

1 Driving the expression of the *Salmonella enterica* sv Typhimurium flagellum using *flhDC*  
2 from *Escherichia coli* results in key regulatory and cellular differences.

3

4 Ayman Albanna<sup>1,2,6</sup>, Martin Sim<sup>1,2§</sup>, Paul A Hoskisson<sup>3</sup>, Colin Gillespie<sup>4</sup>, Christopher V.  
5 Rao<sup>5</sup> and Phillip Aldridge<sup>1,2</sup>

6

7 1: Centre for Bacterial Cell Biology, Baddiley Clark Building, Newcastle University,  
8 Richardson Road, Newcastle upon Tyne, United Kingdom, NE2 4AX. UK.

9 2: Institute for Cell and Molecular Biosciences, Newcastle University, Framlington Place,  
10 Newcastle upon Tyne, United Kingdom, NE2 4HH. UK.

11 3: Strathclyde Institute of Pharmacy and Biomedical Sciences, University of Strathclyde,  
12 Glasgow, United Kingdom G4 0RE. UK.

13 4: School of Mathematics & Statistics, Herschel Building, Newcastle University, Newcastle  
14 upon Tyne, United Kingdom, NE1 7RU. UK.

15 5: Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-  
16 Champaign, Urbana, Illinois, United States, 61801.

17 6: College of environmental science & technology, Mosul University, Iraq 41002

18

19 Running Title: *E. coli flhDC* function in *S. enterica*

20

21 \*Corresponding author. Centre for Bacterial Cell Biology, Baddiley Clark Building,  
22 Newcastle University, Richardson Road, Newcastle upon Tyne, United Kingdom, NE2  
23 4AX. UK. Phone: +44-191-2083218. Email: [phillip.aldridge@ncl.ac.uk](mailto:phillip.aldridge@ncl.ac.uk).

24

25 § Present address: Isomerase Therapeutics Ltd., Chesterford Research Park, Cambridge,  
26 CB10 1XL. UK.

27

28 **ABSTRACT**

29

30 The flagellar systems of *Escherichia coli* and *Salmonella enterica* exhibit a significant level  
31 of genetic and functional synteny. Both systems are controlled by the flagellar specific  
32 master regulator FlhD<sub>4</sub>C<sub>2</sub>. Since the early days of genetic analyses of flagellar systems it  
33 has been known that *E. coli flhDC* can complement a  $\Delta flhDC$  mutant in *S. enterica*. The  
34 genomic revolution has identified how genetic changes to transcription factors and / or  
35 DNA binding sites can impact the phenotypic outcome across related species. We were  
36 therefore interested in asking: using modern tools to interrogate flagellar gene expression  
37 and assembly, what would the impact be of replacing the *flhDC* coding sequences in *S.*  
38 *enterica* for the *E. coli* genes at the *flhDC* *S. enterica* chromosomal locus? We show that  
39 even though all strains created are motile, flagellar gene expression is measurably lower  
40 when *flhDC*<sub>EC</sub> are present. These changes can be attributed to the impact of FlhD<sub>4</sub>C<sub>2</sub> DNA  
41 recognition and the protein-protein interactions required to generate a stable FlhD<sub>4</sub>C<sub>2</sub>  
42 complex. Furthermore, our data suggests that in *E. coli* the internal flagellar FlhT regulatory  
43 feedback loop has a marked difference with respect to output of the flagellar systems. We  
44 argue due diligence is required in making assumptions based on heterologous expression  
45 of regulators and that even systems showing significant synteny may not behave in exactly  
46 the same manner.

47

48 **IMPORTANCE**

49 The bacterial motility organelle known as the flagellum is shared across many bacterial  
50 species. *Escherichia coli* and *Salmonella enterica* have underpinned our appreciation of  
51 how bacteria express and assemble the bacterial flagellum for over half a century. We  
52 show that even though the *E. coli* and *S. enterica* flagellar systems look genetically  
53 identical, they input regulatory signals into the flagellar system differently. Our conclusions

54 are based on experiments where we carefully transfer the master flagellar regulator from  
55 *E. coli* into the *S. enterica* chromosome and measure a range of outputs relating to  
56 flagellar gene expression, assembly and functional output.

57

58

## 59 INTRODUCTION

60

61 The flagellum in the enteric bacteria, *Escherichia coli* and *Salmonella enterica*, has been  
62 studied extensively for over fifty years and provides the canonical example for bacterial  
63 motility. These studies have revealed not only the complex structure of the enteric  
64 flagellum but also its role in host colonization, pathogenesis, and cellular physiology (1-4).  
65 In addition, these studies have identified many of the complex regulatory processes that  
66 coordinate the assembly and control of this exquisitely complex biological machine (3-5).

67

68 The flagellum in *E. coli* and *S. enterica* are structurally very similar and are often tacitly  
69 assumed to be effectively identical aside from differences in the filament structure.  
70 However, in the case of regulation, these assumptions are based more on sequence  
71 similarity rather than on actual experimental data (5) (6). Indeed, a number of studies have  
72 shown that these two systems are regulated in entirely different manners in response to  
73 environmental signals despite strong gene synteny. For example, many common *E. coli*  
74 strains are motile only during growth in nutrient-poor conditions whereas many common *S.*  
75 *enterica* strains are motile only during growth in nutrient-rich conditions (7). In addition, *E.*  
76 *coli* is more motile at 30°C than at 37°C whereas motility *S. enterica* is generally  
77 insensitive to these temperature differences (8). *E. coli flhDC* are transcribed from a single  
78 transcriptional start site that is responsive to OmpR, RcsB and CRP regulation, to name  
79 only a few regulatory inputs (8). In contrast *S. enterica flhDC* transcription is significantly  
80 more complex with up to 5 transcriptional start sites, albeit with only a subset being  
81 responsible for the majority of *flhDC* transcription (9).

82

83 Part of the problem is that different questions have been asked when studying the regulation  
84 of motility in these two bacterial species. Most studies in *E. coli* have focused on the

85 environmental signals and associate regulatory process that induce bacterial motility. In  
86 particular, they have focused on the processes that regulate the expression of the master  
87 flagellar regulator, FlhD<sub>4</sub>C<sub>2</sub> (8). Most studies in *S. enterica*, on the other hand, have focused  
88 on the regulatory processes that coordinate the assembly process following induction (4). In  
89 particular, they have focused on the downstream regulatory processes induced by FlhD<sub>4</sub>C<sub>2</sub>  
90 (3).

91

92 Despite differences in regulation, the protein subunits of master flagellar regulators, FlhC  
93 and FlhD, exhibit high sequence similarity sharing 94 and 92% identity, respectively,  
94 between *E. coli* and *S. enterica*. Given that modifications to transcription factors and/or  
95 promoter structure can lead to divergence in regulatory circuits (10), we were interested in  
96 how FlhD<sub>4</sub>C<sub>2</sub> functions in different genetic backgrounds? Previously, it was shown that *E.*  
97 *coli flhDC* can complement a  $\Delta flhDC$  mutant in *S. enterica*, suggesting that these proteins  
98 are functions identical in the two bacterial species (11). However, it is not clear whether  
99 they are regulated in the same manner. We, therefore, investigated the impact of replacing  
100 the native master regulator in *S. enterica* with the one from *E. coli*. Defining the impact of  
101 known FlhD<sub>4</sub>C<sub>2</sub> regulators such as ClpP, YdiV, FliT and FliZ on the two complexes  
102 suggest that these two species have adapted in how they perceive FlhD<sub>4</sub>C<sub>2</sub>. We argue that  
103 these phenotypic differences arise from adaptations *E. coli* and *S. enterica* have made  
104 during evolution to expand or modify cellular function with respect to movement within  
105 specific environmental niches.

106

107 **RESULTS**

108

109 *Orthologous flhDC from E. coli can functionally complement flhDC in S. enterica*

110

111 Given the similarities between the flagellar systems in *S. enterica* and *E. coli*, we sought to  
112 determine whether the FlhD<sub>4</sub>C<sub>2</sub> master regulator is functionally equivalent in these two  
113 species of bacteria. To test this hypothesis, we replaced the *flhDC* genes in *S. enterica*  
114 (*flhDC*<sub>SE</sub>) with the *flhDC* genes from *E. coli* (*flhDC*<sub>EC</sub>). The reason that we performed these  
115 experiments in *S. enterica* rather than *E. coli* was that the flagellar system is better  
116 characterized in the former, particularly with regards to transcriptional regulation. To avoid  
117 plasmid associated artefacts associated with the ectopic expression of *flhDC*, we replaced  
118 the entire *S. enterica flhDC* operon with the *flhDC* operon from *E. coli* at the native  
119 chromosomal locus (Figure S1).

120

121 We first tested whether *flhDC*<sub>EC</sub> was motile as determined using soft-agar motility plates.  
122 As shown in Figures 1A and B, these strains formed rings similar to the wild type. These  
123 results demonstrate that *flhDC*<sub>EC</sub> is functional in *S. enterica*. However, motility plates  
124 measure both motility and chemotaxis and do not provide any insights regarding possibly  
125 changes in the number of flagella per cell. To determine the impact *flhDC*<sub>EC</sub> had upon  
126 flagellar numbers we used a FliM-GFP fusion as a proxy for flagellar numbers (Figure 1C).  
127 When this fluorescent protein fusion is expressed in cells, it forms spots associated with  
128 nascent C-rings that loosely correlate with the number of flagella (12-14). By counting the  
129 number of spots per cell, we can determine the number of flagella made per cell. As  
130 shown Figure 1C, *flhDC*<sub>EC</sub> did not change flagellar numbers as compared to the wild type.  
131 These results demonstrate *flhDC*<sub>EC</sub> induces flagellar gene expression at similar levels as  
132 the wild type.

133

134 *flhDC* requires a specific transcription rate to maintain optimal flagellar numbers

135

136 The flagellar network in *S. enterica* contains a number of feedback loops to ensure that the  
137 cells regulate the number of flagella produced (4). One possibility is that these feedback  
138 loops mask any differences in FlhD<sub>4</sub>C<sub>2</sub>EC activity. To test this hypothesis, we replaced the  
139 native P<sub>flhD</sub> promoter with the tetracycline-inducible P<sub>tetA/tetR</sub> promoters. We then measured  
140 flagellar gene expression using a luciferase reporter system (15). In this case, a consistent  
141 and significant change in flagellar gene expression was observed when comparing  
142 FlhD<sub>4</sub>C<sub>2</sub>EC to FlhD<sub>4</sub>C<sub>2</sub>SE activity (Figure 2). Maximal expression of P<sub>flgA</sub> and P<sub>flhC</sub>, chosen to  
143 reflect flagellar gene expression at different stages of flagellar assembly (5), for both  
144 complexes was observed between 10 and 25 ng/ml of anhydrotetracycline, when *flhDC*  
145 transcription was from P<sub>tetA</sub> (Figure 2A and B). In contrast, P<sub>tetR</sub>, the weaker of the two  
146 tetracycline inducible promoters, reached a maximal output between 50 to 100 ng/ml  
147 anhydrotetracycline. In both scenarios the output for FlhD<sub>4</sub>C<sub>2</sub>EC control was lower than for  
148 the native FlhD<sub>4</sub>C<sub>2</sub>SE complex (Figure 2A and 2B).

149

150 We also measured the number of FliM-GFP foci at different anhydrotetracycline  
151 concentrations. P<sub>tetR</sub>::*flhDC* expression generated on average of approximately two FliM-  
152 foci per cell at 25 ng/ml of anhydrotetracycline for both FlhD<sub>4</sub>C<sub>2</sub> complexes (Figure 2C). In  
153 contrast, 5 ng/ml induction of the P<sub>tetA</sub>::*flhDC*EC strain was sufficient to generate typical  
154 FliM-foci numbers (approx. 8 flagellar foci per cell). Even with the strong decrease in  
155 average foci per cell at these levels of induction, the number of basal bodies observed is  
156 sufficient to allow motility at comparable levels in the motility agar assay (Figure S2).

157

158 *Replacement of flhC but not flhD in S. enterica with the E. coli orthologs affects motility*

159

160 The hetero-oligomeric regulator FlhD<sub>4</sub>C<sub>2</sub> is unusual in bacteria as the majority of  
161 transcriptional regulators are believed to be homo-oligomeric complexes. To determine the  
162 relative contributions of the two subunits, we individually replaced the *flhC* or *flhD* genes  
163 from *S. enterica* with their ortholog from *E. coli* (Figure S1). When we tested the two  
164 strains using motility plates, we found that motility was inhibited in the strain where *flhC*<sub>EC</sub>  
165 replaced the native *S. enterica flhC* (Figure 3A; blue bars), with an 88% reduction in  
166 swarm diameter when compared to WT *S. enterica*. The introduction of *flhD*<sub>EC</sub> compared  
167 to *flhD*<sub>EC</sub> or *flhD*<sub>SE</sub> produced swarms of a comparable size (Figure 3A; blue bars).

168

169 Using the dose-dependent inducible P<sub>tetA</sub> promoter(16) we observed that P<sub>tetA</sub> expression  
170 of *flhC*<sub>EC</sub> led to reduced P<sub>flgA</sub> transcription and strongly reduced P<sub>flhC</sub> transcription (Figure  
171 4). Strains expressing *flhD*<sub>EC</sub> in *S. enterica* showed a mild increase in P<sub>flgA</sub> gene  
172 expression and a similar response for P<sub>flhC</sub>, although these changes were not significant (P  
173 = 0.32) (Figure 4). These data suggest that the combination of FlhD<sub>SE</sub> and FlhC<sub>EC</sub>  
174 generates an inefficient FlhD<sub>4</sub>C<sub>2</sub> complex, resulting in reduced motility.

175

176 *Orthologous FlhC and FlhD interaction is species specific and a key determinant of*  
177 *promoter recognition by the FlhD<sub>4</sub>C<sub>2</sub> complex*

178

179 The results above demonstrate that *flhC*<sub>EC</sub> is not functionally identical to *flhC*<sub>ST</sub>. One  
180 possibility is that that FlhC<sub>EC</sub> is impaired in FlhD<sub>4</sub>C<sub>2</sub> for DNA-binding. Alternatively, the  
181 stability of the FlhD<sub>4</sub>C<sub>2</sub> complex is reduced in the *flhC*<sub>EC</sub> strain, leading to reduced FlhD<sub>4</sub>C<sub>2</sub>  
182 activity. To test these hypotheses, we purified all combinations of the FlhD<sub>4</sub>C<sub>2</sub> complex  
183 using affinity (Ni<sup>+</sup> and heparin) chromatography (Figure 5A). In each complex, FlhD was  
184 tagged with a carboxy-terminal hexa-histidine to facilitate affinity purification. Such



185 expression constructs have previously been used successfully to purify the FlhD<sub>4</sub>C<sub>2</sub>  
186 complex (17, 18). Using either Ni<sup>+</sup> affinity or heparin purification, we observed complete  
187 complex retrieval for three combinations (Figure 5A). FlhC recovery was less efficient in  
188 the FlhD<sub>SE</sub>/FlhC<sub>EC</sub> complex. In contrast, no FlhD<sub>SE</sub>/FlhC<sub>EC</sub> complex was recovered via  
189 Heparin purification, used to mimic DNA during protein purification of DNA-binding proteins  
190 (Figure 5A). This suggests that the FlhD<sub>SE</sub>/FlhC<sub>EC</sub> complex is less stable, resulting on a  
191 lower yield of complex retrieval.

192

193 We next used the EMSA assays to test all four protein complexes for their ability to bind  
194 the *S. enterica* P<sub>figAB</sub> promoter region. Quantification of the DNA shifts showed that  
195 complexes containing the orthologous FlhC<sub>EC</sub> reduced the P<sub>figAB</sub> promoter binding profile,  
196 compared to FlhC<sub>SE</sub> complexes (Figure 5B). This is consistent with FlhC being the DNA  
197 binding subunit of the complex and the variation in FlhD<sub>4</sub>C<sub>2</sub> activated promoter-binding  
198 sites between *S. enterica* and *E. coli* (19). Therefore, these results suggest that FlhC is a  
199 key determinant of DNA binding ability. Furthermore, the reduction in FlhC<sub>EC</sub> motility and  
200 flagellar gene expression in *S. enterica* is a result of the FlhD<sub>SE</sub>/FlhC<sub>EC</sub> complex being  
201 unstable, ultimately reducing the cellular concentration of the FlhD<sub>4</sub>C<sub>2</sub> complex.

202

203 *FlhD<sub>4</sub>C<sub>2EC</sub> responds to proteolytic regulation*

204

205 *S. enterica* and *E. coli* both regulate the FlhD<sub>4</sub>C<sub>2</sub> complex through ClpXP-mediated  
206 proteolytic degradation. Proteolytic degradation of FlhD<sub>4</sub>C<sub>2</sub> plays a fundamental role in  
207 facilitating rapid responses to environmental changes that require motility (20, 21). The  
208 FlhD<sub>4</sub>C<sub>2</sub> complex has a very short half-life of approximately 2-3 minutes (22). Proteolytic  
209 degradation of FlhD and FlhC is regulated in *E. coli* and *S. enterica* by YdiV (23).

210 However, *ydiV* is not expressed under standard laboratory conditions in model *E. coli*  
211 strains, suggesting that ClpXP activity is modulated in a species-specific manner (7).

212

213 Previous work has shown that YdiV delivers FlhD<sub>4</sub>C<sub>2</sub> complexes to ClpXP for degradation  
214 (24). We have assessed the impact on motility for  $\Delta clpP$  and  $\Delta ydiV$  mutations (Figure 3).

215 The  $\Delta clpP$  and  $\Delta ydiV$  mutants exhibited improved motility and flagellar gene expression,  
216 including the FlhD<sub>SE</sub>/FlhC<sub>EC</sub> strain (Figure 3A and B). These results suggest that

217 proteolytic degradation mechanism of FlhD and FlhC, and its regulation, is common to *E.*  
218 *coli* and *S. enterica*.

219

220 To complement the motility assays, we investigated how  $\Delta clpP$  and  $\Delta ydiV$  mutations  
221 impact the number of FliM-foci in cell. Both  $\Delta clpP$  and  $\Delta ydiV$  mutants showed an increased  
222 number of FliM-foci compared to the wild type (Figure 6 A-C). For *flhC*<sub>EC</sub> strain, FliM-foci  
223 were observed in 13% of the population where individual cells exhibited just one or two  
224 foci. However, the  $\Delta clpP$  or  $\Delta ydiV$  mutants increased the flagellated population of the  
225 *flhC*<sub>EC</sub> strains to 51 and 46 % respectively, albeit with the majority still possessing only a  
226 single FliM focus (Figure 6 B and C).

227

228 *FliT and FliZ regulation of FlhD<sub>4</sub>C<sub>2</sub> complexes*

229

230 FlhD<sub>4</sub>C<sub>2</sub> activity has an additional level of regulation in *S. enterica* via the flagellar-specific  
231 regulators FliT and FliZ. FliT functions as an export chaperone for the filament cap protein,  
232 FliD, and is a regulator of FlhD<sub>4</sub>C<sub>2</sub> activity (17, 25). FliT disrupts the FlhD<sub>4</sub>C<sub>2</sub> complex but  
233 is unable to disrupt a FlhD<sub>4</sub>C<sub>2</sub>:DNA complex. Therefore, FliT modulates availability of  
234 FlhD<sub>4</sub>C<sub>2</sub> complexes for promoter binding (17). In contrast, FliZ is a negative regulator of

235 *ydiV* expression and thus increases the number of FlhD<sub>4</sub>C<sub>2</sub> complexes in *S. enterica* (26,  
236 27).

237

238 In motility assays of  $\Delta fliT$  mutants, we observed a difference between our different *flhDC*  
239 strains. Motility is increased in a  $\Delta fliT$  mutant background in *S. enterica* ((28) and Figure  
240 3A). However, when *flhDC*<sub>EC</sub> and *flhD*<sub>EC</sub> replaced the native genes, a reduced swarm size  
241 was observed (Figure 3A). Furthermore, quantification of P<sub>fliC</sub> activity agreed with the  
242 motility profile for  $\Delta fliT$  mutants, where *flhDC*<sub>EC</sub> and *flhD*<sub>EC</sub> containing strains had reduced  
243 promoter activity compared to wild type (Figure 3B). This suggests that the FlhD<sub>4</sub>C<sub>2</sub>  
244 complexes are being perceived differently by FliT in *S. enterica*. The results for  $\Delta clpP$  and  
245  $\Delta ydiV$  mutants suggests that this is not due to protein stability, as all complex  
246 combinations reacted in a comparable fashion (Figure 3).

247

248 In contrast, the loss of *fliZ* resulted in a consistent reduction in motility, except for the  
249 *flhC*<sub>EC</sub> strain. However, as the *flhC*<sub>EC</sub> strain was already impaired in motility, it is possible  
250 that the resolution of the motility assay was unable to identify differences in  $\Delta fliZ$  mutant.  
251 Flagellar gene expression activity did, however, suggest a 2-fold drop in P<sub>fliC</sub> expression in  
252 the *flhC*<sub>EC</sub>  $\Delta fliZ$  strain as compared to the otherwise wild-type (Figure 3B).

253

254 Analysis of FliM-foci distribution in  $\Delta fliT$  mutant reinforced the observed discrimination of  
255 *flhDC*<sub>EC</sub> and *flhD*<sub>EC</sub> gene replacements. Calculating the average foci per cell, *S. enterica*  
256  $\Delta fliT$  mutants showed an increased average number of foci per cell from 2.9 to 6.3, while  
257 the *flhD*<sub>EC</sub> (*fliT*<sup>+</sup>: 3.4 versus  $\Delta fliT$ : 4.2) and *flhDC*<sub>EC</sub> replacements (*fliT*<sup>+</sup>: 3.6 versus  $\Delta fliT$ :  
258 2.7) exhibited no significant changes (Figure 7A). Interestingly, in a  $\Delta fliZ$  mutant  
259 background, the FliM-foci analysis was able to differentiate *flhDC*<sub>EC</sub> and *flhD*<sub>EC</sub> from the

260 native *S. enterica flhDC* strain. Both replacements exhibited an increase in the average  
261 foci compared to *S. enterica*  $\Delta fliZ$  (Figure 7A).

262

263 These data suggest that there is a fundamental difference in how the FlhD<sub>4</sub>C<sub>2</sub> complexes  
264 in *E. coli* and *S. enterica* respond to, at least, FliT regulation. There are two explanations  
265 for this: a) the *E. coli* combinations are being regulated via an unidentified mechanism in  
266 *S. enterica* or b) that they are insensitive to FliT regulation. Both arguments predict that in  
267 *E. coli* FlhD<sub>4</sub>C<sub>2</sub> may respond differently to FliT regulation. Comparing *S. enterica* and *E.*  
268 *coli* does indeed identify a difference in the response to a  $\Delta fliT$  mutant. While a  $\Delta fliT$   
269 mutant in *S. enterica* leads to a consistent increase in FliM-foci, no significant difference is  
270 noted for an *E. coli*  $\Delta fliT$  mutant compared to *E. coli* wild type (Figure 7B). This suggests  
271 that the regulatory impact of FliT is very different in these two flagellar systems and the  
272 role FliT plays in *S. enterica* is potentially adaptive and species specific.

273

274

275 **DISCUSSION**

276

277 Two model flagellar systems that form the foundation of the flagellar field are those from  
278 the enteric species *E. coli* and *S. enterica*. These two systems have led to key discoveries  
279 in relation to many aspects of flagellar structure, type 3 secretion, flagellar cell biology and  
280 the regulation of flagellar assembly. Textbook explanations suggest that most flagellar  
281 systems are being activated, regulated and built according to the models for *E. coli* and *S.*  
282 *enterica*. Modifications of transcriptional regulatory circuits contribute to the phenotypic  
283 diversity we see in closely related gene sets and we are only now able to investigate this in  
284 depth due to the tools available. Here we have taken a simple step and asked how do  
285 orthologous FlhD<sub>4</sub>C<sub>2</sub> complexes function in the closely related species *E. coli* and *S.*  
286 *enterica*?

287

288 At the onset of our work it was known that FlhD<sub>4</sub>C<sub>2</sub> from *E. coli* could sustain motility in *S.*  
289 *enterica*(11). Our work was focussed on understanding and defining the species-specific  
290 differences in the regulon of two orthologous genes. Here we took advantage of the well-  
291 defined flagellar assembly tools to measure outputs such as, motility, flagellar assembly  
292 per cell and flagellar gene expression. Bioinformatic analysis identifies only an 8 and 6%  
293 identity difference between FlhD and FlhC in *E. coli* and *S. enterica* respectively,  
294 suggesting that these proteins function in an analogous fashion. It is well established that  
295 related taxa usually rely on orthologous regulators to coordinate response to a given signal  
296 (10).

297

298 The fine detail of the differences in the FlhD<sub>4</sub>C<sub>2</sub> complexes only became apparent when  
299 we began to focus on their effect on flagellar gene expression and flagellar assembly. In  
300 all of our assays FlhD<sub>4</sub>C<sub>2</sub><sup>EC</sup> exhibited a reduction in flagellar gene expression compared to

301 FlhD<sub>4</sub>C<sub>2SE</sub>. Biochemical analysis of isolated complexes showed that FlhC<sub>EC</sub> had weaker  
302 DNA binding ability to the P<sub>flgAB</sub> promoter region from *S. enterica*, consistent with previous  
303 investigations into FlhD<sub>4</sub>C<sub>2</sub> DNA binding activity (19). The isolation of FlhD<sub>4</sub>C<sub>2</sub> complexes  
304 from our strains suggested that a key aspect of the phenotypes we observed, was the  
305 stability of the complexes formed.

306  
307 With respect to *flhDC* transcription we show a discrepancy in flagellar numbers defined by  
308 FliM-foci when using P<sub>tetA</sub>/P<sub>tetR</sub>::*flhDC* expression. This was somewhat surprising as all  
309 constructs exhibited good swarming ability on motility agar plates (Figure S2). Original  
310 studies on the regulation of P<sub>tetA</sub>/P<sub>tetR</sub> from Tn10 have shown that these two promoters  
311 have differing activities but both respond to TetR regulation. We show that even though  
312 maximal activity of P<sub>flgA</sub> and P<sub>flhC</sub> can reach 40-50% of P<sub>tetA</sub>::*flhDC* expression for P<sub>tetR</sub>  
313 strains, this results in an average of 2 flagella per cell. This suggests that even though the  
314 majority of the literature states that *E. coli* and *S. enterica* produce between 4 and 8  
315 flagella per cell, only 1 or 2 per cell is needed for an optimal output of the system with  
316 respect to motility agar assays.

317  
318 It has been shown that FliT interacts with FlhC and that in *S. enterica* the output of this  
319 circuit is to destabilize FlhD<sub>4</sub>C<sub>2</sub> complexes that are not bound to DNA. Our data suggests  
320 that this level of regulation does not impact *E. coli* FlhC. The nature of the adaptability  
321 needed by the favourable conditions to drive motility in *E. coli* may have led to the FliT  
322 regulatory input becoming less critical. Similarly, the impact of FliZ regulation becomes  
323 apparent for FlhD<sub>EC</sub> containing complexes when we assess flagellar numbers. FliZ  
324 regulates the transcription of *ydiV* in *S. enterica* (27). It is plausible that the impact in  
325 changing *ydiV* regulation is the source of this differentiation, especially as YdiV is  
326 proposed to interact with FlhD<sub>SE</sub>. Furthermore, we know that *ydiV* is not expressed in

327 model *E. coli* strains, strengthening the argument that FlhD<sub>EC</sub> has adapted to the absence  
328 of YdiV or vice versa FlhD<sub>SE</sub> to YdiV.

329

330 Importantly our analysis shows that even though these two systems are genetically similar,  
331 investigation of FlhD<sub>4C2</sub> activity identifies subtle but key differences into how the FlhD<sub>4C2</sub>  
332 complex is modulated in two closely related species. We argue that this is a valid example  
333 of the caution needed in the age of synthetic biology to exploit heterologous systems in  
334 alternative species or chassis'. Our data shows that even systems showing significant  
335 synteny may not behave in exactly the same manner and due diligence is required in  
336 making assumptions based on heterologous expression.

337

## 338 MATERIALS AND METHODS

339

### 340 *Bacterial Strains and Growth conditions*

341

342 *S. enterica* and *E. coli* strains used in this study have been previously described elsewhere  
343 (12, 15, 17, 28). This study used *S. enterica* serovar Typhimurium strain LT2 as the  
344 chassis for all experiments. *E. coli* genetic material was derived from MG1655. All strains  
345 were grown at either 30°C or 37°C in Luria Bertani Broth (LB) either on 1.5% agar plates  
346 or shaken in liquid cultures at 160 rpm (17). Antibiotics used in this study have been  
347 described elsewhere (29). Motility assays used motility agar (17) incubated at 37°C for 6 to  
348 8 hours. Motility swarms were quantified using images captured on a standard gel doc  
349 system with a ruler in the field of view and quantified using ImageJ to measure the vertical  
350 and horizontal diameter using the average as the swarm size. All motility assays were  
351 performed in triplicate using single batches of motility agar.

352

### 353 *Genetic Manipulations*

354

355 For the replacement of *flhDC* coding sequences the modified lambda red recombination  
356 system described by Blank et al (2011) was used (30). Deletion of *clpP*, *ydiV*, *fliT* and *fliZ*  
357 was performed using the pKD system described by Datsenko and Wanner (2000) (31).  
358  $P_{tetA} / P_{tetR}$  replacements of the  $P_{flhDC}$  region was also performed using Datsenko and  
359 Wanner (2000) with the template being Tn10dTc (32). For Blank et al (2011) replacement  
360 experiments we used autoclaved chlortetracycline instead of anhydrotetracycline as  
361 described for the preparation of Tetracycline sensitive plates (33). All other gene  
362 replacements were performed as previously described (17). All primers used for these  
363 genetic manipulations are available on request.



364

365 *Quantification of flagellar gene expression*

366

367 Flagellar gene expression assays were performed using the plasmids pRG39::cat ( $P_{fliC}$ )  
368 and pRG52::cat ( $P_{fliA}$ ) (15). Both plasmids were transformed into strains using  
369 electroporation. Gene expression was quantified as described previously and analysis was  
370 based on a minimum of  $n = 3$  repeats for each strain tested (15).

371

372 *Quantification of FliM-GFP foci*

373

374 FliM-GFP foci were quantified using Microbetracker on images captured using a Nikon Ti  
375 inverted microscope using filters and exposure times described previously (14). Strains  
376 were grown to an OD600 of 0.5 to 0.6 and cells immobilised using a 1 % agarose pad  
377 containing 10 % LB (14, 17). For each strain a minimum of 5 fields of view were captured  
378 from 3 independent repeats. This allowed analysis of approximately 1000 - 1500 cells per  
379 strain. For the comparison of FliM foci in *E. coli*  $\Delta fliT$  to *S. enterica*  $\Delta fliT$  shown in Figure  
380 7B the chemostat growth system described by Sim et al (2017) was used. For this  
381 experiment the growth rate of both strains was similar to batch culture in LB at 37°C where  
382 the media used was a MinE base with 0.1% Yeast extract and 0.2% glucose (14, 17).

383

384 *Purification of FlhD<sub>4</sub>C<sub>2</sub> complexes*

385

386 Purification of proteins complexes was based on previously described methods (17). Wild  
387 type FlhD<sub>4</sub>C<sub>2</sub><sup>SE</sup> was purified using pPA158. The other 3 complexes were purified from  
388 plasmids generated using the New England Biolabs NEBuilder DNA Assembly kit on the  
389 backbone of pPA158. The *E. coli* strain BL21 was used for all protein induction

390 experiments prior to protein purification using either a pre-equilibrated 5ml His-trap column  
391 or a 5ml heparin column (GE Healthcare). Proteins were visualised using Tricine-based  
392 SDS polyacrylamide gel electrophoresis and standard commassie blue staining (17).

393

394 *Electrophoretic mobility shift assay (EMSA).*

395

396 All EMSA assays were performed using Ni<sup>++</sup> (his-trap) purified proteins as this allowed  
397 analysis of all four complexes (Figure 5A). Buffer exchange from elution buffer to a 100mM  
398 Tris-HCl, 300 mM NaCl 1mM DTT (pH 7.9) buffer was performed through 10 cycles of  
399 protein concentration in VivaSpin columns with 20 ml buffer reduced to 5 ml per round of  
400 centrifugation at 4500 rpm. A protein concentration range of 100 to 700 nM was used with  
401 80 ng / ml of a PCR product containing P<sub>flgAB</sub> from *S. enterica*. After incubation bound and  
402 unbound DNA were resolved using 5% acrylamide gels made with 1x TBE buffer.

403 Quantification of gel images was performed using ImageJ.

404

405 **ACKNOWLEDGEMENTS**

406

407 PDA would like to recognize the internal financial support of ICAMB during this study. The  
408 stipend and research costs for the PhD of AA was provided by The Ministry of Higher  
409 Education and Scientific Research (Iraq). We would like to thank the financial support of  
410 the Newcastle University Faculty of Medicine for providing the John William Luccok and  
411 Ernest Jeffcock Research PhD Studentship to MS for this study. PAH would like to  
412 acknowledge the support of iUK/BBSRC (grant: BB/N023544/1), NERC (grant:  
413 NE/M001415/1), the University of Strathclyde and the Microbiology Society for funding. We  
414 would also like to thank all lab members for feedback on the project during the  
415 experimental and writing phases.

416

## 417 REFERENCES

418

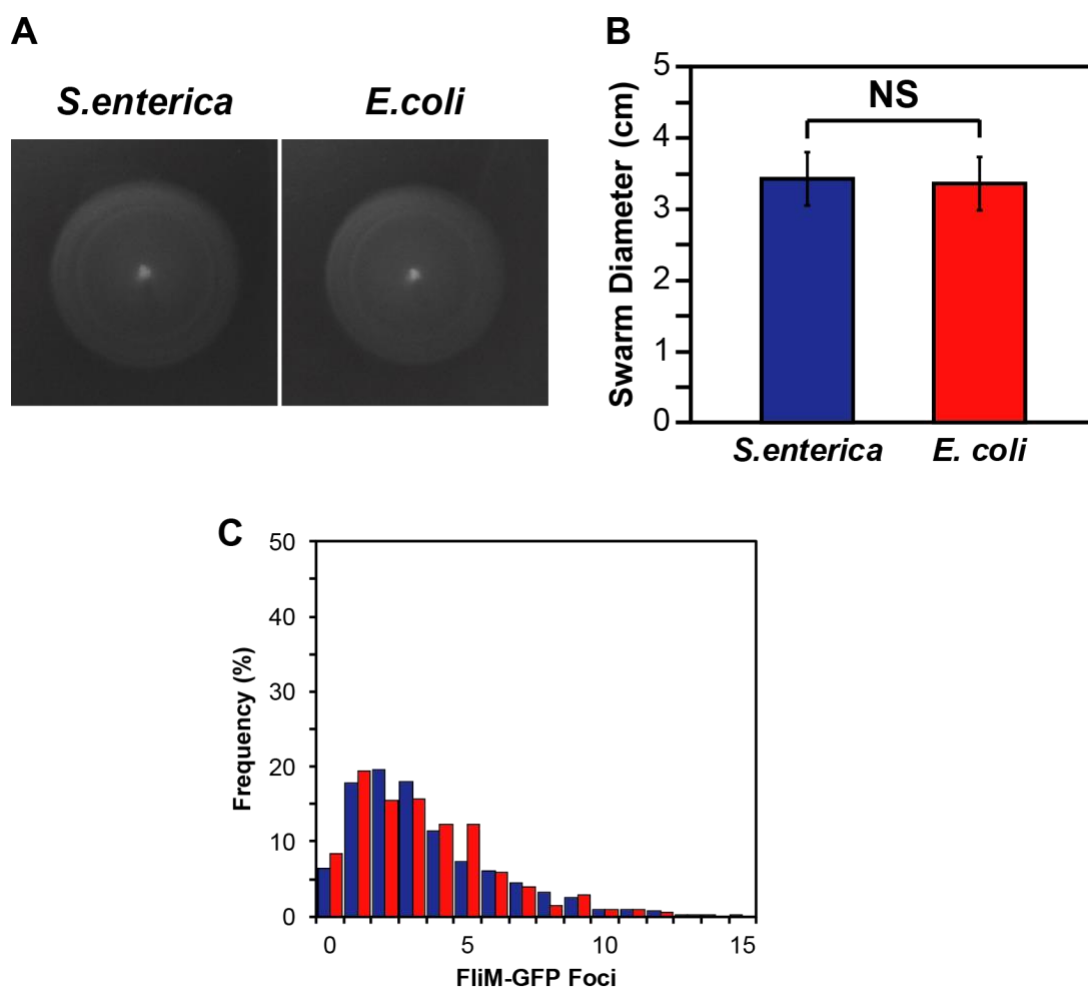
- 419 1. **Duan Q, Zhou M, Zhu L, Zhu G.** 2013. Flagella and bacterial pathogenicity. *J Basic*  
420 *Microbiol* **53**:1–8.
- 421 2. **Minamino T, Imada K, Namba K.** 2008. Mechanisms of type III protein export for  
422 bacterial flagellar assembly. *Mol Biosyst* **4**:1105–1115.
- 423 3. **Chevance FFV, Hughes KT.** 2008. Coordinating assembly of a bacterial  
424 macromolecular machine. *Nat Rev Micro* **6**:455–465.
- 425 4. **Aldridge P, Hughes KT.** 2002. Regulation of flagellar assembly. *Curr Opin*  
426 *Microbiol* **5**:160–165.
- 427 5. **Chilcott GS, Hughes KT.** 2000. Coupling of flagellar gene expression to flagellar  
428 assembly in *Salmonella enterica* serovar typhimurium and *Escherichia coli*. *Microbiol*  
429 *Mol Biol Rev* **64**:694–708.
- 430 6. **Minamino T, Namba K.** 2004. Self-assembly and type III protein export of the  
431 bacterial flagellum. *J Mol Microbiol Biotechnol* **7**:5–17.
- 432 7. **Wada T, Hatamoto Y, Kutsukake K.** 2012. Functional and expressional analyses of  
433 the anti-FlhD4C2 factor gene *ydiV* in *Escherichia coli*. *Microbiology* **158**:1533–1542.
- 434 8. **Soutourina OA, Bertin PN.** 2003. Regulation cascade of flagellar expression in  
435 Gram-negative bacteria. *FEMS Microbiol Rev* **27**:505–523.
- 436 9. **Moulim C, Hughes KT.** 2014. The effect of cell growth phase on the regulatory  
437 cross-talk between flagellar and *Spi1* virulence gene expression. *PLoS Pathog*  
438 **10**:e1003987.
- 439 10. **Perez JC, Groisman EA.** 2009. Evolution of transcriptional regulatory circuits in  
440 bacteria. *Cell* **138**:233–244.
- 441 11. **Kutsukake K, Iino T, Komeda Y, Yamaguchi S.** 1980. Functional homology of *fla*  
442 genes between *Salmonella typhimurium* and *Escherichia coli*. *Mol Gen Genet*  
443 **178**:59–67.
- 444 12. **Aldridge P, Karlinsey JE, Becker E, Chevance FFV, Hughes KT.** 2006. *FliK*  
445 prevents premature secretion of the anti-sigma factor *FlgM* into the periplasm. *Mol*  
446 *Microbiol* **60**:630–643.
- 447 13. **Delalez NJ, Wadhams GH, Rosser G, Xue Q, Brown MT, Dobbie IM, Berry RM,**  
448 **Leake MC, Armitage JP.** 2010. Signal-dependent turnover of the bacterial flagellar  
449 switch protein *FliM*. *Proceedings of the National Academy of Sciences* **107**:11347–  
450 11351.
- 451 14. **Sim M, Koirala S, Picton D, Strahl H, Hoskisson PA, Rao CV, Gillespie CS,**  
452 **Aldridge PD.** 2017. Growth rate control of flagellar assembly in *Escherichia coli*  
453 strain RP437. *Sci Rep* **7**:41189.

- 454 15. **Brown JD, Saini S, Aldridge C, Herbert J, Rao CV, Aldridge PD.** 2008. The rate  
455 of protein secretion dictates the temporal dynamics of flagellar gene expression. *Mol*  
456 *Microbiol* **70**:924–937.
- 457 16. **Bertrand KP, Postle K, Wray LV, Reznikoff WS.** 1984. Construction of a single-  
458 copy promoter vector and its use in analysis of regulation of the transposon Tn10  
459 tetracycline resistance determinant. *J Bacteriol* **158**:910–919.
- 460 17. **Aldridge C, Poonchareon K, Saini S, Ewen T, Soloyva A, Rao CV, Imada K,**  
461 **Minamino T, Aldridge PD.** 2010. The interaction dynamics of a negative feedback  
462 loop regulates flagellar number in *Salmonella enterica* serovar Typhimurium. *Mol*  
463 *Microbiol* **78**:1416–1430.
- 464 18. **Wang S, Fleming RT, Westbrook EM, Matsumura P, McKay DB.** 2006. Structure  
465 of the *Escherichia coli* FlhDC complex, a prokaryotic heteromeric regulator of  
466 transcription. *Journal of Molecular Biology* **355**:798–808.
- 467 19. **Stafford GP, Ogi T, Hughes C.** 2005. Binding and transcriptional activation of non-  
468 flagellar genes by the *Escherichia coli* flagellar master regulator FlhD2C2.  
469 *Microbiology (Reading, Engl)* **151**:1779–1788.
- 470 20. **Kitagawa R, Takaya A, Yamamoto T.** 2011. Dual regulatory pathways of flagellar  
471 gene expression by ClpXP protease in enterohaemorrhagic *Escherichia coli*.  
472 *Microbiology* **157**:3094–3103.
- 473 21. **Tomoyasu T, Ohkishi T, Ukyo Y, Tokumitsu A, Takaya A, Suzuki M, Sekiya K,**  
474 **Matsui H, Kutsukake K, Yamamoto T.** 2002. The ClpXP ATP-dependent protease  
475 regulates flagellum synthesis in *Salmonella enterica* serovar typhimurium. *J*  
476 *Bacteriol* **184**:645–653.
- 477 22. **Claret L, Hughes C.** 2000. Rapid Turnover of FlhD and FlhC, the Flagellar Regulon  
478 Transcriptional Activator Proteins, during *Proteus* Swarming. *J Bacteriol* **182**:833–  
479 836.
- 480 23. **Wada T, Morizane T, Abo T, Tominaga A, Inoue-Tanaka K, Kutsukake K.** 2011.  
481 EAL domain protein YdiV acts as an anti-FlhD4C2 factor responsible for nutritional  
482 control of the flagellar regulon in *Salmonella enterica* Serovar Typhimurium. *J*  
483 *Bacteriol* **193**:1600–1611.
- 484 24. **Takaya A, Erhardt M, Karata K, Winterberg K, Yamamoto T, Hughes KT.** 2012.  
485 YdiV: a dual function protein that targets FlhDC for ClpXP-dependent degradation by  
486 promoting release of DNA-bound FlhDC complex. *Mol Microbiol* **83**:1268–1284.
- 487 25. **Bennett JC, Thomas J, Fraser GM, Hughes C.** 2001. Substrate complexes and  
488 domain organization of the *Salmonella* flagellar export chaperones FlgN and FliT.  
489 *Mol Microbiol* **39**:781–791.
- 490 26. **Saini S, Brown JD, Aldridge PD, Rao CV.** 2008. FliZ Is a posttranslational activator  
491 of FlhD4C2-dependent flagellar gene expression. *J Bacteriol* **190**:4979–4988.
- 492 27. **Wada T, Tanabe Y, Kutsukake K.** 2011. FliZ Acts as a Repressor of the ydiV Gene,  
493 Which Encodes an Anti-FlhD4C2 Factor of the Flagellar Regulon in *Salmonella*  
494 *enterica* Serovar Typhimurium. *J Bacteriol* **193**:5191–5198.

- 495 28. **Aldridge P, Karlinsey J, Hughes KT.** 2003. The type III secretion chaperone FlgN  
496 regulates flagellar assembly via a negative feedback loop containing its chaperone  
497 substrates FlgK and FlgL. *Mol Microbiol* **49**:1333–1345.
- 498 29. **Bonifield HR, Hughes KT.** 2003. Flagellar phase variation in *Salmonella enterica* is  
499 mediated by a posttranscriptional control mechanism. *J Bacteriol* **185**:3567–3574.
- 500 30. **Blank K, Hensel M, Gerlach RG.** 2011. Rapid and highly efficient method for  
501 scarless mutagenesis within the *Salmonella enterica* chromosome. *PLoS ONE*  
502 **6**:e15763.
- 503 31. **Datsenko KA, Wanner BL.** 2000. One-step inactivation of chromosomal genes in  
504 *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA* **97**:6640–6645.
- 505 32. **Rappleye CA, Roth JR.** 1997. A Tn10 derivative (T-POP) for isolation of insertions  
506 with conditional (tetracycline-dependent) phenotypes. *J Bacteriol* **179**:5827–5834.
- 507 33. **Maloy SR, Nunn WD.** 1981. Selection for loss of tetracycline resistance by  
508 *Escherichia coli*. *J Bacteriol* **145**:1110–1111.
- 509

510 **FIGURES and LEGENDS**

511



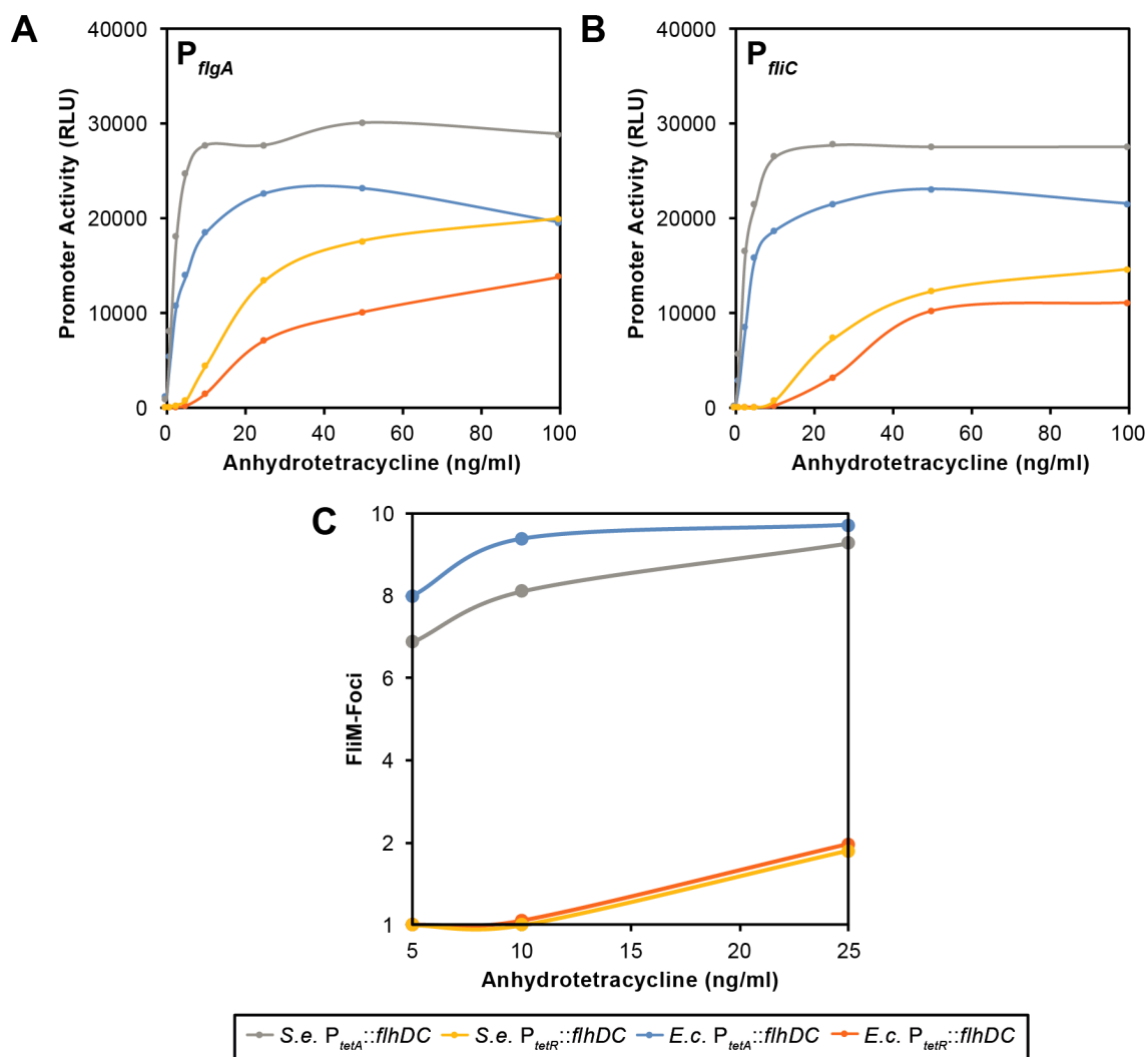
512

513

514 **Figure 1 A.** Motility of *flhDC<sub>ST</sub>* and *flhDC<sub>EC</sub>* driven by  $P_{flhDC}$ . **B.** Quantification of swarms  
515 produced in motility agar after 6 to 8 hours incubation. Error bars indicate calculated  
516 standard deviations. **C.** Percentage frequency of FliM-GFP foci for *flhDC<sub>EC</sub>* compared to *S.*  
517 *enterica* with *flhDC* under the control of  $P_{flhDC}$ . Colors of bars in the graph correspond to  
518 the source of *flhDC* as shown in (B).  
519

520

521



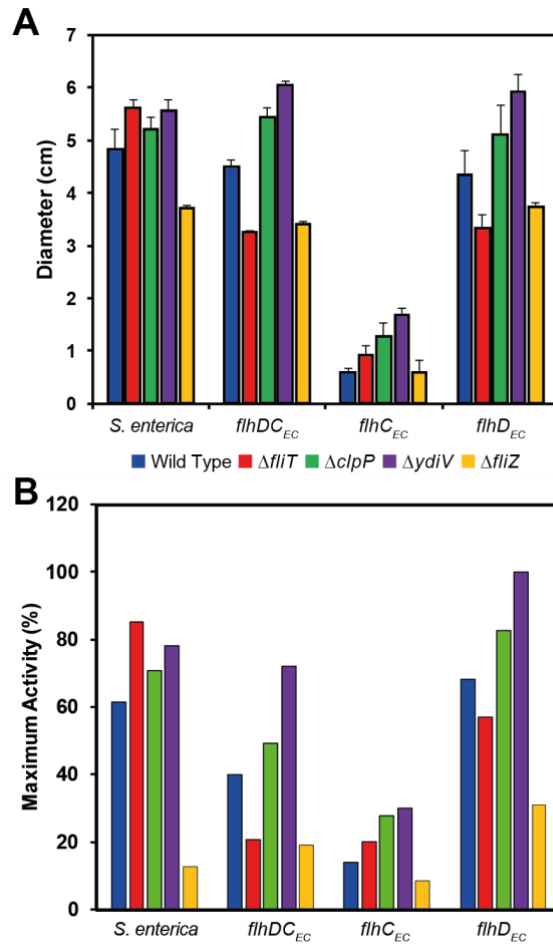
522

523

524 **Figure 2** Titration of  $P_{tetA}::flhDC_{ST/EC}$  and  $P_{tetR}::flhDC_{ST/EC}$  activity suggests a given rate of  
 525 transcription drives optimal flagellar assembly. **A.** Activity of  $P_{flgA}$  in response to  $P_{tetA}$  or  
 526  $P_{tetR}$  transcription of  $flhDC$  from *S. enterica* (S.e.) or *E. coli* (E.c.). **B.** Activity of  $P_{fliC}$  in  
 527 response to  $P_{tetA}$  or  $P_{tetR}$  transcription of  $flhDC$ . **C.** flagellar numbers as defined by FliM-foci  
 528 in response to  $P_{tetA}$  or  $P_{tetR}$  transcription of  $flhDC$ .  
 529



530

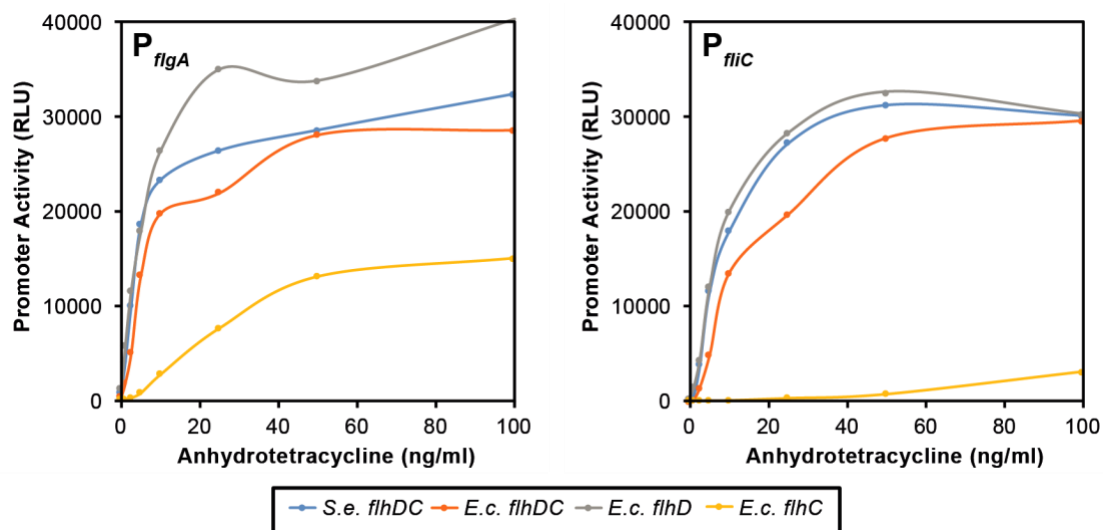


531

532

533 **Figure 3** Motility phenotypes and gene expression of *flhDC<sub>ST</sub>*, *flhDC<sub>EC</sub>*, *flhD<sub>EC</sub>* and *flhC<sub>EC</sub>*  
534 strains in the absence of known FlhD<sub>4C2</sub> regulators. **A.** Quantification of n = 3 swarms per  
535 strain produced in motility agar after 6 to 8 hours incubation at 37°C. Error bars indicate  
536 calculated standard deviations. **B.** Relative activity of P<sub>flhC</sub> in all strains as a percent of the  
537 maximal activity observed in *flhD<sub>EC</sub> ΔydiV*.  
538

539

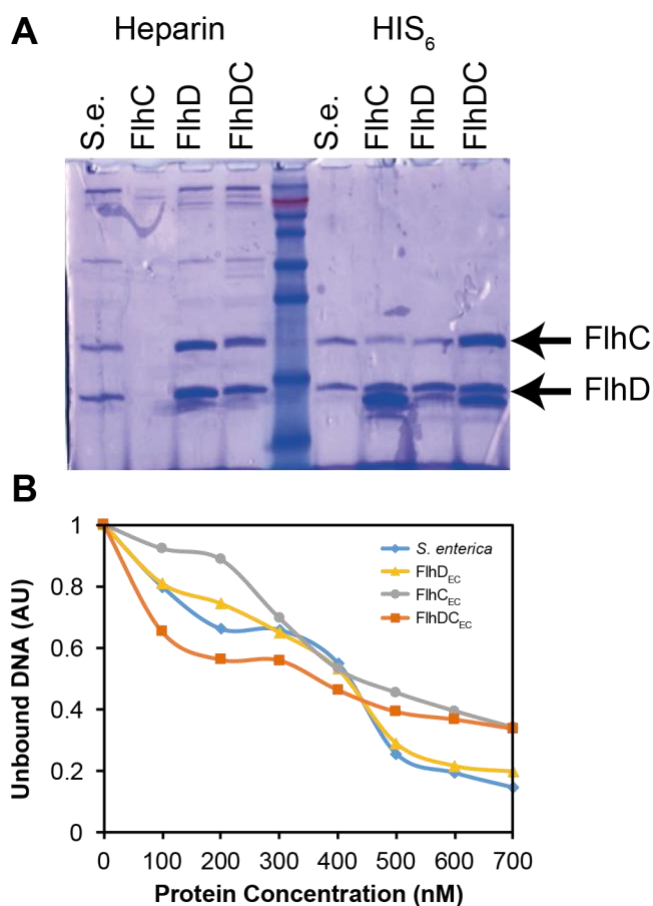


540

541

542 **Figure 4** Titration of  $P_{tetA}::flhDC$  for *S. enterica*,  $flhDC_{EC}$ ,  $flhD_{EC}$  and  $flhC_{EC}$  suggests that  
543  $flhC_{EC}$  exhibits low motility due reduced  $P_{flgA}$  activity and a strong reduction in  $P_{fliC}$  activity.  
544 Inducible expression was driven from the  $P_{tetA}$  promoter within the TetRA cassette of Tn10.  
545 The data shown in both panels is significant using ANOVA statistical analysis  $P < 0.05$ .  
546

547

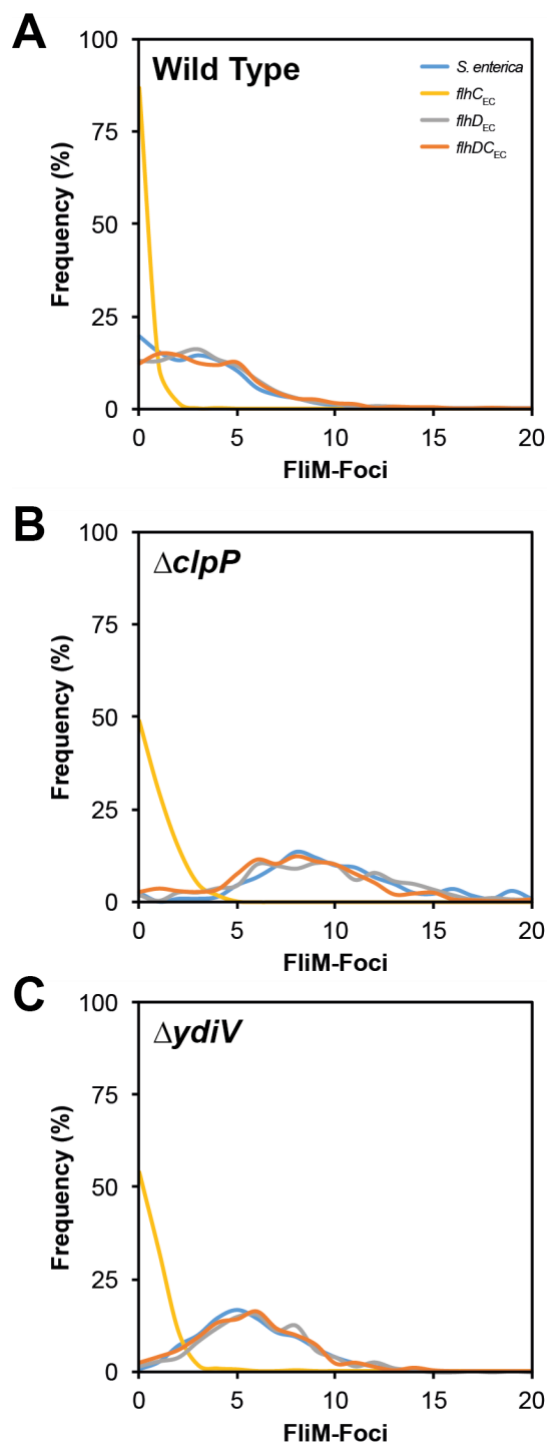


548

549

550 **Figure 5** The FlhD<sub>ST</sub>FlhC<sub>EC</sub> complex is an active but unstable complex. **A.** Protein gel  
551 showing purified complexes with either HIS<sub>6</sub> or Heparin based purification protocols. The  
552 nature of the FlhDC complex allows isolation of both proteins in these assays. Arrows  
553 indicate the FlhC and FlhD bands. **B.** Quantification of EMSA to define the binding ability  
554 of the complex combinations compared to *S. enterica* FlhD<sub>4C2</sub>. Data shows that whenever  
555 FlhC<sub>EC</sub> is present a reduced level of binding to P<sub>flgAB</sub> was observed.  
556

557



558

559

560

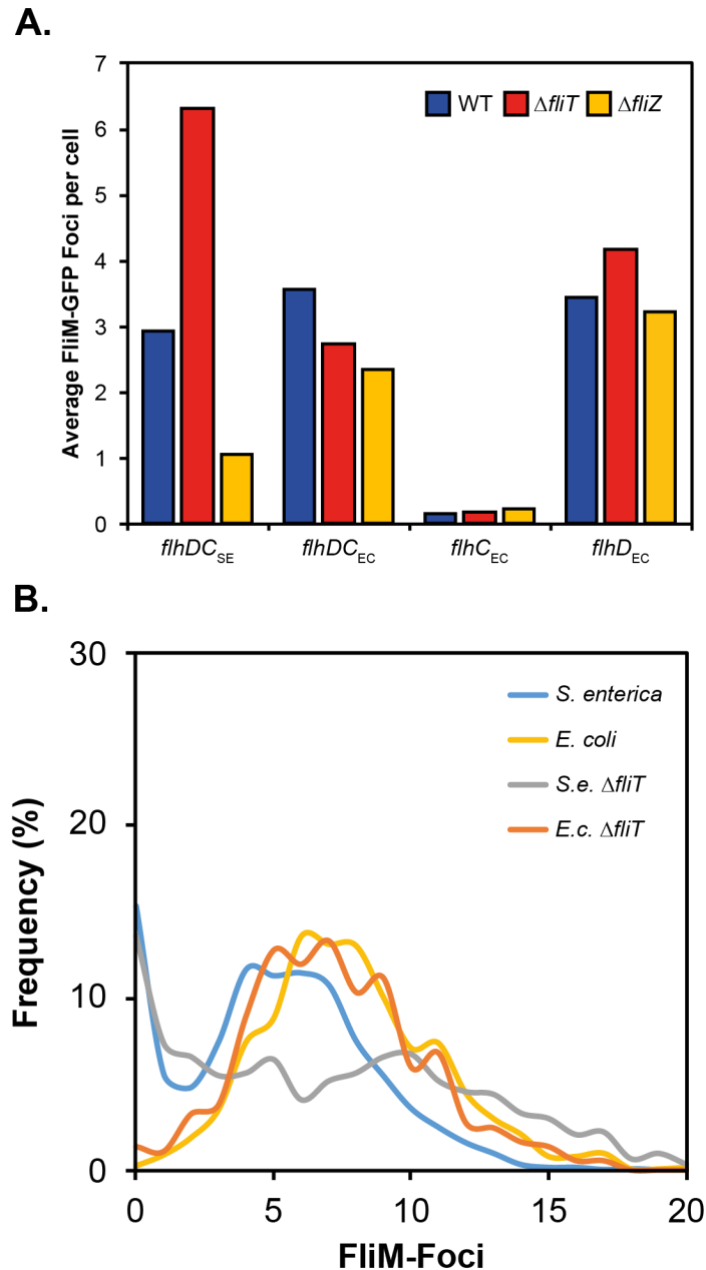
561

562

563

**Figure 6** Impact of protein stability regulators of FlhD<sub>4</sub>C<sub>2</sub> on flagellar numbers as defined by FliM-foci. Quantification of FliM-foci was performed using the semi-automatic protocols defined with in Microbetracker. **A.** Wild Type foci distribution; **B.**  $\Delta clpP$ ; **C.**  $\Delta ydiV$ .

564



565

566

567

568

569

570

571

572

573

574

575

576

**Figure 7** FliT and FliZ regulation reflects when FliC<sub>EC</sub> or FliD<sub>EC</sub> are present. **A.** FliM-Foci quantification is consistent with the observed motility phenotype of  $\Delta fliT$  mutants. For  $\Delta fliZ$  FliM-foci numbers discriminate between the source of FliD, FliD<sub>SE</sub> exhibits a consistent drop in foci while FliD<sub>EC</sub> containing strains show comparable foci averages. **B.** Testing the hypothesis that  $\Delta fliT$  mutants respond differently in *E. coli* compared to *S. enterica*. Note: this experiment in **(B)** uses the species *E. coli* and *S. enterica* not engineered replacements.