorum sensing response regulator OpaR was a negative regulator and AphA was a

positive regulator of *ectABC-asp_ect* gene expression (17). In addition, we showed that OpaR and AphA are positive regulators of *cosR,* which encodes a MarR-type regulator CosR (17, 18). We showed that, similar to *V. cholerae*, CosR is a repressor of *ectABC-asp_ect* indicating that control of ectoine biosynthesis is multilayered and stringent (17).

Production of glycine betaine is a two-step oxidation from the precursor choline, which is acquired exogenously. *De novo* biosynthesis of glycine betaine has been identified in only a few species of halophilic bacteria (19-24). Choline is converted to glycine betaine by the products of two genes *betB* and *betA* (25, 26). In *E. coli*, these genes are encoded by the operon *betIBA,* with the regulator BetI shown to repress its own operon (27, 28). In all *Vibrio* species that biosynthesize glycine betaine, the *betIBA* genes are in an operon with the *proXWV* genes, which encode an ATP-binding cassette (ABC)-type transporter named ProU2 (13, 14, 29). Regulation of glycine betaine biosynthesis has been studied in several species, but few direct mechanisms of regulation have been shown beyond BetI (27, 28, 30-33). Recently, in *V. harveyi*, a close relative of *V. parahaemolyticus*, *betIBAproXWV* was shown to be positively regulated by the quorum sensing master regulator LuxR (32, 33).

It is energetically favorable to the cell to uptake compatible solutes from the environment rather than to biosynthesize them, and Bacteria and Archaea encode multiple osmoregulated transporters (9, 34-39). ABC-type transporters are utilized to import exogenous compatible solutes into the cell and include ProU (encoded by *proVWX*) in *E. coli* and *Pseudomonas syringae*, OpuA in *Lactococcus lactis* and *B. subtilis*, and OpuC in *P. syringae* (39-44). *Vibrio parahaemolyticus* encodes two ProU transporters, one on each chromosome. ProU1 is encoded by *proVWX* (VP1726-VP1728) and ProU2 is encoded by the *betIBAproXWV* operon (VPA1109-VPA1114) (13). ProU1 is a homolog of the *E. coli* K-12 ProU, which in this species was shown

97 to bind glycine betaine with high affinity (41, 45, 46). ProU2 is a homolog of the *P. syringae*
98 *proVXW* (13).

99 The betaine-carnitine-choline transporters (BCCTs) are single component sodium- or
100 proton coupled transporters, the first of which, BetT, discovered in *E. coli*, was shown to
101 transport choline with high-affinity and is divergently transcribed from *betIBA* (47, 48). *Vibrio*
102 *parahaemolyticus* encodes four BCCTs, BCCT1-BCCT3 (VP1456, VP1723, VP1905), and
103 BCCT4 (VPA0356) (13). The *bcct2* (VP1723) gene is the only *bcct* that is not induced by
104 salinity in *V. parahaemolyticus* (14). All four BCCT transporters were shown to transport
105 glycine betaine amongst other compatible solutes (29). A study in *V. cholerae* demonstrated that
106 a *bcct3* homolog is repressed by the regulator CosR in low salt conditions (18).

107 To date there has been no single regulator identified that controls multiple compatible
108 solute systems in bacteria. In this study, we examined whether CosR could have a broader role in
109 the osmotic stress response. First, we examined expression of genes encoding osmotic stress
110 response systems in low salinity and used quantitative real-time PCR to quantify expression of
111 these genes in a Δ*cosR* deletion mutant. This analysis showed that CosR was a negative regulator
112 of both ectoine and glycine betaine biosynthesis systems and two different transporter systems;
113 the ABC-type transporters ProU1 and ProU2 and the sodium-coupled transporters BCCT1 and
114 BCCT3. These data indicate that the CosR regulon is larger than appreciated and expands the
115 role of CosR to that of a global regulator of the osmotic stress response. We determined whether
116 CosR was a direct regulator using DNA binding assays and an *E. coli* plasmid-based reporter
117 assay. We also examined whether *betIBAproXWV* was under the control of the quorum sensing
118 regulator OpaR, which also regulates *cosR*. We showed that OpaR is an activator of

119 *betIBAproXWV* in contrast to its repression of *ectABCasp_ect*. Phylogenetic analysis of CosR
120 showed it is ancestral to the *Vibrionaceae* family, present in all members of the group.
121 Bioinformatics analysis indicated that CosR homologs are also prevalent among *Gamma-*
122 *Proteobacteria* in general. Overall, the data show that CosR is a previously unrecognized global
123 regulator of the osmotic stress response that is widespread among bacteria.

124

125 **RESULTS**

126 **Compatible solute biosynthesis and transport genes are downregulated in low salinity.** We
127 have previously shown that *V. parahaemolyticus* does not produce compatible solutes ectoine
128 and glycine betaine during growth in minimal media (M9G) supplemented with 1% NaCl
129 (M9G1%) (13, 14). Here we quantified expression levels of both biosynthesis operons in
130 M9G1% or M9G3%. RNA was isolated from exponentially growing wild-type *V.*
131 *parahaemolyticus* RIMD2210633 cells, at optical density 595 nm ($OD_{595}$) 0.45, after growth in
132 M9G1% or M9G3%. Real time quantitative PCR (qPCR) showed that ectoine biosynthesis genes
133 *ectA* and *asp_ect* are differentially expressed in M9G1% as compared to expression in M9G3%.
134 *ectA* is significantly downregulated 794.6-fold and *asp_ect* is significantly downregulated 204.9-
135 fold in M9G1% (**Fig. 1A**). The *betIBAproXWV* operon is also significantly repressed in M9G1%,
136 with fold changes of 25.8-fold, 22-fold, 33.7-fold, and 52.8-fold for *betI*, *betB*, *proX*, and *proW*,
137 respectively (**Fig. 1B**).

138 Similarly, the expression of *bcct1*, *bcct3*, *bcct4,* and *proV1* are significantly repressed in
139 M9G1%, 500-fold, 71.4-fold, 11.6-fold, and 2,786-fold, respectively, when compared with
140 expression in M9G3% (**Fig. 1C**). The *bcct2* gene remained unchanged. Previously, we reported
141 that *bcct2* is not induced by salinity (29), and our data indicated that it has a basal level of
142 transcription in the cell based on similar Ct values in both salinities tested (data not shown).
143 Overall, the data demonstrate osmoregulation of *ectABCasp_ect*, *betIBAproXWV*, *bcct1, bcct3,*
144 *bcct4* and *proV1*.

145 **CosR represses compatible solute biosynthesis and transport genes in low salinity.** We know
146 CosR is a repressor of ectoine biosynthesis genes, and we wondered whether it played a broader

role in the regulation of other osmotic stress response genes. Therefore, we examined expression in wild-type and an in-frame deletion mutant of *cosR*. RNA was isolated from the Δ*cosR* mutant strain at mid-exponential phase ($OD_{595}$ 0.45) after growth in M9G1% and compared to wild-type grown under identical conditions. Using qPCR analysis, we determined the expression levels of *ectA* and *asp_ect* and showed they are significantly upregulated, 818.5-fold and 308.2-fold, respectively, in a Δ*cosR* mutant compared to wild-type in M9G1% (**Fig. 2A**), indicating de-repression in the absence of CosR. Next, we examined expression levels of *betIBAproXWV* and showed these genes are significantly de-repressed in the Δ*cosR* mutant (**Fig. 2B**). Similarly, relative expression levels of *bcct1*, *bcct*3 and *proV1* were significantly higher in Δ*cosR* than wild-type, while levels of *bcct2* and *bcct4* were unchanged (**Fig. 2C**). In sum, these data demonstrated that CosR is a repressor of *ectABCasp_ect*, *betIBAproXWV*, *bcct1, bcct3* and *proVWX1* under low salinity conditions. Thus, CosR is a unique example of a regulator that controls multiple compatible solute systems.

**CosR binds directly to the promoter of the *betIBAproXWV* operon and represses transcription.** To determine whether CosR regulation of *betIBAproXWV* is direct, we performed DNA binding assays with purified CosR protein and DNA probes of the regulatory region of this operon. The regulatory region was split into five overlapping probes, P*betI* probes A-E (**Fig. 3A**). CosR bound to probe A, which is directly upstream of the start codon for *betI*, and it also bound to probes B and D (**Fig. 3B**). CosR did not bind to probes C and E, which demonstrated specificity of CosR binding (**Fig. 3B**).

To demonstrate that direct binding by CosR results in transcriptional repression of the *betIBAproXWV* operon, we performed a GFP-reporter assay in *E. coli* strain MKH13. Full-length

169 *cosR* was expressed from a plasmid (pBBR*cosR*) in the presence of a *gfp*-expressing reporter
170 plasmid under the control of the glycine betaine biosynthesis system regulatory region ($P_{betI}$-*gfp*).
171 Relative fluorescence and $OD_{595}$ were measured after overnight growth in M9G1%. Specific
172 fluorescence was calculated by normalizing to OD and compared to specific fluorescence in a
173 strain with an empty expression vector (pBBR1MCS) that also contained the $P_{betI}$-*gfp* reporter
174 plasmid. The activity of the $P_{betI}$-*gfp* reporter was significantly repressed 4.84-fold as compared
175 to the empty vector strain (**Fig. 3C**). This indicates that CosR directly represses transcription of
176 the *betIBAproXWV* genes.

177 **CosR binds directly to the promoter of *bcct1* and *bcct3* and is a direct repressor of**
178 ***bcct3*.** Next, we wanted to investigate whether CosR repression of *bcct1* and *bcct3* was direct.
179 We designed probes upstream of the translational start for *bcct1* and *bcct3*. The 291-bp
180 regulatory region of P*bcct1*, which includes 15-bp of *bcct1* and 276-bp of the intergenic region,
181 was split into three overlapping probes, P*bcct1* probes A, B, and C (**Fig. 4A**). DNA binding
182 assays were performed with increasing concentrations of CosR. CosR bound directly to the
183 P*bcct1* probe B but did not bind to the other probes tested, which indicated direct and specific
184 binding by CosR (**Fig. 4B**). Next, we performed reporter assays in *E. coli* using a GFP
185 expression plasmid under the control of the regulatory region of *bcct1* ($P_{bcct1}$-*gfp*) and a CosR
186 expression plasmid (pBBR*cosR*). Specific fluorescence in the presence of CosR was compared to
187 a strain with empty expression vector (pBBR1MCS). The activity of the $P_{bcct1}$-*gfp* reporter was
188 not significantly different than the strain harboring empty expression vector, which indicates that
189 CosR does not directly repress *bcct1* (**Fig. 4C**). We speculate that CosR may still directly repress
190 *bcct1*, but in our reporter assay the low level of activation of the *bcct1* regulatory region in *E.*
191 *coli* may have affected the significance of the results. In the *E. coli* heterologous background,

192 additional proteins, which are present in the native species, may be necessary for full repression
193 of *bcct1* by CosR.

194 Two overlapping probes designated P*bcct3* probe A and B, were designed encompassing
195 196-bp of the regulatory region of *bcct3* (**Fig. 5A**). Because *bcct3* is divergently transcribed from
196 *cosR*, we used approximately half of the regulatory region for the P*bcct3* EMSA. An EMSA
197 showed that CosR bound directly to the P*bcct3* probe A, which is proximal to the start of the
198 gene, but not probe B (**Fig. 5B**). We then performed reporter assays in *E. coli* using a GFP
199 expression plasmid under the control of the regulatory region of *bcct3*, utilizing the entire 397-bp
200 intergenic region between *bcct3* and *cosR*. Transcriptional activity of the $P_{bcct3}$-*gfp* reporter is
201 repressed in a CosR-expressing strain (**Fig. 5C**), although not to the same extent that we saw in
202 expression analyses in *V. parahaemolyticus*. This is not surprising, given that it appears the
203 regulatory region of *bcct3* is not very active in an *E. coli* background, which made detection of
204 repression more difficult. Additionally, other proteins are likely necessary for full repression of
205 the regulatory region of *bcct3* that are not present in an *E. coli* background. The *E. coli* GFP
206 assay did show a direct interaction between CosR and the *bcct3* regulatory region that resulted in
207 repression of transcription. This result, in combination with expression analyses in a *cosR* mutant
208 (**Fig. 2C**) and binding assays which demonstrated direct binding (**Fig. 5B**), indicate that CosR
209 directly represses *bcct3*. In addition, we showed that CosR does not bind to the regulatory region
210 of *bcct2* and *bcct4* (**Fig. 5D**), which is in agreement with the *cosR* mutant expression data (**Fig.**
211 **2C**). These data suggest that *bcct2* and *bcct4* are under the control of a yet to be described
212 regulator.

213 **CosR is a direct repressor of *proVWX1*.** The regulatory region upstream of the *proV1* gene was
214 divided into four probes (**Fig. 6A**). A DNA binding assay was performed with increasing

concentrations of CosR and 30 ng of each probe. A shift in the DNA bands of probe D, which is proximal to the start codon of *proV1*, indicated that CosR binds directly to this region (**Fig. 6B**). CosR did not bind to the other probes tested, which indicated that CosR binding is specific.

We also performed a reporter assay in *E. coli* utilizing the *cosR* expression plasmid (pBBRcosR) and a GFP reporter plasmid ($P_{proV1}$-*gfp*). In a CosR-expressing strain, expression of the $P_{proV1}$-*gfp* reporter was repressed when compared to an empty expression vector strain (**Fig. 6C**). This repression was to a lesser extent than is seen in *V. parahaemolyticus* but recapitulation of the same magnitude of repression in the heterologous background is not to be expected given the potential absence of additional factors present in the native background. Overall, the results of the *E. coli* reporter assay, taken together with expression analyses in the native background (**Fig. 2C**) and the DNA binding assay (**Fig. 6B**), indicate that CosR is a direct repressor of the *proVWX1* operon.

**CosR is not autoregulated.** In *V. cholerae*, expression levels of *cosR* were upregulated in 0.5 M NaCl as compared to levels in 0.2 M NaCl (18). It was suggested that one reason for the upregulation of *cosR* in higher salinity could be that it is involved in an autoregulatory feedback loop (18). In *V. parahaemolyticus,* we found that levels of *cosR* were not significantly upregulated in 3% NaCl as compared to 1% NaCl (data not shown). We have already shown that CosR binds to the intergenic region between *bcct3* and *cosR*, but the binding site location is proximal to the start codon of *bcct3*, more than 300-bp upstream of the *cosR* gene (**Fig. 5A & B**). Therefore, to investigate CosR autoregulation, we designed two probes, 105-bp and 142-bp, which comprise a 220-bp portion of the regulatory region upstream of *cosR* (VP1906) (**Fig. 7A**) and used this in a DNA binding assay with various concentrations of purified CosR (**Fig. 7B**). There were no shifts observed in the binding assay, which indicated that CosR does not bind

238 (Fig. 7B). We then performed a GFP reporter assay in *E. coli*, utilizing the entire 397-bp
239 intergenic region between *bcct3* and *cosR*, to determine if CosR directly represses transcription
240 of its own gene. The transcriptional activity of $P_{cosR}$-*gfp* in the presence of CosR was not
241 significantly different from the empty-vector strain (p=0.09) (**Fig. 7C**). Because we cannot
242 assess expression of *cosR* in a Δ*cosR* mutant, we examined this in a GFP reporter assay in wild-
243 type and a Δ*cosR* mutant after growth in M9G1%. We found that the activity of a $P_{cosR}$-*gfp*
244 reporter was not different between wild-type and the *cosR* mutant (**Fig. 7D**). Taken together, lack
245 of CosR binding in the EMSA and both *in vivo* and *E. coli* reporter assays lead us to conclude
246 that under these conditions, CosR does not autoregulate, and that the CosR binding site proximal
247 to the *bcct3* gene does not affect transcription of the *cosR* gene.

248 **BetI represses its own operon *betIBA*.** Previously, it was shown that BetI represses its own
249 operon in several bacterial species and this repression is relieved in the presence of choline (27,
250 31, 32). To demonstrate BetI regulates its own operon in *V. parahaemolyticus*, we performed a
251 reporter assay utilizing the $P_{betI}$-*gfp* reporter in wild-type and a Δ*betI* mutant strain. Strains were
252 grown overnight in M9G3%, with and without choline, and specific fluorescence was calculated.
253 Expression of the reporter was de-repressed in the Δ*betI* mutant when no choline was present,
254 indicating that BetI is a negative regulator of its own operon (**Fig. 8A**). In the presence of
255 choline, there was no longer a difference in reporter activity between the wild-type strain and the
256 Δ*betI* mutant strain, indicating that repression by BetI was relieved (**Fig. 8B**). To confirm
257 regulation of *betIBAproXWV* by BetI is direct, we performed a GFP reporter assay in *E. coli*
258 MKH13 strain. The $P_{betI}$-*gfp* reporter was transformed into *E. coli* MKH13 (which lacks the
259 *betIBA* operon) along with an expression vector harboring full-length *betI* under the control of an
260 IPTG-inducible promoter. In the BetI-expressing strain, $P_{betI}$-*gfp* expression was significantly

261 repressed, which indicated that BetI is a direct repressor of its own operon in *V.*
262 *parahaemolyticus* (**Fig. 8C**).

263 **The quorum sensing LuxR homolog OpaR is a positive regulator of *betIBAproXWV* in *V.***
264 ***parahaemolyticus*.** We examined expression of the $P_{betI}$-*gfp* reporter in wild-type and the Δ*opaR*
265 mutant in *V. parahaemolyticus*. Expression of the reporter was significantly down-regulated in
266 Δ*opaR*, indicating that OpaR is a positive regulator of the glycine betaine biosynthesis operon
267 (**Fig. 9A**). We also examined whether regulation of P*betI* by OpaR was direct utilizing an EMSA
268 with purified OpaR protein. OpaR bound to P*betI* probes A, B, C and E, and very weakly to
269 probe D, which indicated that regulation of *betIBAproXWV* by OpaR is direct (**Fig. 9B**). These
270 results are in agreement with a previous study, which also showed direct positive regulation of
271 *betIBAproVWX* by LuxR in *V. harveyi* (32). Thus, it appears that the quorum sensing master
272 regulator may be a conserved regulatory mechanism of glycine betaine biosynthesis among
273 *Vibrio* species.

274 **Motif identification and phylogenetic distribution of CosR.** CosR bound to eight of the
275 probes tested in our DNA binding analyses, including two probes of the *ectABCasp_ect*
276 regulatory region, as shown previously (17). We utilized these sequences in MEME (multiple
277 EM for motif elicitation) analysis (49), and identified a 24-bp pseudo-palindromic motif present
278 in each of the eight sequences (**Fig. 10A**). The motif is an imperfect inverted repeat separated by
279 two bp (TTTGA-NN-TCTAA). The alignment of the motifs found within each sequence is
280 shown in **Figure 10B**.

281 CosR, a MarR-type regulator, is a 158 amino acid protein that is divergently transcribed
282 from *bcct3* on chromosome 1 in *V. cholerae* and *V. parahaemolyticus*, two distantly related

283 species. Bioinformatics analysis showed that a CosR homolog is present in over 50 *Vibrio*
284 species and in all cases the *cosR* homolog was divergently transcribed from a *bcct* transporter
285 (**Fig. 11**). Within these *Vibrio* species, similarity ranged from 98% to 73% amino acid identity
286 and showed that CosR was present in phylogenetically divergent *Vibrio* species. We found that
287 in *V. splendidus, V. crassostreae, V. cyclitrophicus, V. celticus*, *V. lentus* and *Aliivibrio wodanis,*
288 the CosR homolog is present directly downstream of the *betIBAproXWV* operon and in *V.*
289 *tasmaniensis* strains and *Vibrio sp.* MED222, the ectoine biosynthesis operon clustered in the
290 same genome location (**Fig. 11**). Collectively, these data indicated that CosR function is
291 conserved among this divergent group of species and that CosR is an important regulator of the
292 osmotic stress response. In all strains of *Aliivibrio fischeri*, the *cosR* homolog (which shares 73%
293 amino acid identity with CosR from *V. parahaemolyticus*) clusters with two uncharacterized
294 transporters. A recent phylogenomics study of distribution of ectoine biosynthesis genes and a
295 homolog of CosR showed the presence of this regulator in species of the *Alpha-Proteobacteria,*
296 *Beta-Proteobacteria* and *Gamma-Proteobacteria* (50). This again suggests that the role of CosR
297 in the osmotic stress response is conserved and phylogenetically widespread.

298 Incidentally, a second MarR-type regulator, a 141 amino acid protein, which we name
299 *ectR*, clusters with the ectoine biosynthesis genes in *Aliivibrio fischeri* (**Fig.11**). EctR shares only
300 31% identity with less than 60% query coverage to CosR from *V. parahaemolyticus* and a similar
301 level of low amino acid identity to EctR1 from *Methylmicrobium alcaliphilum*. EctR was
302 clustered with the *ectABCasp_ect* genes in all strains of *Aliivibrio finisterrensis, Aliivibrio sifiae,*
303 and most *A. wodanis* strains. Thus, in *Aliivibrio* species, it appears that the ectoine biosynthesis
304 gene cluster has a new uncharacterized MarR-type regulator.

305

## DISCUSSION

Here we have shown that the compatible solute biosynthesis and transport genes are downregulated in *V. parahaemolyticus* in low salinity and these genes are de-repressed in a *cosR* mutant. Our genetic analyses, binding analyses, and reporter assays demonstrated that CosR is a regulator of *betIBAproXWV, bcct1*, *bcct3*, and *proVWX* (**Fig. 13**). Additionally, we showed that under the conditions tested, CosR is not autoregulated. To date, it has now been demonstrated that CosR is a regulator of six different compatible solute systems; two biosynthesis systems ectoine and glycine betaine and four transporters (**Fig. 13**). Our phylogenetic and bioinformatics analyses indicated that CosR is universal within the *Vibrionaceae* and widespread in *Gamma-Proteobacteria* in general suggesting a conserved previously unrecognized global regulator of the osmotic stress response in bacteria.

The physiological importance of CosR repression of compatible solute biosynthesis in low salinity is to protect levels of key intracellular metabolites. Ectoine biosynthesis requires the precursor aspartate and this affects the level of glutamate, acetyl-CoA, and oxaloacetate (51, 52). Thus, tight regulation of ectoine biosynthesis is essential for cellular fitness. CosR characterized from *Vibrio* species show ~50% amino acid identity to EctR1 first described in the halotolerant methanotroph *Methylmicrobium alcaliphilum* that repressed ectoine biosynthesis (53). In this species, *ectR1* is divergently transcribed from the same promoter as *ectABC-ask*. Mustakhimov and colleagues showed that EctR1 is a direct repressed expression of the *ectABC-ask* operon in response to low salinity (53). EctR repression of the ectoine biosynthesis genes was also shown in both *Methylophaga alcalica* and *Methylophaga thalassica*, two moderately halophilic methylotrophs (54, 55). Czech and colleagues showed that CosR/EctR1 was phylogenetically widespread and clustered with *ect* genes in some species (50). In *V. cholerae*, CosR was also

identified as a repressor of ectoine biosynthesis genes though it does not cluster with *ectABCasp_ect* (18). The *cosR* gene in *V. cholerae* is divergently transcribed from the *opuD* gene (a *bcct3* homolog), which was also repressed by CosR (18). Similarly, in *V. parahaemolyticus,* the *cosR* (VP1906) homolog is divergently transcribed from *bcct3* (VP1905) and is a direct negative regulator of both *bcct3* and *ectABCasp_ect* (17). Our phylogenetic analysis found that the CosR homolog was present in all members of the *Vibrionaceae*, and among many *Vibrio* species was clustered with a *bcct3* homolog. The phylogeny of CosR mirrored the branching pattern of the relationships of members of the group for other housekeeping genes. These data indicate that CosR is ancestral to the group and the conservation of genomic context suggests functional conservation (**Fig. 11 and Fig. 12**). We used CosR from *Aeromonas* species to root the tree and similar to *Vibrionaeae* species, CosR is present in all members of this group. Indeed, bioinformatics analysis indicated that a CosR homolog is present in many other *Gamma-Proteobacteria* suggesting that it is an under-appreciated player in the osmotic stress response in bacteria (data not shown). In several *Vibrio* species, the CosR homolog was clustered with the *betIBAproXWV* operon, which is further suggestive of its role in regulation of compatible solute biosynthesis among *Vibrio* species.

The MarR family of transcriptional regulators, first characterized in *E. coli*, are important regulators of a number of cellular responses, typically responding to a change in the external environment (56-58). The literature suggests that MarR-type regulators form dimers and bind to a 20-45 bp pseudo-palindromic site in the intergenic region of genes they control (56, 59-61). We utilized the regulatory regions of each of the osmotic stress response genes that CosR regulates and identified a pseudo-palindromic CosR DNA binding motif (**Fig. 10**). This motif is similar to the binding sequence of the CosR homolog EctR1 identified previously in *M.*

*alcaliphilum*, which was also pseudo-palindromic with 2-bp separating inverted repeats (53). The activity of MarR-type regulators can be modulated by the presence of a chemical signal, either a ligand, metal ion, or reactive oxygen species. Binding of these signals causes the protein to undergo a conformational change, thereby affecting DNA binding capability (56, 62, 63). We modeled a CosR homodimer using SWISS-MODEL and did not identify a ligand binding pocket (data not shown). In *V. cholerae*, CosR activity was not affected by the presence of exogenous compatible solutes including ectoine, glycine betaine and proline, and *opuD* (*bcct* homolog) transcripts were unchanged in a *cosR* mutant (50). Hence, the environmental or cellular signals that modulate the activity of CosR remain unidentified, as was noted by Czech and colleagues (50). Interestingly, our modelling of the EctR regulator clustered with ectoine genes identified in *Aliivibrio* species indicated it also does not have a ligand-binding pocket (data not shown). Autoregulation was shown for several MarR family regulators, including *ectR1* in *M. alcaliphilum* (53, 56). In *V. parahaemolyticus,* we showed CosR does not bind to its own regulatory region, and our reporter assays suggested that CosR does not autoregulate. It is interesting to note that EctR1 does not participate in an autoregulatory feedback loop in *M. thalassica* (55, 64).

Similar to ectoine biosynthesis gene expression, few direct regulators of glycine betaine biosynthesis genes have been identified. In *E. coli*, expression of *betIBA* was repressed by BetI and repression was relieved in the presence of choline (27). BetI was shown to directly regulate transcription at this locus via DNA binding assays (28). ArcA was shown to repress the *betIBA* operon under anaerobic conditions in *E. coli*, although direct binding was not shown (27). In *V. harveyi,* it was shown that *betIBAproXWV* were repressed 2- to 3-fold when *betI* was overexpressed from a plasmid. Purified BetI bound directly to the regulatory region of the

375 *betIBAproXWV* operon in DNA binding assays (32, 33). In these studies, it was also shown that
376 the quorum sensing response regulator LuxR, along with the global regulator IHF, activated
377 expression of *betIBAproXWV* (32, 33). Here, we have shown that BetI represses its own operon
378 in *V. parahaemolyticus*, as expected, and we identified a novel regulator of glycine betaine
379 biosynthesis genes, CosR, which directly represses under low salinity conditions (**Fig. 13**). We
380 also confirm that, similar to *V. harveyi*, the quorum sensing master regulator OpaR directly
381 induced *betIBAproXWV* expression in *V. parahaemolyticus*, indicating this mechanism is likely
382 conserved in *Vibrio* species (**Fig. 13**).

383 Biosynthesis of compatible solutes is an energetically costly process for bacteria (35).
384 *Vibrio parahaemolyticus* does not accumulate compatible solutes in low salinity (13, 14, 29), and
385 therefore the transcription of biosynthesis and transport genes is unnecessary. CosR represses
386 these genes involved in the osmotic stress response in *V. parahaemolyticus*. The high
387 conservation of the CosR protein across *Vibrionaceae* and *Gamma-Proteobacteria* and its
388 genomic context indicates that regulation by CosR of compatible solute systems is widespread in
389 bacteria.

390
391 **MATERIALS AND METHODS**

392 **Bacterial strains, media and culture conditions.** Listed in Table 1 are all strains and plasmids
393 used in this study. A previously described streptomycin-resistant clinical isolate of *V.*
394 *parahaemolyticus*, RIMD2210633, was used as the wild-type strain (65, 66). *Vibrio*
395 *parahaemolyticus* strains were grown in either lysogeny broth (LB) (Fisher Scientific, Fair
396 Lawn, NJ) supplemented with 3% NaCl (wt/vol) (LBS) or in M9 minimal medium (47.8 mM
397 $Na_2HPO_4$, 22 mM $KH_2PO_4$, 18.7 mM $NH_4Cl$, 8.6 mM NaCl) (Sigma-Aldrich, USA)
398 supplemented with 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 20 mM glucose as the sole carbon source

(M9G) and 1% or 3% NaCl (wt/vol) (M9G1%, M9G3%). *E. coli* strains were grown in LB supplemented with 1% NaCl (wt/vol) or M9G1% where indicated. *E. coli* β2155, a diaminopimelic acid (DAP) auxotroph, was supplemented with 0.3 mM DAP and grown in LB 1% NaCl. All strains were grown at 37˚C with aeration. Antibiotics were used at the following concentrations (wt/vol) as necessary: ampicillin (Amp), 50 µg/ml; chloramphenicol (Cm), 12.5 µg/ml; tetracycline (Tet), 1 µg/mL; and streptomycin (Str), 200 µg/ml. Choline was added to media at a final concentration of 1 mM, when indicated.

**Construction of the *betI* deletion mutant.** An in-frame *betI* (VPA1114) deletion mutant was constructed as described previously (17). Briefly, the Gibson assembly protocol, using NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs, Ipswich, MA), followed by allelic exchange, was used to generate an in-frame 63-bp truncated, non-functional *betI* gene (67, 68). Two fragments, AB and CD, were amplified from the RIMD2210633 genome using primers listed in Table 2. These were ligated with pDS132, which had been digested with SphI, via Gibson assembly to produce suicide vector pDS132 with a truncated *betI* allele (pDSΔ*betI*). pDSΔ*betI* was transformed into *E. coli* strain β2155 λ*pir*, followed by conjugation with *V. parahaemolyticus*. The suicide vector pDS132 must be incorporated into the *V. parahaemolyticus* genome via homologous recombination, as *V. parahaemolyticus* lacks the *pir* gene required for replication of the vector. Growth without chloramphenicol induces a second recombination event which leaves behind either the truncated mutant allele or the wild-type allele. Colonies were plated on sucrose for selection, as pDS132 harbors a *sacB* gene, which makes sucrose toxic to cells still carrying the plasmid and colonies appear soapy. Healthy colonies were screened via PCR and sequenced to confirm an in-frame deletion of the *betI* gene.

**RNA isolation and qPCR.** *Vibrio parahaemolyticus* RIMD2210633 and Δ*cosR* were grown with aeration at 37 °C overnight in LBS. Cells were pelleted, washed twice with 1X PBS, diluted 1:50 into M9G3% or M9G1% and grown with aeration to mid-exponential phase ($OD_{595}$ 0.45). RNA was extracted from 1 mL of culture using Trizol, following the manufacturer's protocol (Invitrogen, Carlsbad, CA). The samples were treated with Turbo DNase (Invitrogen), followed by heat inactivation of the enzyme as per manufacturer's protocol. Final RNA concentration was quantified using a Nanodrop spectrophotometer (Thermo Scientific, Waltham, MA). 500 ng of RNA were used for cDNA synthesis by priming with random hexamers using SSIV reverse transcriptase (Invitrogen). Synthesized cDNA was diluted 1:25 and used for quantitative real-time PCR (qPCR). qPCR experiments were performed using PowerUp SYBR master mix (Life Technologies, Carlsbad, CA) on an Applied Biosystems QuantStudio6 fast real-time PCR system (Applied Biosystems, Foster City, CA). Reactions were set up with the following primer pairs listed in Table 2: VPbcct1Fwd/Rev, VPbcct2Fwd/Rev, VPbcct3Fwd/Rev, VPbcct4Fwd/Rev, VPectAFwd/Rev, VPasp_ectFwd/Rev, VPproV1Fwd/Rev, VPAbetIFwd/Rev, VPAbetBFwd/Rev, VPAproXFwd/Rev, VPAproWFwd/Rev, and 16SFwd/Rev for normalization. Expression levels were quantified using cycle threshold (CT) and were normalized to 16S rRNA. Differences in gene expression were determined using the ΔΔCT method (69).

**Protein purification of CosR.** CosR was purified as described previously (17). Briefly, full-length *cosR* (VP1906) was cloned into the protein expression vector pET28a (+) containing an IPTG-inducible promoter and a C-terminal 6x-His tag (Novagen). Expression of CosR-His was then induced in *E. coli* BL21 (DE3) with 0.5 mM IPTG at $OD_{595}$ of 0.4 and grown overnight at room temperature. Cells were harvested, resuspended in lysis buffer (50 mM NaPO4, 200 mM

NaCl, 20 mM imidazole buffer [pH 7.4]) and lysed using a microfluidizer. CosR-His was bound to a Ni-NTA column and eluted with 50 mM NaPO4, 200 mM NaCl, 500 mM imidazole buffer [pH 7.4] after a series of washes to remove loosely bound protein. Protein purity was determined via SDS-PAGE. OpaR was purified as described previously (70).

**Electrophoretic Mobility Shift Assay.** Five overlapping DNA fragments, designated P*betI* probe A (125-bp), probe B (112-bp), probe C (142-bp), probe D (202-bp) and probe E (158-bp), were generated from the *betIBAproXWV* regulatory region (includes 36 bp of the coding region and the 594-bp upstream intergenic region) using primer sets listed in Table 2. Three overlapping DNA fragments, designated P*bcct1* probe A (120-bp), probe B (110-bp), and probe C (101-bp), were generated from the *bcct1* regulatory region (includes 15 bp of the coding region and the 276-bp upstream intergenic region) using primer sets listed in Table 2. Two overlapping DNA fragments, designated P*bcct3* probe A (108-bp) and probe B (107-bp), were generated from the *bcct3* regulatory region (includes 17 bp of the coding region and 179-bp of the upstream intergenic region) using primer sets listed in Table 2. Four overlapping DNA fragments, designated P*proV1* probe A (160-bp), probe B (134-bp), probe C (108-bp), and probe D (109-bp), were generated from the *proV1* regulatory region (includes 9-bp of the coding region and the 438-bp upstream intergenic region) using primer sets listed in Table 2. Fragments designated P*bcct2* (233-bp) and P*bcct4* (244-bp) were generated from the *bcct2* and *bcct4* regulatory regions, respectively, using primers listed in Table 2. Two overlapping DNA fragments, designated P*cosR* probe A (105-bp) and probe B (142-bp), were generated from the *cosR* regulatory region (includes 4 bp of the coding region and 216-bp of the upstream intergenic region) using primer sets listed in Table 2. The concentration of purified CosR-His and OpaR was determined using a Bradford assay. CosR or OpaR was incubated for 20 minutes with 30 ng

of each DNA fragment in a defined binding buffer (10 mM Tris, 150 mM KCl, 0.5 mM dithiothreitol, 0.1 mM EDTA, 5% polyethylene glycol [PEG] [pH 7.9 at 4°C]). A 6% native acrylamide gel was pre-run for 2 hours at 4C (200 V) in 1 X TAE buffer. Gels were loaded with the DNA:protein mixtures (10 µL), and run for 2 hours at 4°C (200 V). Finally, gels were stained in an ethidium bromide bath for 15 min and imaged.

**Reporter Assays.** A GFP reporter assay was conducted using the *E. coli* strain MKH13 (71). GFP reporter plasmids were constructed as previously described (17). Briefly, each regulatory region of interest was amplified using primers listed in Table 2 and ligated via Gibson assembly protocol with the promoterless parent vector pRU1064, which had been digested with SpeI, to generate reporter plasmids with GFP under the control of the regulatory region of interest. Complementary regions for Gibson assembly are indicated in lower case letters in the primer sequence (Table 2). Reporter plasmid $P_{betI}$-*gfp* encompasses 594-bp upstream of the *betIBAproXWV* operon. Reporter plasmid $P_{bcct1}$-*gfp* encompasses 278-bp upstream of the P*bcct1* regulatory region. Reporter plasmid $P_{bcct3}$-*gfp* encompasses 397-bp upstream of the P*bcct3* regulatory region. Reporter plasmid $P_{proV1}$-*gfp* encompasses 438-bp upstream of the P*proV1* regulatory region. Reporter plasmid $P_{cosR}$-*gfp* encompasses 397-bp upstream of the P*cosR* regulatory region. The full-length *cosR* was then expressed from an IPTG-inducible promoter in the pBBR1MCS expression vector. Relative fluorescence (RFU) and $OD_{595}$ were measured; specific fluorescence was calculated by dividing RFU by $OD_{595}$. Strains were grown overnight with aeration at 37°C in LB1% with ampicillin (50 µg/mL) and chloramphenicol (12.5 µg/mL), washed twice with 1X PBS, then diluted 1:1000 in M9G1%. Expression of *cosR* was induced with 0.25 mM IPTG, and strains were grown for 20 hours at 37°C with aeration under antibiotic selection. GFP fluorescence was measured with excitation at 385 and emission at 509 nm in

black, clear-bottom 96-well plates on a Tecan Spark microplate reader with Magellan software (Tecan Systems Inc., San Jose, CA). Specific fluorescence was calculated for each sampled by normalizing fluorescence intensity to $OD_{595}$. Two biological replicates were performed for each assay.

A GFP reporter assay was conducted in RIMD2210633 wild-type, Δ*betI*, Δ*opaR*, or Δ*cosR* mutant strains. The $P_{betI}$-*gfp* or $P_{cosR}$-*gfp* reporter plasmid was transformed into *E. coli* β2155 λ*pir* and conjugated into wild-type, Δ*betI,* Δ*opaR,* or Δ*cosR* mutant strains. Strains were grown overnight with aeration at 37°C in LB3% with tetracycline (1 µg/mL). Cells were then pelleted, washed two times with 1X PBS, diluted 1:100 into M9G3% and grown for 20 hours with antibiotic selection. Choline was added to a final concentration of 1 mM, where indicated. GFP fluorescence was measured with excitation at 385 and emission at 509 nm in black, clear-bottom 96-well plates on a Tecan Spark microplate reader with Magellan software (Tecan Systems Inc.). Specific fluorescence was calculated for each sampled by normalizing fluorescence intensity to $OD_{595}$. Two biological replicates were performed for each assay.

**Bioinformatics and phylogenetic analyses.** Sequences of EMSA probes P*ectA* A and B, P*betI* A, B and D, P*bcct1* B, P*bcct3* A, and P*proV1* D to which CosR bound were input into the MEME (Multiple EM for Motif Elicitation) tool (meme-suite.org/tools/meme) (49). We set the parameters to search for one occurrence of one motif per sequence with a minimum width of 18-bp and a maximum width of 35-bp. The *V. parahaemolyticus* protein CosR (NP_798285) was used as a seed for BLASTp to identify homologs in the *Vibrionaceae* family in the NCBI database. Sequences of representative strains were downloaded from NCBI and used in a Python-based program Easyfig to visualize gene arrangements (72). Accession numbers for

select strains were: *V. parahaemolyticus* RIMD (BA00031), *V. crassotreae* 9CS106 (CP016229), *V. splendidus* BST398 (CP031056), *V. celticus* CECT7224 (NZ_FLQZ01000088), *V. lentus* 10N.286.51.B9 (NZ_MCUE01000044), *V. tasmaniensis* LGP32 (FM954973), *V. cyclitrophicus* ECSMB14105 (CO039701), *Aliivibrio fischeri* ES114 (CP000021), *A. fischeri* MJ11 (CP001133), *A. wodanis* AWOD1 (LN554847), *A. wodanis* 06/09/160 (CP039701). The *V. parahaemolyticus* RIMD2201633 CosR and *A. fischeri* ES114 EctR protein sequences were retrieved from NCBI using accession numbers NP_798285 and AAW88191.1, respectively, and input into the SWISS-MODEL workspace, which generated a 3D model of a homodimer to identify putative ligand-binding pockets (73-77). Evolutionary analysis was performed on the CosR protein from all species within the family *Vibrionaceae* with completed genome sequence and as an outgroup we used CosR from members of the genus *Aeromonas*. Protein sequences were obtained from NCBI database and aligned using the Clustal W algorithm (78). Aligned protein sequences were used to generate a Neighbor-Joining tree with a bootstrap value of 1000 (79, 80). The evolutionary distances were computed using the JTT matrix-based method [3] and are in the units of the number of amino acid substitutions per site (81). The rate variation among sites was modeled with a gamma distribution (shape parameter = 5). This analysis involved 96 amino acid sequences. All ambiguous positions were removed for each sequence pair using the pairwise deletion option. There were a total of 173 positions in the final dataset. Evolutionary analyses were conducted in MEGA X (82).

## ACKNOWLEDGEMENTS

This research was supported by a National Science Foundation grant (award IOS-1656688) to E.F.B. G.J.G. was funded in part by a University of Delaware graduate fellowship award. We

535 thank members of the Boyd Group and three anonymous reviewers for constructive feedback on
536 the manuscript.

## REFERENCES


1. Galinski EA. 1995. Osmoadaptation in bacteria. Adv Microb Physiol 37:272-328.
2. Csonka LN. 1989. Physiological and genetic responses of bacteria to osmotic stress. Microbiol Rev 53:121-47.
3. Wood JM. 2011. Bacterial osmoregulation: a paradigm for the study of cellular homeostasis. Annu Rev Microbiol 65:215-38.
4. da Costa MS, Santos H, Galinski EA. 1998. An overview of the role and diversity of compatible solutes in Bacteria and Archaea. Adv Biochem Eng Biotechnol 61:117-53.
5. Galinski EA, Oren A. 1991. Isolation and structure determination of a novel compatible solute from the moderately halophilic purple sulfur bacterium *Ectothiorhodospira marismortui*. Eur J Biochem 198:593-598.
6. Sleator RD, Hill C. 2002. Bacterial osmoadaptation: the role of osmolytes in bacterial stress and virulence. FEMS Microbiol Rev 26:49-71.
7. Roberts MF. 2004. Osmoadaptation and osmoregulation in archaea: update 2004. Front Biosci 9:1999-2019.
8. Roberts MF. 2005. Organic compatible solutes of halotolerant and halophilic microorganisms. Saline Systems 1:5.
9. Kempf B, Bremer E. 1998. Uptake and synthesis of compatible solutes as microbial stress responses to high-osmolality environments. Arch Microbiol 170:319-30.
10. Record MT, Jr., Courtenay ES, Cayley DS, Guttman HJ. 1998. Responses of *E. coli* to osmotic stress: large changes in amounts of cytoplasmic solutes and water. Trends Biochem Sci 23:143-8.
11. Wood JM. 1999. Osmosensing by bacteria: signals and membrane-based sensors. Microbiol Mol Biol Rev 63:230-62.
12. Widderich N, Hoppner A, Pittelkow M, Heider J, Smits SH, Bremer E. 2014. Biochemical properties of ectoine hydroxylases from extremophiles and their wider taxonomic distribution among microorganisms. PLoS One 9:e93809.
13. Naughton LM, Blumerman SL, Carlberg M, Boyd EF. 2009. Osmoadaptation among Vibrio species and unique genomic features and physiological responses of *Vibrio parahaemolyticus*. Appl Environ Microbiol 75:2802-10.
14. Ongagna-Yhombi SY, Boyd EF. 2013. Biosynthesis of the osmoprotectant ectoine, but not glycine betaine, is critical for survival of osmotically stressed *Vibrio parahaemolyticus* cells. Appl Environ Microbiol 79:5038-49.
15. Louis P, Galinski EA. 1997. Characterization of genes for the biosynthesis of the compatible solute ectoine from *Marinococcus halophilus* and osmoregulated expression in *Escherichia coli*. Microbiology 143 ( Pt 4):1141-9.
16. Lo CC, Bonner CA, Xie G, D'Souza M, Jensen RA. 2009. Cohesion group approach for evolutionary analysis of aspartokinase, an enzyme that feeds a branched network of many biochemical pathways, p 594-651, Microbiol Mol Biol Rev, 73.
17. Gregory GJ, Morreale DP, Carpenter MR, Kalburge SS, Boyd EF. 2019. Quorum sensing regulators AphA and OpaR control expression of the operon responsible for biosynthesis of the compatible solute ectoine. Appl Environ Microbiol 85.
18. Shikuma NJ, Davis KR, Fong JN, Yildiz FH. 2013. The transcriptional regulator, CosR, controls compatible solute biosynthesis and transport, motility and biofilm formation in *Vibrio cholerae*. Environ Microbiol 15:1387-99.

19. Roberts MF, Lai MC, Gunsalus RP. 1992. Biosynthetic pathways of the osmolytes N epsilon-acetyl-beta-lysine, beta-glutamine, and betaine in *Methanohalophilus* strain FDF1 suggested by nuclear magnetic resonance analyses. J Bacteriol 174:6688-93.
20. Lai MC, Yang DR, Chuang MJ. 1999. Regulatory factors associated with synthesis of the osmolyte glycine betaine in the halophilic methanoarchaeon *Methanohalophilus portucalensis*. Appl Environ Microbiol 65:828-33.
21. Nyyssölä A, Kerovuo J, Kaukinen P, Weymarn Nv, Reinikainen T. 2000. Extreme halophiles synthesize betaine from glycine by methylation. J. Biol. Chem. 275:22196-22201.
22. Waditee R, Tanaka Y, Aoki K, Hibino T, Jikuya H, Takano J, Takabe T. 2003. Isolation and functional characterization of N-methyltransferases that catalyze betaine synthesis from glycine in a halotolerant photosynthetic organism *Aphanothece halophytica*. J Biol Chem 278:4932-42.
23. Lu WD, Chi ZM, Su CD. 2006. Identification of glycine betaine as compatible solute in *Synechococcus sp.* WH8102 and characterization of its N-methyltransferase genes involved in betaine synthesis. Arch Microbiol 186:495-506.
24. Kimura Y, Kawasaki S, Yoshimoto H, Takegawa K. 2010. Glycine betaine biosynthesized from glycine provides an osmolyte for cell growth and spore germination during osmotic stress in *Myxococcus xanthus*. J Bacteriol 192:1467-70.
25. Landfald B, Strom AR. 1986. Choline-glycine betaine pathway confers a high level of osmotic tolerance in *Escherichia coli*. J Bacteriol 165:849-55.
26. Andresen PA, Kaasen I, Styrvold OB, Boulnois G, Strom AR. 1988. Molecular cloning, physical mapping and expression of the bet genes governing the osmoregulatory choline-glycine betaine pathway of *Escherichia coli*. J Gen Microbiol 134:1737-46.
27. Lamark T, Rokenes TP, McDougall J, Strom AR. 1996. The complex bet promoters of *Escherichia coli*: regulation by oxygen (ArcA), choline (BetI), and osmotic stress. J Bacteriol 178:1655-62.
28. Rokenes TP, Lamark T, Strom AR. 1996. DNA-binding properties of the BetI repressor protein of *Escherichia coli*: the inducer choline stimulates BetI-DNA complex formation. J Bacteriol 178:1663-70.
29. Ongagna-Yhombi SY, McDonald ND, Boyd EF. 2015. Deciphering the role of multiple betaine-carnitine-choline transporters in the Halophile *Vibrio parahaemolyticus*. Appl Environ Microbiol 81:351-63.
30. Eshoo MW. 1988. lac fusion analysis of the bet genes of *Escherichia coli*: regulation by osmolarity, temperature, oxygen, choline, and glycine betaine. J Bacteriol 170:5208-15.
31. Scholz A, Stahl J, de Berardinis V, Muller V, Averhoff B. 2016. Osmotic stress response in *Acinetobacter baylyi*: identification of a glycine-betaine biosynthesis pathway and regulation of osmoadaptive choline uptake and glycine-betaine synthesis through a choline-responsive BetI repressor. Environ Microbiol Rep 8:316-22.
32. van Kessel JC, Rutherford ST, Cong JP, Quinodoz S, Healy J, Bassler BL. 2015. Quorum sensing regulates the osmotic stress response in *Vibrio harveyi*. J Bacteriol 197:73-80.
33. Chaparian RR, Olney SG, Hustmyer CM, Rowe-Magnus DA, van Kessel JC. 2016. Integration host factor and LuxR synergistically bind DNA to coactivate quorum-sensing genes in *Vibrio harveyi*. Mol Microbiol 101:823-40.
34. Ventosa A, Nieto JJ, Oren A. 1998. Biology of moderately halophilic aerobic bacteria. Microbiol Mol Biol Rev 62:504-44.

35. Oren A. 1999. Bioenergetic aspects of halophilism. Microbiol Mol Biol Rev 63:334-48.
36. Wood JM. 2007. Bacterial osmosensing transporters. Methods Enzymol 428:77-107.
37. Culham DE, Henderson J, Crane RA, Wood JM. 2003. Osmosensor ProP of *Escherichia coli* responds to the concentration, chemistry, and molecular size of osmolytes in the proteoliposome lumen. Biochemistry, 42:410-420.
38. Rübenhagen R, Morbach S, Krämer R. 2001. The osmoreactive betaine carrier BetP from *Corynebacterium glutamicum* is a sensor for cytoplasmic K+. EMBO J, 20:5412-20.
39. van der Heide T, Stuart MC, Poolman B. 2001. On the osmotic signal and osmosensing mechanism of an ABC transport system for glycine betaine. EMBO J, 20:7022-32.
40. Cairney J, Booth IR, Higgins CF. 1985. Osmoregulation of gene expression in *Salmonella typhimurium*: proU encodes an osmotically induced betaine transport system. J Bacteriol 164:1224-32.
41. May G, Faatz E, Villarejo M, Bremer E. 1986. Binding protein dependent transport of glycine betaine and its osmotic regulation in *Escherichia coli* K12. Mol Gen Genet 205:225-33.
42. Kempf B, Bremer E. 1995. OpuA, an osmotically regulated binding protein-dependent transport system for the osmoprotectant glycine betaine in *Bacillus subtilis*. J Biol Chem 270:16701-13.
43. Mahmood NA, Biemans-Oldehinkel E, Patzlaff JS, Schuurman-Wolters GK, Poolman B. 2006. Ion specificity and ionic strength dependence of the osmoregulatory ABC transporter OpuA. J Biol Chem 281:29830-9.
44. Chen C and Beattie GA. 2007. Characterization of the osmoprotectant transporter OpuC from *Pseudomonas syringae* and demonstration that cystathionine-beta-synthase domains are required for its osmoregulatory function. J Bacteriol 189:6901-12.
45. Lucht JM and Bremer E. 1994. Adaptation of *Escherichia coli* to high osmolarity environments: osmoregulation of the high-affinity glycine betaine transport system proU. FEMS Microbiol Rev 14:3-20.
46. Gul N and Poolman B. 2013. Functional reconstitution and osmoregulatory properties of the ProU ABC transporter from *Escherichia coli.* Mol Membr Biol 30:138-48.
47. Lamark T, Kaasen I, Eshoo MW, Falkenberg P, McDougall J, Strom AR. 1991. DNA sequence and analysis of the bet genes encoding the osmoregulatory choline-glycine betaine pathway of *Escherichia coli.* Mol Microbiol 5:1049-64.
48. Ziegler C, Bremer E, Kramer R. 2010. The BCCT family of carriers: from physiology to crystal structure. Mol Microbiol 78:13-34.
49. Bailey TL, Elkan C. 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol 2:28-36.
50. Czech L, Hermann L, Stoveken N, Richter AA, Hoppner A, Smits SHJ, Heider J, Bremer E. 2018. Role of the Extremolytes Ectoine and Hydroxyectoine as Stress Protectants and Nutrients: Genetics, Phylogenomics, Biochemistry, and Structural Analysis. Genes (Basel) 9.
51. Shao Z, Deng W, Li S, He J, Ren S, Huang W, Lu Y, Zhao G, Cai Z, Wang J. 2015. GlnR-Mediated Regulation of *ectABCD* transcription expands the role of the GlnR regulon to osmotic stress management. J Bacteriol 197:3041-7.
52. Pastor JM, Bernal V, Salvador M, Argandona M, Vargas C, Csonka L, Sevilla A, Iborra JL, Nieto JJ, Canovas M. 2013. Role of central metabolism in the osmoadaptation of the halophilic bacterium *Chromohalobacter salexigens*. J Biol Chem 288:17769-81.

53. Mustakhimov, II, Reshetnikov AS, Glukhov AS, Khmelenina VN, Kalyuzhnaya MG, Trotsenko YA. 2010. Identification and characterization of EctR1, a new transcriptional regulator of the ectoine biosynthesis genes in the halotolerant methanotroph *Methylomicrobium alcaliphilum* 20Z. J Bacteriol 192:410-7.
54. Mustakhimov, II, Reshetnikov AS, Khmelenina VN, Trotsenko YA. 2009. EctR--a novel transcriptional regulator of ectoine biosynthesis genes in the haloalcaliphilic methylotrophic bacterium *Methylophaga alcalica*. Dokl Biochem Biophys 429:305-8.
55. Mustakhimov, II, Reshetnikov AS, Fedorov DN, Khmelenina VN, Trotsenko YA. 2012. Role of EctR as transcriptional regulator of ectoine biosynthesis genes in *Methylophaga thalassica*. Biochemistry (Mosc) 77:857-63.
56. Perera IC, Grove A. 2010. Molecular mechanisms of ligand-mediated attenuation of DNA binding by MarR family transcriptional regulators. J Mol Cell Biol 2:243-54.
57. Sulavik MC, Gambino LF, Miller PF. 1995. The MarR repressor of the multiple antibiotic resistance (mar) operon in *Escherichia coli*: prototypic member of a family of bacterial regulatory proteins involved in sensing phenolic compounds. Mol Med 1:436-46.
58. Cohen SP, Hachler H, Levy SB. 1993. Genetic and functional analysis of the multiple antibiotic resistance (mar) locus in *Escherichia coli*. J Bacteriol 175:1484-92.
59. Hong M, Fuangthong M, Helmann JD, Brennan RG. 2005. Structure of an OhrR-ohrA operator complex reveals the DNA binding mechanism of the MarR family. Mol Cell 20:131-41.
60. Kumarevel T, Tanaka T, Umehara T, Yokoyama S. 2009. ST1710-DNA complex crystal structure reveals the DNA binding mechanism of the MarR family of regulators. Nucleic Acids Res 37:4723-35.
61. Dolan KT, Duguid EM, He C. 2011. Crystal structures of SlyA protein, a master virulence regulator of *Salmonella*, in free and DNA-bound states. J Biol Chem 286:22178-85.
62. Hao Z, Lou H, Zhu R, Zhu J, Zhang D, Zhao BS, Zeng S, Chen X, Chan J, He C, Chen PR. 2014. The multiple antibiotic resistance regulator MarR is a copper sensor in *Escherichia coli*. Nat Chem Biol 10:21-8.
63. Deochand DK, Grove A. 2017. MarR family transcription factors: dynamic variations on a common scaffold. Crit Rev Biochem Mol Biol 52:595-613.
64. Reshetnikov AS, Khmelenina VN, Mustakhimov, II, Kalyuzhnaya M, Lidstrom M, Trotsenko YA. 2011. Diversity and phylogeny of the ectoine biosynthesis genes in aerobic, moderately halophilic methylotrophic bacteria. Extremophiles 15:653-63.
65. Whitaker WB, Parent MA, Naughton LM, Richards GP, Blumerman SL, Boyd EF. 2010. Modulation of responses of *Vibrio parahaemolyticus* O3:K6 to pH and temperature stresses by growth at different salt concentrations. Appl Environ Microbiol 76:4720-9.
66. Makino K, Oshima K, Kurokawa K, Yokoyama K, Uda T, Tagomori K, Iijima Y, Najima M, Nakano M, Yamashita A, Kubota Y, Kimura S, Yasunaga T, Honda T, Shinagawa H, Hattori M, Iida T. 2003. Genome sequence of *Vibrio parahaemolyticus*: a pathogenic mechanism distinct from that of *V. cholerae*. Lancet 361:743-9.
67. Horton RM HH, Ho SN, Pullen JK, Pease LR. 1989. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. Gene 77:61-68.
68. Gibson DG. 2011. Enzymatic assembly of overlapping DNA fragments. Methods Enzymol 498:349-61.

69. Pfaffl MW. 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 29:e45.
70. Kalburge SS, Carpenter MR, Rozovsky S, Boyd EF. 2017. Quorum Sensing Regulators Are Required for Metabolic Fitness in *Vibrio parahaemolyticus*. Infect Immun 85:e00930-16.
71. Haardt M, Kempf B, Faatz E, Bremer E. 1995. The osmoprotectant proline betaine is a major substrate for the binding-protein-dependent transport system ProU of *Escherichia coli* K-12. Mol Gen Genet 246:783-6.
72. Sullivan MJ, Petty NK, Beatson SA. 2011. Easyfig: a genome comparison visualizer. Bioinformatics 27:1009-10.
73. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, Lepore R, Schwede T. 2018. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 46:W296-w303.
74. Guex N, Peitsch MC, Schwede T. 2009. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. Electrophoresis 30 Suppl 1:S162-73.
75. Bienert S, Waterhouse A, de Beer TA, Tauriello G, Studer G, Bordoli L, Schwede T. 2017. The SWISS-MODEL Repository-new features and functionality. Nucleic Acids Res 45:D313-d319.
76. Benkert P, Biasini M, Schwede T. 2011. Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics 27:343-50.
77. Bertoni M, Kiefer F, Biasini M, Bordoli L, Schwede T. 2017. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. Sci Rep 7:10480.
78. Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673-80.
79. Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406-25.
80. Felsenstein J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39:783-791.
81. Jones DT, Taylor WR, Thornton JM. 1992. The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci 8:275-82.
82. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol Biol Evol 35:1547-1549.
83. Dehio C, Meyer M. 1997. Maintenance of broad-host-range incompatibility group P and group Q plasmids and transposition of Tn5 in *Bartonella henselae* following conjugal plasmid transfer from *Escherichia coli*. J Bacteriol 179:538-40.
84. Philippe N, Alcaraz JP, Coursange E, Geiselmann J, Schneider D. 2004. Improvement of pCVD442, a suicide plasmid for gene allele exchange in bacteria. Plasmid 51:246-55.
85. Kovach ME, Phillips RW, Elzer PH, Roop RM, 2nd, Peterson KM. 1994. pBBR1MCS: a broad-host-range cloning vector. Biotechniques 16:800-2.

765 86. Karunakaran R, Mauchline TH, Hosie AH, Poole PS. 2005. A family of promoter probe
766 vectors incorporating autofluorescent and chromogenic reporter proteins for studying
767 gene expression in Gram-negative bacteria. Microbiology 151:3249-56.

768

769

770

771 **Table 1. Strains and Plasmids**

| Strain | Genotype or description | Reference or Source |
|---|---|---|
| *Vibrio parahaemolyticus* | | |
| RIMD2210633 | O3:K6 clinical isolate, Str$^r$ | (65, 66) |
| Δ*cosR* | RIMD2210633 Δ*cosR* (VP1906), Str$^r$ | (17) |
| Δ*betI* | RIMD2210633 Δ*betI* (VPA1114), Str$^r$ | This study |
| SSK2516 (Δ*opaR*) | RIMD2210633 Δ*opaR* (VP2516), Str$^r$ | (70) |
| *Escherichia coli* | | |
| DH5α λ*pir* | Δ*lac pir* | ThermoFisher Scientific |
| β2155 λ*pir* | Δ*dapA*::*erm pir* for bacterial conjugation | (83) |
| BL21(DE3) | Expression strain | ThermoFisher Scientific |
| MKH13 | MC4100 (Δ*betTIBA*) Δ(*putPA*)*101* Δ(*proP*)*2* Δ(*proU*); Sp$^r$ | (71) |
| **Plasmids** | | |
| pDS132 | Suicide plasmid; Cm$^R$Cm$^r$; SacB | (84) |
| pBBR1MCS | Expression vector; *lacZ* promoter; Cm$^r$ | (85) |
| pBBR*cosR* | pBBR1MCS harboring full-length *cosR* (VP1906) | (17) |
| pRU1064 | promoterless-*gfp*UV, Amp$^r$, Tet$^r$, IncP origin | (86) |
| pRUP*ectA* | pRU1064 with P*ectA-gfp,* Amp$^r$, Tet$^r$ | (17) |
| pRUP*betI* | pRU1064 with P*betI-gfp,* Amp$^r$, Tet$^r$ | This study |
| pRUP*bcct1* | pRU1064 with P*bcct1-gfp,* Amp$^r$, Tet$^r$ | This study |
| pRUP*bcct3* | pRU1064 with P*bcct3-gfp,* Amp$^r$, Tet$^r$ | This study |
| pRUP*proV1* | pRU1064 with P*proV-gfp,* Amp$^r$, Tet$^r$ | This study |
| pRUP*cosR* | pRU1064 with P*cosR-gfp,* Amp$^r$, Tet$^r$ | (17) |
| pET28a+ | Expression vector, 6xHis; Kan$^r$ | Novagen |
| pET*cosR* | Pet28a+ harboring *cosR*, Kan$^r$ | (17) |

772

773

774 **Table 2.** ***Primers used in this study***

| Use | Sequence (5'-3') | bp |
|---|---|---|
| **Mutant** | | |
| VPbetIA | gcttcttctagaggtaccgcatgcGCCAGTTTTATGTGCTCACC | 580 |
| VPbetIB | atattttatgagaCATCCCCACCTTTGGCATTTTG | |
| VPbetIC | gatgcctgaaCTCGACAAGCAGCTAACG | 688 |
| VPbetID | ggagagctcgatatcgcatgcTCTGCCCTACCCGGTAATC | |
| VPbetIFLFwd | AGCATAGCACAATAAGAGTCG | 1895 |
| VPbetIFLRev | CCTGATTCGCCAGTGAACGA | |
| **EMSA** | | |
| VPbetIFwdA | CGGTTTTCTGATTTCAGGC | 125 |
| VPbetIRevA | CTTTTAATGATAAATCGTTTGAGTTCG | |
| VPbetIFwdB | ATGCCAAAAATTTAGTTCGAAC | 112 |
| VPbetIRevB | GGTCTTTGAATGGATGGTAGGG | |
| VPbetIFwdC | CCCTACCATCCATTCAAAGACC | 142 |
| VPbetIRevC | CTAAGGCTTCTACATTGCTTTC | |
| VPbetIFwdD | GAAAGCAATGTAGAAGCCTTAG | 202 |
| VPbetIRevD | GAACTTGGATATGCGTCCATT | |
| VPbetIFwdE | AATGGACGCATATCCAAGTTC | 158 |
| VPbetIRevE | AGCATAGCACAATAAGAGTCG | |
| VPbcct1FwdA | ACCGCAAACTTCCCGATC | 120 |
| VPbcct1RevA | CGGTATTCAGTACAAAAGAA | |
| VPbcct1FwdB | TTCTTTTGTACTGAATACCG | 110 |
| VPbcct1RevB | TGTCTTCAACTCACAAGAAT | |
| VPbcct1FwdC | ATTCTTGTGAGTTGAAGACA | 101 |
| VPbcct1RevC | AGCGAATTTTATCACCAATCACA | |
| VPbcct3FwdA | CGCTTTTTGTAATGCAAATTACC | 107 |
| VPbcct3RevA | CCCGTGAAAGCGGAAGATC | |
| VPbcct3FwdB | GATCTTCCGCTTTCACGGG | 108 |
| VPbcct3RevB | TCTATACCCTTTGTCATCGTTCCTC | |
| VPcosRFwdA | CAAATCTCCACACCATTAATTAG | 105 |
| VPcosRRevA | CGTCTTTGGTGATTTCTTTTTATTCG | |
| VPcosRFwdB | GCGAATAAAAAGAAATCACCAAAGACG | 142 |
| VPcosRRevB | CCAATTTTTTCATCCAGTCTGTAGGG | |
| VPproU1FwdA | TCTTTATTCCATGCGTTG | 160 |
| VPproU1RevA | AGAGGCAGAAAGAACAGTGAA | |
| VPproU1FwdB | TTCACTGTTCTTTCTGCCTCT | 134 |
| VPproU1RevB | GGTTATGAATGTGTTCGTTTGT | |
| VPproU1FwdC | ACAAACGAACACATTCATAACC | 108 |
| VPproU1RevC | TGGCTTGGCTTATTGGTGTTC | |
| VPproU1FwdD | GAACACCAATAAGCCAAGCCA | 109 |
| VPproU1RevD | GGGATCCATGTTAATTGTCCTTTG | |
| VPbcct2Fwd | ACCGAGACATGCCAATTTCTG | 233 |
| VPbcct2Rev | CGGTGCTCACGAATAATCTCC | |
| VPbcct4Fwd | AGAACAGGTTGGCTCAATGT | 244 |
| VPbcct4Rev | TTCCCCTCACATCAAGTCG | |
| **Expression** | | |
| PbetIFwd | TCTAAGCTTGCATAGCACAATAAGAGTCGC | 594 |
| PbetIRev | TATACTAGTTTTGCGTCCTTGTTATTTTTAATTG | |
| Pbcct1Fwd | tagatagagagagagagagaAAACCGCAAACTTCCCGATC | 278 |

| | | |
|---|---|---|
| Pbcct1Rev | actcatttttttcttcctccaCAATCACAAATTTATGCAAAAATGAC | |
| Pbcct3Fwd | tagatagagagagagagagaAATTTTTTCATCCAGTCTGTAGG | 397 |
| Pbcct3Rev | actcatttttttcttcctccaCGTTCCTCTCTATTTTTGTATTATTTTTTTC | |
| PproU1Fwd | tagatagagagagagagagaTCTTTATTCCATGCGTTG | 438 |
| PproU1Rev | actcatttttttcttcctccaGTTAATTGTCCTTTGTTATGTG | |
| PcosRFwd | tagatagagagagagagagaCGTTCCTCTCTATTTTTGTATTATTTTTTC | 397 |
| PcosRRev | cggccgctctagaactagtgTTATTCTGGTTTGGTGATG | |
| **RT-PCR primers** | | |
| VPbcct1Fwd | GTTCGGTCTTGCGACTTCTC | 246 |
| VPbcct1Rev | CCCATCGCAGTATCAAAGGT | |
| VPbcct2Fwd | AACAAAGGGTTGCCACTGAC | 167 |
| VPbcct2Rev | TTCAAACCTGTTGCTGCTTG | |
| VPbcct3Fwd | TGGACGGTATTCTACTGGGC | 202 |
| VPbcct3Rev | CGCCTAACTCGCCTACTTTG | |
| VPectAFwd | TCGAAAGGGAAGCGCTGAG | 125 |
| VPectARev | AGTGCTGACTTGGCCATGAT | |
| VPasp_ectFwd | CGATGATTCCATTCGCGACG | 126 |
| VPasp_ectRev | GTCATCTCACTGTAGCCCCG | |
| VPproV1Fwd | GCATCGTTTCTCTCGACTCC | 163 |
| VPproV1Rev | TGCTCATCGACTACTGGCAC | |
| VPAbcct4Fwd | CAAGGCGTAGGCCGCATGGT | 234 |
| VPAbcct4Rev | ACCGCCCACGATGCTGAACC | |
| VPAbetIFwd | ACTTCGGTGGTAAGCATGGG | 138 |
| VPAbetIRev | TGCCGTCAATAATGGCGTTG | |
| VPAbetBFwd | TGGAAATCAGCACCAGCACT | 160 |
| VPAbetBRev | TCTGCCCTACCCGGTAATCA | |
| VPAproXFwd | TTCCTTGGTAACTGGATGCC | 216 |
| VPAproXRev | ATCGTTACCTGGTTCGATGC | |
| VPAproWFwd | ATCACAGCGGCACTGGCTTGG | 190 |
| VPAproWRev | GGCGATGCGCTGCCATGATC | |
| 16SFwd | ACCGCCTGGGGAGTACGGTC | 234 |
| 16SRev | TTGCGCTCGTTGCGGGACTT | |

775

776

777

778

779

780

781

782

783

784 **Figure legends**

785 **Figure 1.** RNA was isolated from RIMD2210633 after growth in M9G 1%NaCl and M9G
786 3%NaCl at an $OD_{595}$ of 0.45. Expression analysis of **(A)** *ectA, asp_ect,* **(B)** *betI, betB, proX2,*
787 *proW2* **(C)** *bcct1, bcct2, bcct3, bcct4* and *proV1* by quantitative real time PCR (qPCR). 16S was
788 used for normalization. Expression levels shown are levels in M9G1% relative to M9G3%. Mean
789 and standard error of two biological replicates are shown. Statistics were calculated using a
790 Student's t-test (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$).

791 **Figure 2**. RNA was isolated from RIMD2210633 and Δ*cosR* after growth in M9G 1%NaCl at an
792 $OD_{595}$ of 0.45. Expression analysis of **(A)** *ectA, asp_ect,***(B)** *betI, betB, proX2, proW2* **(C)** *bcct1,*
793 *bcct2, bcct3, bcct4* and *proV1* by qPCR. 16S was used for normalization. Expression levels
794 shown are levels in Δ*cosR* relative to wild-type. Mean and standard error of two biological
795 replicates are shown. Statistics were calculated using a Student's t-test (*, $P < 0.05$; **, $P <$
796 0.01).

797 **Figure 3. (A)** The regulatory region of *betIBAproXWV* was divided into five probes for EMSAs,
798 P*betI* A-E, 125-bp, 112-bp, 142-bp, 202-bp and 158-bp, respectively. The regulatory region used
799 for the GFP reporter assay is indicated with a bracket. **(B)** An EMSA was performed with
800 purified CosR-His (0 to 0.62 μM) and 30 ng of each P*betI* probe, with DNA:protein molar ratios
801 of 1:0, 1:1, 1:5, and 1:10. **(C)** A $P_{betI}$-*gfp* reporter assay was performed in *E. coli* strain MKH13
802 containing an expression plasmid with full-length *cosR* (p*cosR*). Specific fluorescence of the
803 CosR-expressing strain was compared to a strain harboring empty expression vector. Mean and
804 standard deviation of two biological replicates are shown. Statistics were calculated using a
805 Student's t-test (***, $P < 0.001$).

806 **Figure 4. (A)** The regulatory region of *bcct1* was divided into three similarly sized probes for
807 EMSAs, P*bcct1* A-C, 120-bp, 110-bp, and 101-bp, respectively. The regulatory region used for
808 the GFP reporter assay is indicated with a bracket. **(B)** An EMSA was performed with purified
809 CosR-His (0 to 0.69 µM) and 30 ng of P*bcct1* probe with DNA:protein molar ratios of 1:0, 1:1,
810 1:5, and 1:10. **(C)** A $P_{bcct1}$-*gfp* reporter assay was performed in *E. coli* strain MKH13 containing
811 an expression plasmid with full-length *cosR* (p*cosR*). Specific fluorescence of the CosR-
812 expressing strain was compared to a strain harboring empty expression vector (pBBR1MCS).
813 Mean and standard deviation of two biological replicates are shown. Statistics were calculated
814 using a Student's t-test (**, $P < 0.01$).

815 **Figure 5. (A)** A 196-bp portion of the regulatory region of *bcct3* was split into two probes for
816 EMSAs, P*bcct3* A and B, 108-bp and 107-bp, respectively. The regulatory region used for the
817 GFP reporter assay is indicated with a bracket. **(B)** An EMSA was performed with purified
818 CosR-His (0 to 0.65 µM) and 30 ng of P*bcct3* probe with DNA:protein molar ratios of 1:0, 1:1,
819 1:5, and 1:10. **(C)** $P_{bcct3}$-*gfp* reporter assay was performed in *E. coli* strain MKH13 containing an
820 expression plasmid with full-length *cosR* (p*cosR*). Specific fluorescence of the CosR-expressing
821 strain was compared to a strain harboring empty expression vector (pBBR1MCS). Mean and
822 standard deviation of two biological replicates are shown. Statistics were calculated using a
823 Student's t-test (**, $P < 0.01$). **(D)** Diagrams indicating the regulatory regions of *bcct2* and *bcct4*
824 that were used as probes in a CosR EMSA. **(E)** An EMSA was performed with CosR-His (0 to
825 0.18 µM) and probes of the regulatory regions of *bcct2* and *bcct4*. Each lane contains 30 ng of
826 DNA and DNA:protein molar ratios of 1:0, 1:1, 1:5, and 1:10.

827 **Figure 6. (A)** The 447-bp regulatory region of the *proV1* gene was divided into four probes for
828 EMSAs, P*proV1* A-D, 160-bp, 134-bp, 108-bp and 109-bp, respectively. The regulatory region

used for the GFP reporter assay is indicated with a bracket. **(B)** An EMSA was performed with purified CosR-His (0 to 0.64 µM) and 30 ng of each P*proV1* probe with DNA:protein molar ratios of 1:0, 1:1, 1:5, and 1:10. **(C)** A reporter assay was conducted in *E. coli* MKH13 harboring the $P_{proV1}$-gfp reporter plasmid and the expression plasmid p*cosR*. Specific fluorescence of the CosR-expressing strain was compared to an empty vector strain. Mean and standard deviation of two biological replicates are shown. Statistics were calculated using a Student's t-test (*, $P < 0.05$).

**Figure 7. (A)** A 220-bp section of the regulatory region of *cosR* was split into two similarly sized probes for EMSAs, P*cosR* A and B, 105-bp and 142-bp, respectively. The regulatory region used for the GFP reporter assay is indicated with a bracket. **(B)** An EMSA was performed with increasing concentrations of purified CosR-His (0 to 0.66 µM) and 30 ng of each probe with DNA:protein molar ratios of 1:0, 1:1, 1:5, and 1:10. **(C)** A $P_{cosR}$-*gfp* reporter assay was performed in *E. coli* strain MKH13 the p*cosR* expression plasmid. Specific fluorescence of the CosR-expressing strain was compared to a strain harboring empty expression vector. Mean and standard deviation of two biological replicates are shown. **(D)** A $P_{cosR}$-*gfp* reporter assay was performed in *V. parahaemolyticus* WT and Δ*cosR* mutant strains. Mean and standard deviation of three biological replicates are shown.

**Figure 8. (A)** Expression of a $P_{betI}$-*gfp* transcriptional fusion reporter in wild-type and a Δ*betI* mutant. Relative fluorescence intensity (RFU) and $OD_{595}$ were measured after growth in **(A)** M9G3% or **(B)** M9G3% with the addition of choline. Specific fluorescence was calculated by dividing RFU by OD. Mean and standard deviation of two biological replicates are shown. Statistics were calculated using a Student's t-test (*, $P < 0.05$). **(C)** A reporter assay was conducted in *E. coli* MKH13 using the $P_{betI}$-*gfp* reporter plasmid and an expression plasmid with

852 full-length *betI* (p*betI*). The specific fluorescence was calculated and compared to a strain with
853 an empty expression vector (pBBR1MCS). Mean and standard deviation of two biological
854 replicates are shown. Statistics were calculated using a Student's t-test (***, $P < 0.001$).

855 **Figure 9. (A)** Expression of a $P_{betI}$-*gfp* transcriptional fusion reporter in wild-type and Δ*opaR*
856 mutant strains. Relative fluorescence intensity (RFU) and $OD_{595}$ were measured after growth in
857 M9G3%. Specific fluorescence was calculated by dividing RFU by OD. Mean and standard
858 deviation of two biological replicates are shown. Statistics were calculated using a one-way
859 ANOVA with a Tukey-Kramer *post hoc* test (**, $P < 0.01$). **(B)** An EMSA was performed with
860 30 ng of each P*betI* probe A-E utilized previously in the CosR EMSA and purified OpaR protein
861 (0 to 0.41 µM) in various DNA:protein molar ratios (1:0, 1:1, and 1:5 for probe A; 1:0, 1:1, 1:10
862 for all other probes).

863 **Figure 10. (A)** A CosR DNA binding motif was created using MEME analysis with sequences of
864 the EMSA probes that were bound by CosR. **(B)** An alignment of the motif sites found in each
865 sequence used for MEME analysis with corresponding p-values.

866 **Figure 11.** Schematic of the genomic context of CosR homologs from select *Vibrionaceae*
867 species. Open reading frames are designated by arrows.

868 **Figure 12.** Phylogenetic and distribution analysis of CosR. The phylogeny of CosR was inferred
869 using the Neighbor-Joining method. The optimal tree with the sum of branch length =
870 5.19649696 is shown. The percentage of replicate trees in which the associated taxa clustered
871 together in the bootstrap test (1000 replicates) are shown next to the branches. The tree is drawn
872 to scale, with branch lengths in the same units as those of the evolutionary distances used to infer
873 the phylogenetic tree. Brackets represent groups based on genus.

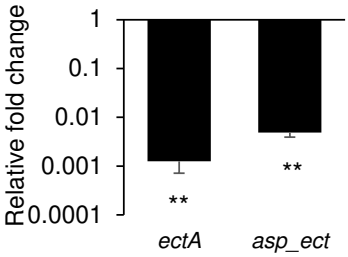
874 **Figure 13.** A model of CosR regulation of the osmotic stress response and its known regulators.
875 Solid arrows indicate direct positive regulation, dashed arrows, indirect positive regulation, solid
876 hammers represent direct repression and dashed hammers, indirect repression. Transporters
877 colored purple are osmotic responsive. The quorum sensing regulators OpaR and AphA were
878 shown in previous studies to directly and indirectly positively regulate CosR, respectively, and in
879 addition, directly regulate ectoine and glycine betaine biosynthesis operons.

880 **Table 1.** Strains and plasmids used in this study

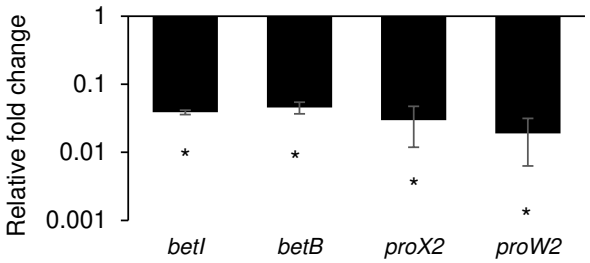
881 **Table 2.** Primers used in this study

882

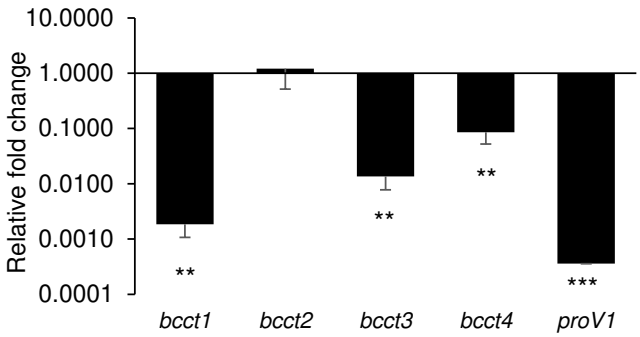
*A.* *ectA* and *asp_ect* expression in M9G 1%NaCl relative to M9G 3%NaCl



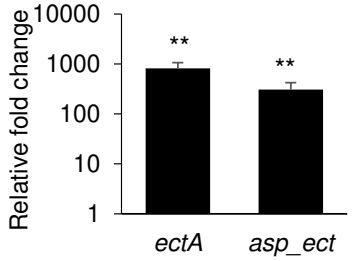

**B.** *betIBAproXWV* expression in M9G 1%NaCl relative to M9G 3%NaCl



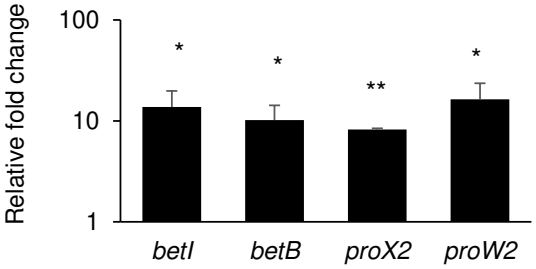

**C.** *bcct1-bcct4* & *proV1* expression in M9G 1%NaCl relative to M9G 3%NaCl



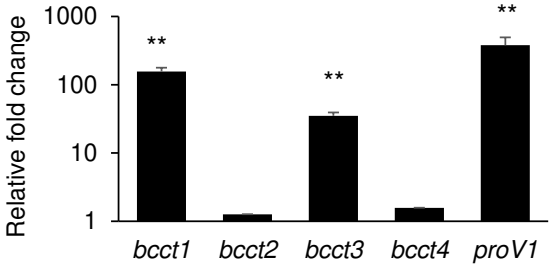
**A.** Expression of *ectA* and *asp_ect* in Δ*cosR* relative to WT in M9G1%NaCl



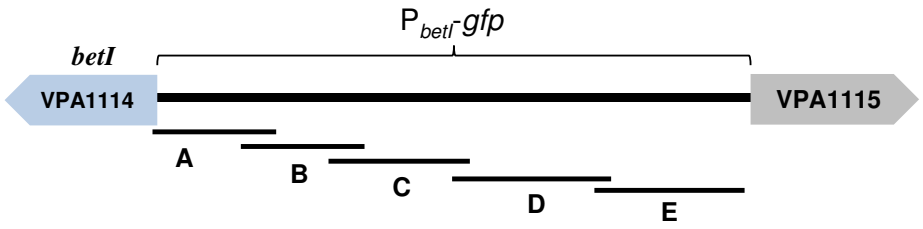

**B.** Expression of *betIBAproXWV* in Δ*cosR* relative to WT in M9G1%NaCl



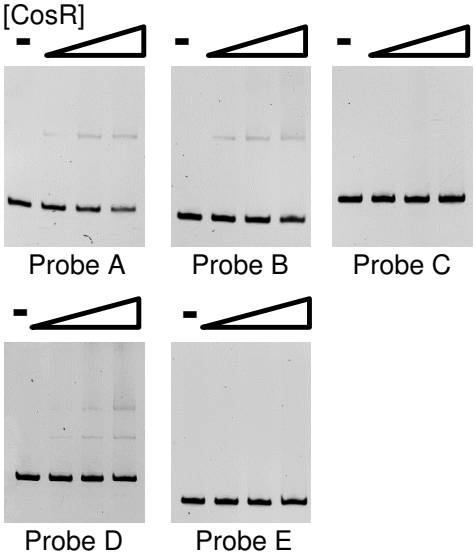

**C.** Expression of *bccts* and *proV1* in Δ*cosR* relative to WT in M9G1%NaCl



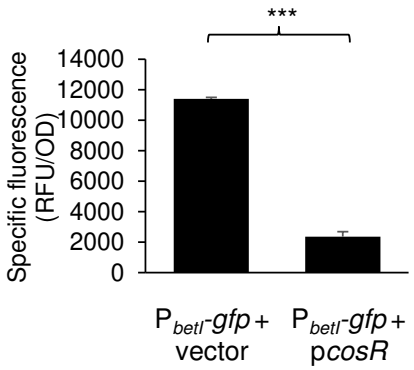
**A.** *betIBAproXWV* regulatory region



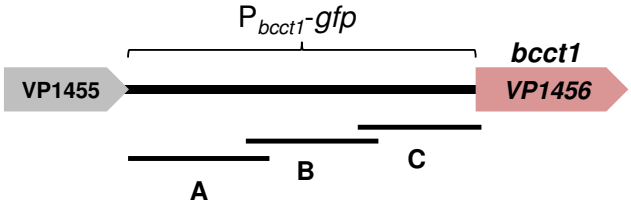

**B.** P*betI* CosR EMSA



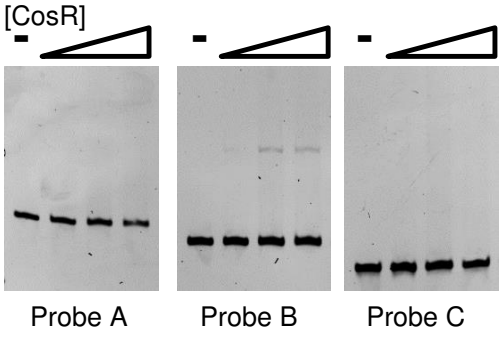

**C.** $P_{betI}$-*gfp* reporter assay in *E. coli*



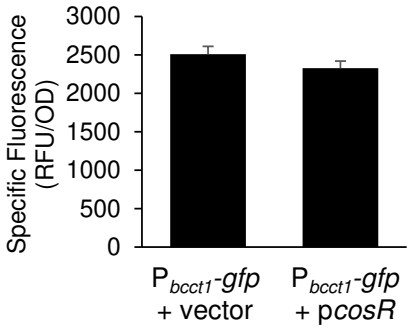
**A.** *bcct1* regulatory region

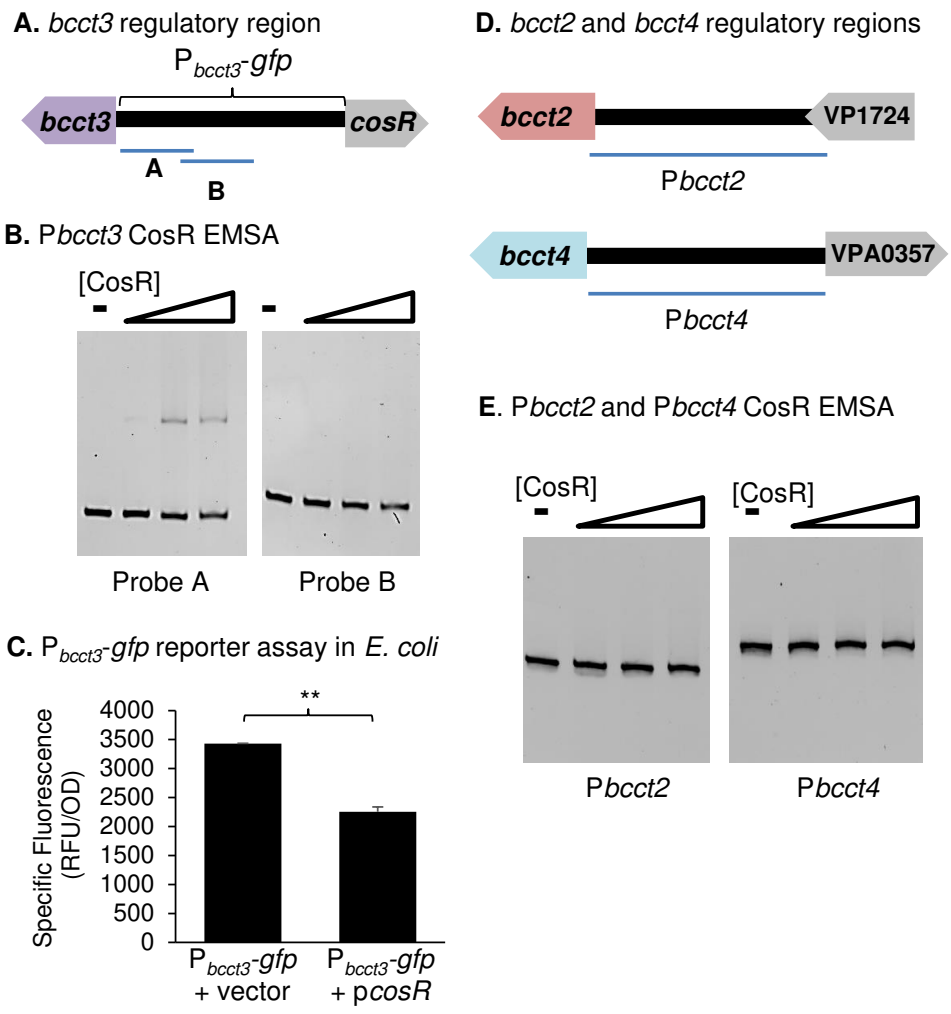


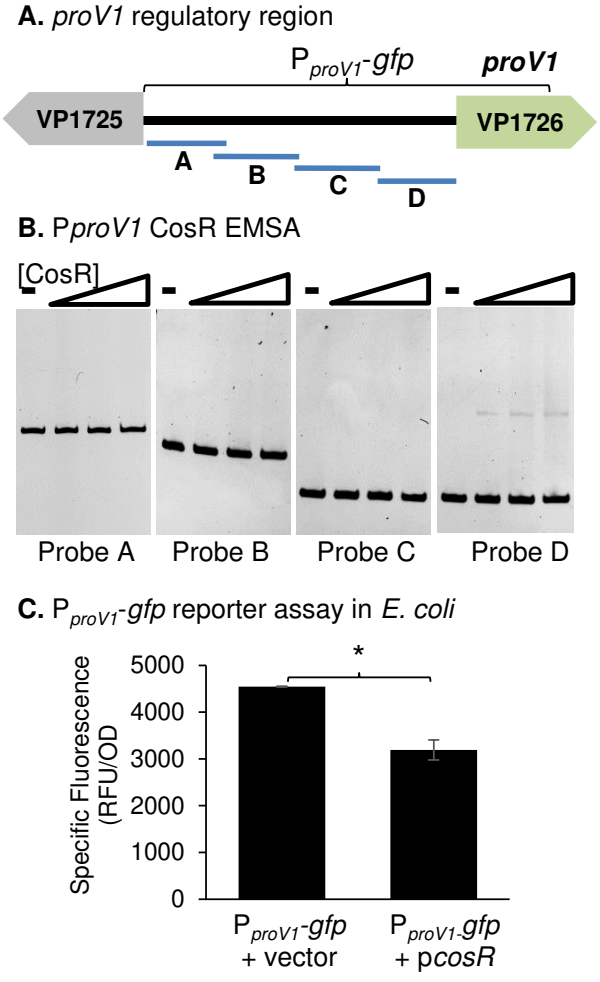

**B.** P*bcct1* CosR EMSA

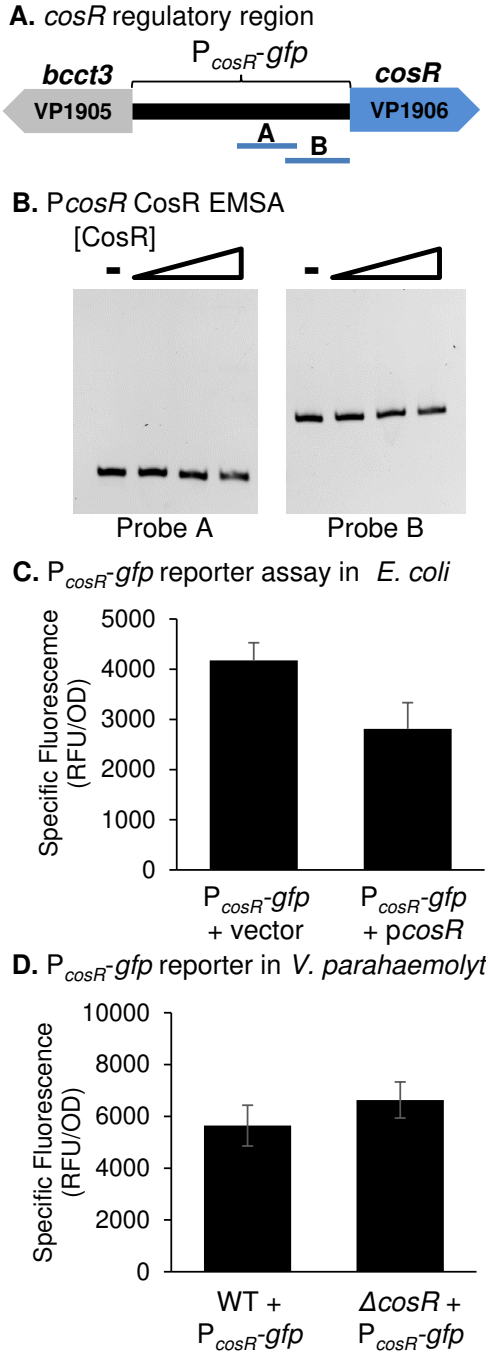



**C.** P *bcct1*-*gfp* reporter assay in *E. coli*

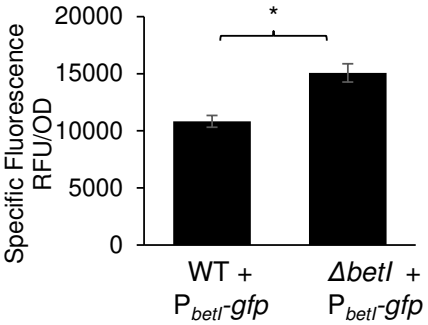




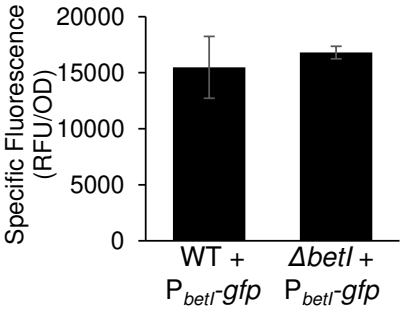




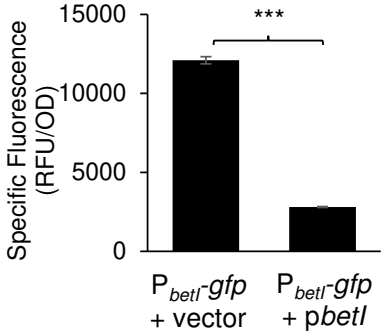
**A.** $P_{betI}$-*gfp* reporter assay in *V. parahaemolyticus*



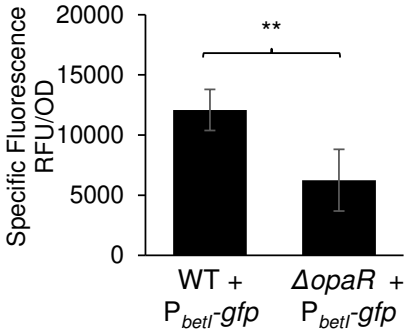

**B.** $P_{betI}$-*gfp* reporter assay in *V. parahaemolyticus* with choline



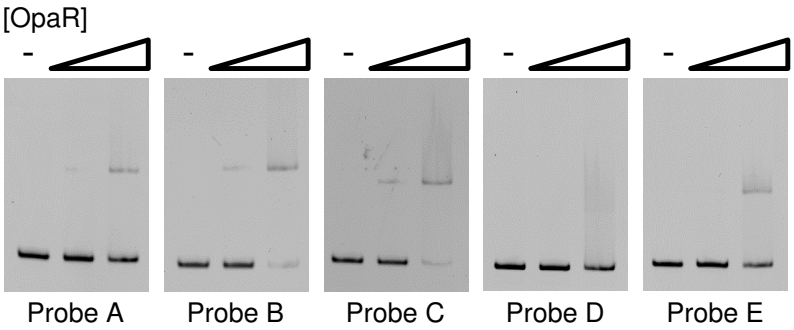

**C.** $P_{betI}$-*gfp* reporter assay in *E. coli*



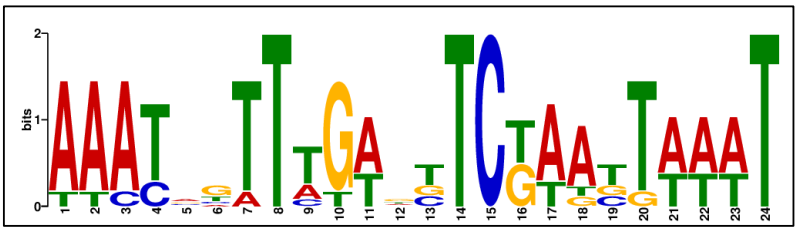
**A.** $P_{betI}$-*gfp* reporter assay in *V. parahaemolyticus*




**B.** P*betI* OpaR EMSA



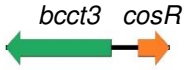
**A.** CosR DNA binding motif from MEME analysis of bound EMSA probes



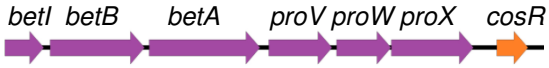

**B.** Alignment of motif location in regulatory regions of osmotic stress response genes

| Probe | p-value | Sites |
|---|---|---|
| P*ectA* A | 1.40e-8 | TTTGAAATAC ATATCCTTTGACGTCTAATTAAAT TTCGTATTTA |
| P*ectA* B | 9.81e-10 | CCTAAATTTA AAATAGTTTGAACTCTAATTTATT AAAGCATCAT |
| P*betI* A | 4.37e-13 | TTTTAATGAT AAATCGTTTGAGTTCGAACTAAAT TTTTGGCATT |
| P*betI* B | 7.68e-7 | AAATTGCATT AACTTGTTTTATTTCTTGTTTTTT GGGCGGGTCT |
| P*betI* D | 2.85e-8 | AACACAACAC AAATAATTAGTGTTCGATGTAAAT TTTTTCTAAG |
| P*bcct1* B | 6.18e-9 | CAATTTCGAA AAATAGTTAGATGTCGTAGTTTAT TCTTGTGAGT |
| P*bcct3* A | 1.14e-7 | ACGGGATGTA AAACGTATCGTGTTCGAACGATTT TTTGTTTGGT |
| P*proV1* D | 3.64e-8 | TATCACGAGT TAACGTTTTGTACTCTAATTAAAT TGATAAAAGT |

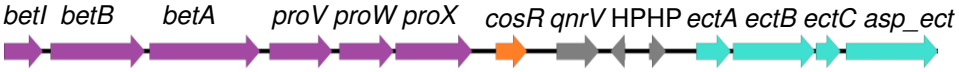
***Vibrio parahaemolyticus* and more than 50 *Vibrio* species**



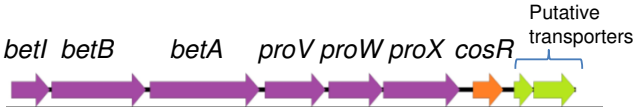

***V. cyclitrophicus/V. splendidus/V. crassotreae/V. lentus/V. celticus***



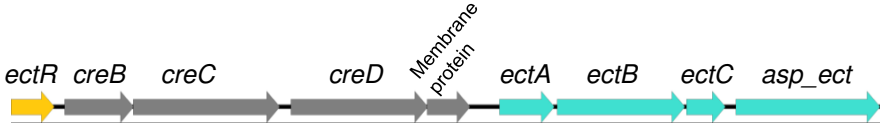

***Vibrio tasmaniensis/Vibrio sp.* MED222**




***A. wodanis* AWOD1/*A. wodanis* 06/09/160**




***A. wodanis* AWOD1/*A. wodanis* 06/09/160**




***A. fischeri* MJ11/*A. fischeri* ES114**